ACTA UNIVERSITATIS SZEGEDIENSIS

ACTA SCIENTIARUM MATHEMATICARUM

TOMUS 34

SZEGED, 1973

INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS

A JÓZSEF ATTILA TUDOMÁNYEGYETEM KÖZLEMÉNYEI

ACTA SCIENTIARUM MATHEMATICARUM

34. KÖTET

SZEGED, 1973. JÚNIUS

JÓZSEF ATTILA TUDOMÁNYEGYETEM BOLYAI INTÉZETE

ACTA UNIVERSITATIS SZEGEDIENSIS

ACTA SCIENTIARUM MATHEMATICARUM

TOMUS 34

TOMUM CURAVERUNT

I. KOVÁCS ET K. TANDORI

SZEGED, 1973

INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS

A JÓZSEF ATTILA TUDOMÁNYEGYETEM KÖZLEMÉNYEI

ACTA SCIENTIARUM MATHEMATICARUM

34. KÖTET

E KÖTETET SZERKESZTETTE
KOVÁCS ISTVÁN ÉS TANDORI KÁROLY

SZEGED, 1973. JÚNIUS

JÓZSEF ATTILA TUDOMÁNYEGYETEM BOLYAI INTÉZETE

SZŐKEFALVI-NAGY BÉLA

TOMUS XXXIV — 1973 — 34. KÖTET

On the convergence of function series

By G. ALEXITS in Budapest

Dedicated to Béla Szőkefalvi-Nagy on his 60th birthday

1. Let X be a measurable space with a positive measure μ and $\{f_n(x)\}$ a sequence of μ-measurable functions in X. On the measurable set $E \subset X$, consider the Lebesgue functions of the system $\{f_n(x)\}$:

$$L_n(x) = \int_E \left| \sum_{k=0}^{n} f_k(x) f_k(y) \right| d\mu(y),$$

and for an index sequence $\nu_1 < \nu_2 < \dots$ set

$$L_{\nu_n}(E) = \int_E \max_{0 \leq j \leq n} L_{\nu_j}(x) \, d\mu(x).$$

Recently we have proved the following theorem ([2], Theorem 2):

If E is of finite measure and $L_n(E) \leq K$ $(n=0, 1, \dots)$, further if $\{a_n\}$ is a sequence of real numbers such that $\sum a_n^2 < \infty$, then the series $\sum a_n f_n(x)$ converges on E a.e.

If no more than the uniform boundedness of the subsequence $\{L_{\nu_n}(E)\}$ is required then for the subsequence $\{s_{\nu_n}(x)\}$ of the partial sums

$$s_{\nu_n}(x) = \sum_{k=0}^{\nu_n} a_k f_k(x)$$

a similar statement could be proved only under a rather restrictive subcondition ([2], Theorem 3). But it seems that an analogous statement without any restriction could have a certain importance. In the following we shall prove it by suppressing, besides the mentioned subcondition, also the inutile condition that E should have a finite measure. More exactly, we shall prove the following

Theorem 1. *Let $\{a_n\}$ be an arbitrary sequence of real numbers with $\sum a_n^2 < \infty$ and $\{f_n(x)\}$ an arbitrary sequence of μ-integrable functions defined on the measurable set $E \subset X$. Then the condition $L_{\nu_n}(E) \leq K$ $(n=1, 2, \dots)$ implies the convergence of the sequence $\{s_{\nu_n}(x)\}$ on E a.e.*

Theorem 1 has different consequences of various kind; one of them could open a new way to the study of the convergence properties of certain function series even if the corresponding Lebesgue functions do not form a bounded sequence. One of these consequences concerns the series of weakly multiplicative functions, a notion we introduced occasionally [3] and which is a vigorous generalization of the stochastically independent functions.

Definition. A system $\{\varphi_n(x)\}$ of μ-integrable functions on E is called *weakly multiplicative*, if the integrals $\int_E \varphi_{\nu_1}(x)\varphi_{\nu_2}(x)\dots\varphi_{\nu_n}(x)\,d\mu(x)$ exist for all finite collections of indices $\nu_1<\nu_2<\dots<\nu_n$ and

$$\sum\left|\int_E \varphi_{\nu_1}(x)\varphi_{\nu_2}(x)\dots\varphi_{\nu_n}(x)\,d\mu(x)\right|<\infty,$$

where the summation has to be taken for all finite collections of $\nu_1<\nu_2<\dots<\nu_n$.

We shall prove the convergence a.e. of the series $\sum c_n\varphi_n(x)$ if $\sum c_n^2<\infty$ and $\{\varphi_n(x)\}$ is bounded. (We have already proved this [3] assuming the validity of our present Theorem 1.) Then we shall study also the absolute convergence of such series.

The convergence a.e. of $\sum c_n\varphi_n(x)$ under $\sum c_n^2<\infty$ generalizes a theorem we proved earlier ([1], Theorem 1). Our present result is much stronger than that earlier one; this can be seen by the following fact: FIEDLER and TRAUTNER proved [4] the existence of a complete bounded orthonormal system which does not contain any infinite subsystem of multiplicatively orthogonal functions (to which our earlier theorem refers). Moreover, FRIESS and TRAUTNER [5] proved that the bounded complete orthogonal systems containing an infinite multiplicatively orthogonal subsystem are in some sense "rare". Whereas we shall see that every bounded infinite orthonormal system contains an infinite weakly multiplicative subsystem.

2. Turning to the proof of our Theorem 1 we first prove an inequality which plays a similar role as the Rademacher—Menchov inequality in the theory of orthogonal series.

Let $n\geqq m$ be two fixed positive integers and denote by $m(x)$ and $n(x)$ measurable functions taking only integer values between m and n, i.e. $m\leqq m(x)\leqq n(x)\leqq n$. If $r_n(t)$ denotes the kth Rademacher function defined in $0\leqq t\leqq 1$, i.e. $r_k(t)=$ $=\operatorname{sign}\sin 2^n\pi t$, we can write

$$s_{\nu_{n(x)}}(x)-s_{\nu_{m(x)}}(x)=\int_0^1\sum_{k=\nu_m}^{\nu_n}a_k r_k(t)\sum_{k=\nu_{m(x)}+1}^{\nu_{n(x)}}r_k(t)f_k(x)\,dt.$$

Hence, denoting by P and N the sets on which $s_{\nu_{n(x)}}(x)-s_{\nu_{m(x)}}(x)\geqq 0$ or <0, respectively, we get by Schwarz's inequality

$$I_{m,n}(P)=\int_P [s_{\nu_{n(x)}}(x)-s_{\nu_{m(x)}}(x)]\,d\mu(x)\leqq$$

$$\leqq\left\{\int_0^1\left[\sum_{k=\nu_m}^{\nu_n} a_k r_k(t)\right]^2 dt\int_0^1\left[\int_P \sum_{k=\nu_{m(x)}+1}^{\nu_{n(x)}} r_k(t) f_k(x)\,d\mu(x)\right]^2 dt\right\}^{\frac{1}{2}}=$$

$$=\left\{\sum_{k=\nu_m}^{\nu_n} a_k^2 \int_P\int_P\int_0^1 \sum_{k=\nu_{m(x)}+1}^{\nu_{n(x)}} r_k(t) f_k(x) \sum_{k=\nu_{m(y)}+1}^{\nu_{n(y)}} r_k(t) f_k(y)\,dt\,d\mu(x)\,d\mu(y)\right\}^{\frac{1}{2}}=$$

$$=\left\{\sum_{k=\nu_m}^{\nu_n} a_k^2 \int_P\int_P \sum_{k=\nu_{m(x,y)}+1}^{\nu_{n(x,y)}} f_k(x) f_k(y)\,d\mu(x)\,d\mu(y)\right\}^{\frac{1}{2}},$$

where $\nu_{m(x,y)}=\max\{\nu_{m(x)},\nu_{m(y)}\}$ and $\nu_{n(x,y)}=\min\{\nu_{n(x)},\nu_{n(y)}\}$. Write the sum in the last integral in the following form:

$$\sum_{k=\nu_{m(x,y)}+1}^{\nu_{n(x,y)}}=\sum_{k=0}^{\nu_{n(x,y)}}-\sum_{k=0}^{\nu_{m(x,y)}}.$$

Then we get by definition of $L_{\nu_n}(x)$ and $L_{\nu_n}(E)$

$$I_{m,n}(P)\leqq\left\{\sum_{k=\nu_m}^{\nu_n} a_k^2\int_P\int_P\left(\left|\sum_{k=0}^{\nu_{n(x,y)}} f_k(x)f_k(y)\right|+\left|\sum_{k=0}^{\nu_{m(x,y)}} f_k(x)f_k(y)\right|\right)d\mu(x)\,d\mu(y)\right\}^{\frac{1}{2}}\leqq$$

$$\leqq\left\{\sum_{k=\nu_m}^{\nu_n} a_k^2\left(2\int_P\int_P\left|\sum_{k=0}^{\nu_{n(x)}} f_k(x)f_k(y)\right|d\mu(x)\,d\mu(y)+\right.\right.$$

$$\left.\left.+2\int_P\int_P\left|\sum_{k=0}^{\nu_{m(x)}} f_k(x)f_k(y)\right|d\mu(x)\,d\mu(y)\right)\right\}^{\frac{1}{2}}\leqq$$

$$\leqq\left\{\sum_{k=\nu_m}^{\nu_n} a_k^2\left(2\int_P L_{\nu_{n(x)}}(x)\,d\mu(x)+2\int_P L_{\nu_{m(x)}}(x)\,d\mu(x)\right)\right\}^{\frac{1}{2}}\leqq\left\{4\sum_{k=\nu_m}^{\nu_n} a_k^2\cdot L_{\nu_n}(E)\right\}^{\frac{1}{2}}.$$

The same estimate holds true for the integral of $s_{\nu_{m(x)}}(x)-s_{\nu_{n(x)}}(x)$ extended over the set N, so we get finally

$$\int_E |s_{\nu_{n(x)}}(x)-s_{\nu_{m(x)}}(x)|\,d\mu(x)\leqq\left\{16L_{\nu_n}(E)\sum_{k=\nu_m}^{\nu_n} a_k^2\right\}^{\frac{1}{2}}. \tag{1}$$

This is the inequality we intended to prove.

3. From (1) the proof of Theorem 1 follows. Indeed, choose for $m(x)$ the least integer $\geqq m$ and for $n(x)$ the largest integer $\leqq n$ such that

$$|s_{\nu_{n(x)}}(x)-s_{\nu_{m(x)}}(x)|=\max_{m\leqq i\leqq j\leqq n}|s_{\nu_j}(x)-s_{\nu_i}(x)|.$$

Denote by $A_{m,n}$ the set on which

$$|s_{\nu_{n(x)}}(x) - s_{\nu_{m(x)}}(x)| \geqq \varepsilon \qquad (\varepsilon > 0).$$

From (1) and the inequality

$$\varepsilon |A_{m,n}| \leqq \int_E |s_{\nu_{n(x)}}(x) - s_{\nu_{m(x)}}(x)| \, d\mu(x)$$

one gets the estimate

$$|A_{m,n}| \leqq \varepsilon^{-1} \left\{ 16 L_{\nu_n}(E) \sum_{k=\nu_m}^{\nu_n} a_k^2 \right\}^{\frac{1}{2}},$$

where $|A_{m,n}|$ denotes the μ-measure of $A_{m,n}$. Since $\sum a_k^2 < \infty$, for every $\varepsilon > 0$ there exists an index m_ε such that

$$\text{(2)} \qquad \sum_{k=\nu_m}^{\nu_n} a_k^2 < \frac{\varepsilon^4}{16K} \qquad (m \geqq m_\varepsilon),$$

where K is the common bound of the numbers $L_{\nu_n}(E)$. Hence $|A_{m,n}| < \varepsilon$ for every $m \geqq m_\varepsilon$ and $n \geqq m$. From the definition of $m(x)$ and $n(x)$ it follows that, for m fixed, the sequence $\{|s_{\nu_{n(x)}}(x) - s_{\nu_{m(x)}}(x)|\}$ is not decreasing if $n \to \infty$. Then, for m fixed, the sequence of sets $\{A_{m,n}\}$ is also not decreasing. Therefore the set

$$A(m) = \lim_{n \to \infty} A_{m,n}$$

exists and has measure $|A(m)| \leqq \varepsilon$ for an arbitrary $m \geqq m_\varepsilon$.

Put $m_1 > m$, then

$$|s_{\nu_{n(x)}}(x) - s_{\nu_{m(x)}}(x)| \geqq |s_{\nu_{n(x)}}(x) - s_{\nu_{m_1(x)}}(x)|,$$

hence $A(m_1) \subset A(m)$. Or if $x \notin A(m)$ we have

$$|s_{\nu_n}(x) - s_{\nu_m}(x)| \leqq |s_{\nu_{n(x)}}(x) - s_{\nu_{m(x)}}(x)| < \varepsilon$$

for an arbitrary $n \geqq m$. So we got finally the estimate

$$\text{(3)} \qquad |s_{\nu_n}(x) - s_{\nu_m}(x)| < \varepsilon$$

for every $m \geqq m_\varepsilon$ and an arbitrary $n \geqq m$, provided $x \notin A(m_\varepsilon)$. The measure of $A(m_\varepsilon)$ being $\leqq \varepsilon$, the inequality (3) holds true except the points of a set of measure $\leqq \varepsilon$.

Repeating the same order of ideas with $\varepsilon/2$, $\varepsilon/4$, ... instead of ε, we obtain a sequence of sets $A(m_{\varepsilon/2})$, $A(m_{\varepsilon/4})$, ... with measures $\leqq \varepsilon/2$, $\leqq \varepsilon/4$, ... on the complements of which (3) holds true with $\varepsilon/2$, $\varepsilon/4$, ... instead of ε. Form the set

$$A = \bigcup_{k=0}^{\infty} A(m_{\varepsilon/2^k}),$$

then $|A| \leqq 2\varepsilon$ and, for $x \notin A$, we have

$$|s_{\nu_n}(x) - s_{\nu_m}(x)| < \frac{\varepsilon}{2^k} \qquad (k = 0, 1, \dots)$$

for every $n \geqq m \geqq m_{\varepsilon/2^k}$. This means that $\{s_{\nu_n}(x)\}$ converges except perhaps on the set A of measure $\leqq 2\varepsilon$ and the proof is complete.

4. We say that the function system $\{\varphi_n(x)\}$ can be extended to a $L_{\nu_n}(E)$-bounded system $\{f_n(x)\}$, if $f_{\nu_n+1}(x) = \varphi_n(x)$ and the system $\{f_n(x)\}$ has the property $L_{\nu_n}(E) \leqq K$ $(n = 1, 2, \dots)$. From Theorem 1 we deduce immediately the following

Corollary. *If a system $\{\varphi_n(x)\}$ can be extended to a $L_{\nu_n}(E)$-bounded system $\{f_n(x)\}$, then the series $\sum c_n \varphi_n(x)$ converges on E a.e. under the sole condition $\sum c_n^2 < \infty$.*

Indeed, if we set $a_k = c_n$ for $k = \nu_n + 1$ $(n = 1, 2, \dots)$ and $a_k = 0$ for every other k, then we have

$$\sum_{k=0}^{\nu_n+1} a_k f_k(x) = \sum_{k=0}^{n} c_k \varphi_k(x)$$

and the corollary follows from Theorem 1.

We would like to emphasize that this corollary contains eventually a possible way for the study of the convergence properties of different series $\sum c_n \varphi_n(x)$. Considering namely the circumstance that we do not need more than the μ-integrability of the functions $f_k(x)$, it might be possible that, by a suitable choice of the indices ν_n and the functions $f_k(x)$ which we insert between $\varphi_n(x)$ and $\varphi_{n+1}(x)$, one could extend different systems $\{\varphi_n(x)\}$ to a $L_{\nu_n}(E)$-bounded system $\{f_n(x)\}$, and so conclude the convergence a.e. of $\sum c_n \varphi_n(x)$ if $\sum c_n^2 < \infty$. It would be very interesting if one could apply this method to some classical orthogonal system.

5. We defined in Sec. 1 the notion of a weekly multiplicative system $\{\varphi_n(x)\}$. For such systems we can apply the above sketched method to prove the following

Theorem 2. *If $\{\varphi_n(x)\}$ is weakly multiplicative on the set $E \subset X$ of finite measure, further if $|\varphi_n(x)| \leqq M_n$ with $M_n \geqq 1$, then the condition $\sum c_n^2 M_n^2 < \infty$ implies the convergence of the series $\sum c_n \varphi_n(x)$ on E a.e.*

Denote by $\{\psi_n(x)\}$ the product system of $\{\varphi_n(x)/M_n\}$, i.e. $\psi_0(x) \equiv 1$ and $\psi_n(x) = (\varphi_{\nu_1+1}(x) \dots \varphi_{\nu_k+1}(x))/(M_{\nu_1+1} \dots M_{\nu_k+1})$ for $n = 2^{\nu_1} + 2^{\nu_2} + \dots + 2^{\nu_k}$. Then

$$\psi_{2^{n-1}}(x) = \varphi_n(x)/M_n,$$

and it is easy to see that

$$\sum_{k=0}^{2^n-1} \psi_k(x)\psi_k(y) = \prod_{k=1}^{n} \left(1 + \frac{\varphi_k(x)\varphi_k(y)}{M_k^2}\right). \tag{4}$$

We want to show that the product system $\{\psi_n(x)\}$ is $L_{2^n-1}(E)$-bounded, hence $\{\varphi_n(x)/M_n\}$ is imbedded in a $L_{2^n-1}(E)$-bounded system. Taking into account $|\varphi_n(x)|/M_n \leqq 1$, the right hand side of (4) is non-negative; so we can omit the sign of absolute value in the integral defining $L_{2^n-1}(x)$, hence

$$L_{2^n-1}(x) = \int_E \sum_{k=0}^{2^n-1} \psi_k(x)\psi_k(y)\,d\mu(y) \leqq$$

$$\leqq \sum_{k=0}^{2^n-1} |\psi_k(x)| \left|\int_E \psi_k(y)\,d\mu(y)\right| \leqq \sum_{k=0}^{\infty} \left|\int_E \psi_k(y)\,d\mu(y)\right| \leqq C_1,$$

where $C_1, C_2, \ldots$ are absolute constants. The last inequality is a consequence of the weak multiplicativity of $\{\varphi_n(x)\}$. In fact, denoting by $\{\psi_n^*(x)\}$ the product system of $\{\varphi_n(x)\}$ we have by assumption

$$\sum_{n=0}^{\infty} \left|\int_E \psi_n^*(y)\,d\mu(y)\right| \leqq C_2$$

and, because of $M_n \geqq 1$,

$$\sum_{n=0}^{\infty} \left|\int_E \psi_n(y)\,d\mu(y)\right| = \sum_{n=0}^{\infty} \frac{1}{M_{\nu_1} M_{\nu_2} \ldots M_{\nu_n}} \left|\int_E \psi_n^*(y)\,d\mu(y)\right| \leqq$$

$$\leqq \sum_{n=0}^{\infty} \left|\int_E \psi_n^*(y)\,d\mu(y)\right| = C_2.$$

The sequence $\{L_{2^n-1}(x)\}$ being uniformly bounded on E, we get $L_{2^n-1}(E) \leqq C_3$ $(n=1, 2, \ldots)$ by the finiteness of $|E|$. Therefore we can apply our corollary to the series $\sum c_n M_n \dfrac{\varphi_n(x)}{M_n}$ and our statement follows.

The property to be a weakly multiplicative system is, of course, independent of the order of the terms. Hence, in the statement of Theorem 2 we can say *unconditional* convergence a.e. instead of simple convergence a.e. Theorem 2 immediately implies different forms of the strong law of great numbers (see [3]). But P. Révész [6] proved that also the law of iterated logarithm can be extended, in a proper form, to weakly independent systems.

6. Now we are looking for the absolute convergence of expansions in the functions $\varphi_n(x)$ of a weakly multiplicative system.

Theorem 3. *Let $\{\varphi_n(x)\}$ be a bounded weakly multiplicative system on the set $E \subset X$ of finite measure and assume*

$$\varliminf_{n\to\infty} \int_E |\varphi_n(x)|\,d\mu(x) \geqq q > 0. \tag{5}$$

If the μ-integrable function $f(x)$ is one-sided bounded and the expansion coefficients of $f(x)$ in the product functions $\psi_n(x)$

$$a_n = \int_E f(x)\psi_n(x)\,d\mu(x)$$

vanish except perhaps the coefficients

$$a_{2^{n-1}} = c_n = \int_E f(x)\varphi_n(x)\,d\mu(x),$$

then the series $\sum |c_n|$ is convergent.

We may assume without restricting the generality that $|\varphi_n(x)| \leqq 1$ for all n. Indeed we have $|\varphi_n(x)| \leqq C_4$ by assumption and the absolute convergence of the series $\sum c_n \varphi_n(x)$ is equivalent to that of $C_4 \sum c_n \varphi_n(x)/C_4$.

Rearrange $\{\varphi_n(x)\}$ in an arbitrary way: $\{\varphi_{\nu_k}(x)\}$, and put

$$s_n(\{\nu_k\}, x) = \sum_{k=1}^{n} c_{\nu_k}\varphi_{\nu_k}(x).$$

Denote by $\{\psi_n^*(x)\}$ the rearranged product system of $\{\varphi_n(x)\}$ corresponding to the arrangement $\{\varphi_{\nu_k}(x)\}$. Since the expansion coefficients of $f(x)$ in $\psi_k^*(x)$ vanish for $\psi_k^*(x) \neq \varphi_n(x)$, i.e. for $k \neq 2^{n-1}$, we get

$$(6) \qquad s_n(\{\nu_k\}, x) = \int_E f(t) \sum_{k=1}^{n} \varphi_{\nu_k}(t)\varphi_{\nu_k}(x)\,d\mu(t) = \int_E f(t) \sum_{k=0}^{2^n-1} \psi_k^*(t)\psi_k^*(x)\,d\mu(t).$$

By assumption $f(t)$ is bounded from one side, for instance $f(t) \leqq M$, so we infer from (6) and

$$\sum_{k=0}^{2^n-1} \psi_k^*(t)\psi_k^*(x) = \prod_{k=1}^{n} [1+\varphi_{\nu_k}(t)\varphi_{\nu_k}(x)] \geqq 0$$

the estimate

$$(7) \qquad s_n(\{\nu_k\}, x) \leqq M \sum_{k=0}^{2^n-1} |\psi_k^*(x)| \left|\int_E \psi_k^*(t)\,d\mu(t)\right| \leqq M \sum_{k=0}^{\infty} \left|\int_E \psi_k^*(t)\,d\mu(t)\right| \leqq C_5 M.$$

Furthermore, in a similar way we obtain

$$(8) \qquad -s_n(\{\nu_k\}, x) = \int_E f(t) \prod_{k=0}^{n} [1-\varphi_{\nu_k}(t)\varphi_{\nu_k}(x)]\,d\mu(t) \leqq$$

$$\leqq M \sum_{k=0}^{\infty} |\psi_k^*(x)| \left|\int_E \psi_k^*(t)\,d\mu(t)\right| \leqq C_5 M.$$

The estimates (7) and (8) give the result

$$|s_n(\{\nu_k\}, x)| \leqq C_5 M \qquad (n = 1, 2, \ldots)$$

and this common bound holds good for every rearrangement of the series $\sum c_n \varphi_n(x)$.

Hence, according to a classical theorem of Riemann, the convergence of the series $\sum|c_n\varphi_n(x)|$ follows. So the sum of this series is bounded on any $E'\subseteqq E$ with $|E'|>0$ therefore

$$\sum_{n=1}^{\infty}|c_n|\int_{E'}|\varphi_n(x)|\,d\mu(x)<\infty,$$

and so we get by (5)

$$\frac{1}{q}\sum_{n=1}^{\infty}|c_n|<\infty$$

as we have stated.

7. In section 1 we mentioned that the bounded complete orthonormal systems containing an infinite multiplicatively orthogonal subsystem are "rare" in some sense. Now we will show that *every bounded infinite orthonormal system on a set of finite measure, even if it is not complete, contains an infinite weakly multiplicative system.*

Let $\{\Phi_n(x)\}$ be a bounded infinite orthogonal system on the set E. The expansion coefficients of every L^2_μ-integrable function tend to zero, hence there exists an index n_1 such that

$$\left|\int_E \Phi_{n_1}(x)\,d\mu(x)\right|\leqq\frac{1}{2^2}.$$

Set $\varphi_1(x)=\Phi_{n_1}(x)$. Suppose, we have chosen the functions $\varphi_1(x),\varphi_2(x),\dots,\varphi_{n-1}(x)$ from $\{\Phi_n(x)\}$ in such a way that for every product $\varphi_{\nu_1}\varphi_{\nu_2}\dots\varphi_{\nu_k}$ with indices $\nu_1<$ $<\nu_2<\dots<\nu_k\leqq n-1$

$$\left|\int_E\prod_{j=1}^{k}\varphi_{\nu_j}(x)\,d\mu(x)\right|\leqq\frac{1}{2^{2(n-1)}}$$

holds true. All the finite products of the φ_k's being L^2_μ-integrable, for every product $\varphi_{\nu_1}\varphi_{\nu_2}\dots\varphi_{\nu_k}$ there exists a number n_m, depending on the choice of the product, such that

$$\left|\int_E\Phi_n(x)\prod_{j=1}^{k}\varphi_{\nu_j}(x)\,d\mu(x)\right|\leqq\frac{1}{2^{2n}}$$

for every $n\geqq n_m$. There are 2^{n-1} different products of this form, hence at most 2^{n-1} indices n_m. Denote by n_N the greatest of them and set $\varphi_n(x)=\Phi_{n_N}(x)$. Then

$$\left|\int_E\varphi_n(x)\prod_{j=1}^{k}\varphi_{\nu_j}(x)\,d\mu(x)\right|\leqq\frac{1}{2^{2n}}.$$

In this way we defined the infinite system $\{\varphi_n(x)\}$ by induction. To see that this system is weakly multiplicative, form all different products $\varphi_{v_1}\varphi_{v_2}\dots\varphi_{v_k}$ of the first n functions $\varphi_1, \varphi_2, \dots, \varphi_n$. There are 2^n such products, hence

$$\sum \Big| \int_E \varphi_{v_1}(x)\,\varphi_{v_2}(x)\dots\varphi_{v_k}(x)\,d\mu(x)\Big| \leqq \frac{1}{2^n} \tag{9}$$

where the sum has to be taken over all 2^n different products. The sum S of the absolute values of integrals of all possible finite products formed with the functions of the system $\{\varphi_n(x)\}$ is less than the sum of the sums (9) taken for $n=1, 2, \dots$. Hence

$$S < \sum_{n=1}^{\infty} \frac{1}{2^n} = 1.$$

This estimate means just that $\{\varphi_n(x)\}$ is weakly multiplicative.

Literature

[1] G. Alexits—A. Sharma, On the convergence of multiplicatively orthogonal series, *Acta Math. Sci. Hung.*, **22** (1971), 257—266.

[2] ——— The influence of Lebesgue functions on the convergence and summability of function series, *Acta Sci. Math.*, **33** (1972), 1—10.

[3] G. Alexits, Stochastische Unabhängigkeit und Orthogonalität, *Mitteilungen aus dem Mathem. Seminar Giessen*, **92** (1971), 1—10.

[4] H. Fiedler—R. Trautner, Zur Existenz vollständiger Orthogonalsysteme, welche kein multiplikativ orthogonales Teilsystem enthalten, *Periodica Math. Hung.* (to appear).

[5] L. Friess—R. Trautner, Zur Häufigkeit multiplikativ orthogonaler Funktionensysteme, *Periodica Math. Hung.* (to appear).

[6] P. Révész, A new law of the iterated logarithm for multiplicative systems and its applications, *Acta Sci. Math.*, **34** (1973), 349—358.

(Received May 3, 1972)

Structure of operators with numerical radius one

By T. ANDO in Sapporo (Japan)

Dedicated to Professor Béla Sz.-Nagy on his sixtieth birthday

1. Introduction

The *numerical radius* $w(T)$ of a bounded linear operator T on a Hilbert space $\mathfrak{H}$ is defined by

$$w(T) = \sup \{|(Th, h)| : \|h\| \leqq 1\}.$$

Important characterization of operators with numerical radius not greater than one was discovered by BERGER and subsequently generalized by SZ.-NAGY and FOIAŞ (see [1] I-1): $w(T) \leqq 1$ if and only if there is a unitary operator W on a Hilbert space $\mathfrak{K}$, containing $\mathfrak{H}$ as a subspace, such that for all $h \in \mathfrak{H}$ and $n \geqq 1$

$$T^n h = 2PW^n h$$

where P is the projection from $\mathfrak{K}$ to $\mathfrak{H}$. W is called a *unitary 2-dilation* of T.

The key result of the present paper is an intrinsic characterization of operators with numerical radius not greater than one (Theorem 1): $w(T) \leqq 1$ if and only if there are a selfadjoint contraction A and a contraction B such that $T = (1+A)^{\frac{1}{2}} B (1-A)^{\frac{1}{2}}$. This factorization theorem makes it possible to construct a unitary 2-dilation in simple matricial form just as the Schäffer description of a unitary dilation of a contraction (Theorem 2).

2. Factorization

$\mathfrak{H}$ is a Hilbert space, and $\bigoplus_{j=-\infty}^{\infty} \mathfrak{H}_j$ denotes direct sum of copies of $\mathfrak{H}$ in which $\mathfrak{H}$ is identified with $\cdots \oplus 0 \oplus \mathfrak{H}_0 \oplus 0 \oplus \cdots$ in the canonical way. A bounded linear operator S on $\bigoplus_j \mathfrak{H}_j$ can be represented by its matricial components, $[S_{j,k}]$, where $S_{j,k}$ is an operator on $\mathfrak{H}$, considered as an operator from $\mathfrak{H}_k$ to $\mathfrak{H}_j$.

Lemma 1. *If* $w(T) \leqq 1$, *there is a positive contraction* X *such that*

$$(Xh, h) = \inf_g \left(\begin{bmatrix} 1 & \frac{1}{2}T^* \\ \frac{1}{2}T & X \end{bmatrix} \begin{bmatrix} h \\ g \end{bmatrix}, \begin{bmatrix} h \\ g \end{bmatrix} \right). \tag{1}$$

Moreover X is the maximum of all positive contractions Y for which $\begin{bmatrix} 1-Y & \frac{1}{2}T^* \\ \frac{1}{2}T & Y \end{bmatrix}$ *is positive.*

Proof. Suppose that W is a unitary 2-dilation of T on a Hilbert space $\mathfrak{K} \supseteq \mathfrak{H}$. Let $X_0 = 1$ and X_n be the compression of $1 - Q_n$ to $\mathfrak{H}$ where Q_n is the projection to $\bigvee_{j=1}^{n} W^{*j}(\mathfrak{H})$. This means that

$$(X_n h, h) = \inf_{h_1, \ldots, h_n \in \mathfrak{H}} \left\| h + \sum_{j=1}^{n} W^{*j} h_j \right\|^2 =$$

$$= \inf_{h_1, \ldots, h_n} \left\{ (h, h) + \left(h, \sum_{j=1}^{n} W^{*j} h_j \right) + \left(\sum_{j=1}^{n} W^{*j} h_j, h \right) + \sum_{j,k=1}^{n} (W^{*j} h_j, W^{*k} h_k) \right\} =$$

$$= \inf_{h_1, \ldots, h_n} \left(\begin{bmatrix} 1 & \frac{1}{2}T^* & \cdots & \frac{1}{2}T^{*n} \\ \frac{1}{2}T & 1 & & \vdots \\ \vdots & & \ddots & \vdots \\ \vdots & & 1 & \frac{1}{2}T^* \\ \frac{1}{2}T^n & \cdots & \frac{1}{2}T & 1 \end{bmatrix} \begin{bmatrix} h \\ h_1 \\ \vdots \\ h_n \end{bmatrix}, \begin{bmatrix} h \\ h_1 \\ \vdots \\ h_n \end{bmatrix} \right).$$

Since

$$\begin{bmatrix} 1 & \frac{1}{2}T^* & \cdots & \frac{1}{2}T^{*n} \\ \frac{1}{2}T & 1 & & \vdots \\ \vdots & & \ddots & \vdots \\ \vdots & & 1 & \frac{1}{2}T^* \\ \frac{1}{2}T^n & \cdots & \frac{1}{2}T & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & T^* & \cdots & T^{*n} \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & \vdots \\ \vdots & & 1 & T^* \\ 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2}T^* & 0 & \cdots & 0 \\ -\frac{1}{2}T & 1 & \ddots & \ddots & \vdots \\ & \ddots & \ddots & \ddots & 0 \\ 0 & & \ddots & 1 & -\frac{1}{2}T^* \\ \vdots & \ddots & & \ddots & \\ 0 & \cdots & 0 & -\frac{1}{2}T & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ T & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 0 \\ T^n & \cdots & \cdots & T & 1 \end{bmatrix},$$

X_n admits the alternative representation

$$(X_n h, h) = \inf_{h_1, \ldots, h_n} \left(\begin{bmatrix} 1 & \frac{1}{2}T^* & 0 & \cdots & 0 \\ \frac{1}{2}T & 1 & \ddots & \ddots & \vdots \\ & \ddots & \ddots & \ddots & 0 \\ 0 & & \ddots & \ddots & \frac{1}{2}T^* \\ \vdots & \ddots & & \ddots & \\ 0 & \cdots & 0 & \frac{1}{2}T & 1 \end{bmatrix} \begin{bmatrix} h \\ h_1 \\ \vdots \\ h_n \end{bmatrix}, \begin{bmatrix} h \\ h_1 \\ \vdots \\ h_n \end{bmatrix} \right).$$

The corresponding representation for X_{n-1} yields

$$(2)\qquad (X_n h, h) = \inf_g \left(\begin{bmatrix} 1 & \frac{1}{2}T^* \\ \frac{1}{2}T & X_{n-1} \end{bmatrix} \begin{bmatrix} h \\ g \end{bmatrix}, \begin{bmatrix} h \\ g \end{bmatrix} \right).$$

Since X_n converges decreasingly to some X, transfer to limit in (2) leads to the relation (1).

Now if $\begin{bmatrix} 1-Y & \frac{1}{2}T^* \\ \frac{1}{2}T & Y \end{bmatrix}$ is positive, $X_{n-1} \geqq Y$ implies by (1) and (2)

$$(X_n h, h) \geqq \inf_g \left(\begin{bmatrix} 1 & \frac{1}{2}T^* \\ \frac{1}{2}T & Y \end{bmatrix} \begin{bmatrix} h \\ g \end{bmatrix}, \begin{bmatrix} h \\ g \end{bmatrix} \right) \geqq (Yh, h),$$

hence $X_n \geqq Y$. Now $X \geqq Y$ follows from $X_0 = 1 \geqq Y$.

Theorem 1. *The numerical radius of T is not greater than one if and only if T admits a factorization*

$$T = (1+A)^{\frac{1}{2}} B (1-A)^{\frac{1}{2}}$$

with a selfadjoint contraction A and a contraction B. Moreover in the set of such A there exist the maximum $A_{\max}$ and the minimum $A_{\min}$ and the corresponding $B_{\max}$ (resp. *the adjoint of $B_{\min}$*) *is isometric on the range of* $1-A_{\max}$ (resp. *that of* $1+A_{\min}$).

Proof. Suppose that T admits the factorization. Then

$$|(Th, h)| = |(B(1-A)^{\frac{1}{2}}h, (1+A)^{\frac{1}{2}}h)| \leqq \tfrac{1}{2}\{\|(1-A)^{\frac{1}{2}}h\|^2 + \|(1+A)^{\frac{1}{2}}h\|^2\} = \|h\|^2,$$

which shows $w(T) \leqq 1$.

Conversely, if $w(T) \leqq 1$, by Lemma 1 there is a positive contraction X with (1). Since (1) is equivalent to

$$\|(1-X)^{\frac{1}{2}}h\| = \sup_g \frac{|(\frac{1}{2}Th, g)|}{\|X^{\frac{1}{2}}g\|}$$

with convention $0/0=0$, to each h there corresponds uniquely f in the closure of the range of X such that

$$X^{\frac{1}{2}}f = \tfrac{1}{2}Th \quad \text{and} \quad \|f\| = \|(1-X)^{\frac{1}{2}}h\|.$$

Thus there is a contraction $B_{\max}$ which is isometric on the range of $1-X$ and $\frac{1}{2}T = X^{\frac{1}{2}}B_{\max}(1-X)^{\frac{1}{2}}$. Now $A_{\max} = 2X-1$ meets the requirement. Given any factorization with A and B it follows with $Y = \frac{1}{2}(1+A)$ that

$$\begin{bmatrix} 1-Y & \frac{1}{2}T^* \\ \frac{1}{2}T & Y \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1-A & T^* \\ T & 1+A \end{bmatrix} =$$

$$= \frac{1}{2}\begin{bmatrix} (1-A)^{\frac{1}{2}} & 0 \\ 0 & (1+A)^{\frac{1}{2}} \end{bmatrix}\begin{bmatrix} 1 & B^* \\ B & 1 \end{bmatrix}\begin{bmatrix} (1-A)^{\frac{1}{2}} & 0 \\ 0 & (1+A)^{\frac{1}{2}} \end{bmatrix}.$$

Since $\|B\| \leqq 1$ is equivalent to $\begin{bmatrix} 1 & B^* \\ B & 1 \end{bmatrix} \geqq 0$, $\begin{bmatrix} 1-Y & \frac{1}{2}T^* \\ \frac{1}{2}T & Y \end{bmatrix}$ is positive. Then Lemma 1 shows $Y \leqq X$, hence $A \leqq A_{\max}$. The minimum operator $A_{\min}$ can be obtained so as $-A_{\min}$ is the maximum operator for T^*. This completes the proof.

In general $A_{\max}$ is different from $A_{\min}$. This is shown with the simple 2×2 matrix $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$. In this case

$$A_{\max} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad A_{\min} = \begin{bmatrix} -1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}.$$

In terms of the unitary 2-dilation W the maximum and the minimum operators are given by the following formulas:

$$A_{\max} = 1 - 2PQ_- P \quad \text{and} \quad A_{\min} = 2PQ_+ P - 1$$

where P is the projection to $\mathfrak{H}$ and Q_+(resp. Q_-) is the projection to $\bigvee_{n=1}^{\infty} W^n(\mathfrak{H})$ $\left(\text{resp. } \bigvee_{n=1}^{\infty} W^{-n}(\mathfrak{H})\right)$.

Theorem 2. *The numerical radius of T is not greater than one if and only if there is a contraction C such that*

$$T = 2(1 - C^* C)^{\frac{1}{2}} C. \tag{3}$$

Under such factorization a unitary 2-dilation W is realized as an operator on $\bigoplus_{j=-\infty}^{\infty} \mathfrak{H}_j$ with components:

$$W_{k+1,k} = 1 \quad \text{for} \quad k \geqq 1 \quad \text{or} \quad k \leqq -3,$$

$$W_{-1,-2} = (1 - CC^*)^{\frac{1}{2}}, \quad W_{-1,-1} = C(1 - CC^*)^{\frac{1}{2}}, \quad W_{-1,0} = C^2,$$

$$W_{0,-2} = -C^*, \quad W_{0,-1} = (1 - C^* C)^{\frac{1}{2}}(1 - CC^*)^{\frac{1}{2}}, \quad W_{0,0} = (1 - C^* C)^{\frac{1}{2}} C,$$

$$W_{1,-1} = -C^*, \quad W_{1,0} = (1 - C^* C)^{\frac{1}{2}}$$

and $W_{j,k} = 0$ for other j, k.

Proof. If $w(T) \leqq 1$, by Theorem 1

$$T = (1 + A_{\max})^{\frac{1}{2}} B_{\max} (1 - A_{\max})^{\frac{1}{2}}$$

and $B_{\max}$ is isometric on the range of $1 - A_{\max}$. Let $C = 2^{-\frac{1}{2}} B_{\max}(1 - A_{\max})^{\frac{1}{2}}$. Then

$$1 - C^* C = 1 - \tfrac{1}{2}(1 - A_{\max}) = \tfrac{1}{2}(1 + A_{\max});$$

hence

$$T = 2(1 - C^* C)^{\frac{1}{2}} C.$$

Suppose conversely that T admits a factorization (3). It is well known (see [1]

I-5) that a unitary dilation U of C is realized as the operator on $\bigoplus_{j=-\infty}^{\infty} \mathfrak{H}_j$ with components

$$U_{k+1,k} = 1 \quad \text{for} \quad k \geqq 1 \quad \text{or} \quad k \geqq -2,$$
$$U_{0,0} = C, \qquad U_{0,-1} = (1-CC^*)^{\frac{1}{2}},$$
$$U_{1,0} = (1-C^*C)^{\frac{1}{2}}, \qquad U_{1,-1} = -C^*,$$

and $U_{j,k}=0$ for other j, k. Then W in the assertion is written in the form $W=VU^2$ where V is the backward shift, that is,

$$V_{j,k} = \delta_{j,k-1} \quad \text{for all} \quad j,\ k;$$

hence W is unitary. W is a unitary 2-dilation of T if

(4) $$(W^n)_{0,0} = \tfrac{1}{2}T^n \qquad (n = 1, 2, \ldots).$$

To prove (4) by induction, assume that

$$(W^n)_{-k,0} = 0 \quad (k \geqq 2), \quad (W^n)_{-1,0} = C^2T^{n-1} \quad \text{and} \quad (W^n)_{0,0} = \tfrac{1}{2}T^n,$$

which is valid for $n=1$ by definition. Matrix multiplication shows

$$(W^{n+1})_{-k,0} = (W^n)_{-k-1,0} = 0 \quad (k \geqq 2),$$
$$(W^{n+1})_{-1,0} = C(1-CC^*)^{\frac{1}{2}}(W^n)_{-1,0} + C^2(W^n)_{0,0} =$$
$$= C(1-CC^*)^{\frac{1}{2}}C^2T^{n-1} + \tfrac{1}{2}C^2T^n = \tfrac{1}{2}C^2T^n + \tfrac{1}{2}C^2T^n = C^2T^n$$

and

$$(W^{n+1})_{0,0} = (1-C^*C)^{\frac{1}{2}}(1-CC^*)^{\frac{1}{2}}(W^n)_{-1,0} + \tfrac{1}{2}T(W^n)_{0,0} =$$
$$= (1-C^*C)^{\frac{1}{2}}(1-CC^*)^{\frac{1}{2}}C^2T^{n-1} + \tfrac{1}{4}T^{n+1}$$
$$= \tfrac{1}{4}T^{n+1} + \tfrac{1}{4}T^{n+1} = \tfrac{1}{2}T^{n+1}.$$

Here, besides the relation $(1-C^*C)^{\frac{1}{2}}C = \frac{1}{2}T$, the well-known formula (see [1] I-5)

$$(1-CC^*)^{\frac{1}{2}}C = C(1-C^*C)^{\frac{1}{2}}$$

is used. This completes the proof.

Reference

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

RESEARCH INSTITUTE OF APPLIED ELECTRICITY
HOKKAIDO UNIVERSITY, SAPPORO, JAPAN

(Received September 8, 1972)

Ergodic theory and the measure of sets in the Bohr group

By J. R. BLUM,*) B. EISENBERG and L.-S. HAHN in Albuquerque (New Mexico, U.S.A.)

To Professor Béla Szőkefalvi-Nagy in honor of his 60th birthday

Introduction. Let U be a strongly continuous unitary representation of a locally compact group G on a Hilbert space H. If μ is a probability measure on G, $\int U_g\,d\mu$ is defined weakly so that $(\int U_g f\,d\mu, h) = \int (U_g f, h)\,d\mu$ for all h in H. For notation and terminology see [3].

Proposition 1. *If G is compact and μ is Haar measure then $\int U_g f\,d\mu = Pf$, where P is the orthogonal projection on the space $K = \{k \mid U_g k = k,\ U_g k = k, g \in G\}$.*

Proof. $\int U_g f\,d\mu$ is invariant since

$$(U_{g_0}\int U_g f d\mu, h) = (\int U_g f d\mu, U_{g_0}^* h) = \int (U_{g_0 g} f, h) = \int (U_g f, h)\,d\mu = (\int U_g f d\mu, h)$$

by the invariance of Haar measure. Since this holds for all $h \in H$, $U_{g_0}\int U_g f\,d\mu = \int U_g f\,d\mu$.

To complete the proof it must be shown that if $k \in K$ then $f - \int U_g f\,d\mu \perp k$. But

$$(f - \int U_g f d\mu, k) = (f, k) - \int (f, U_{g^{-1}} k)\,d\mu = (f, k) - \int (f, k)\,d\mu = 0.$$

Let μ_n be a sequence of probability measures on a locally compact Abelian group G. In BLUM and EISENBERG [2] the following theorem and corollary are proved.

Theorem. *The following are equivalent:*

(i) *For every continuous unitary representation U of G and every f in H, $\int U_g f\,d\mu_n$ converges in mean to Pf.*

(ii) *For every character x on G except that identically 1 the Fourier transforms $\hat{\mu}_n(x) = \int \langle x, g\rangle\,d\mu_n$ converge to 0.*

(iii) *μ_n considered as restrictions of measures on the Bohr compactification $\bar{G}$ of G converge weakly to Haar measure on $\bar{G}$.*

---

*) Research supported by N.S.F. Grant GP-25 736.

(iv) *For every character x of infinite order the measures μ_n^x induced by x on the unit circle in the complex plane converge weakly to normalized Lebesgue measure on the circle and for every character x of order m, $m=0, 1, 2, \dots$ the measures μ_n^x converge weakly to Haar measure on the m^{th} roots of unity.* (The measure μ_n^x induced by x is defined so that $\mu_n^x(B)=\mu_n(\{g|(x, g)\in B\})$.)

Corollary. *If E_n is a sequence of sets in G such that for all g in G, $\frac{\mu(E_n\cap gE_n)}{\mu(E_n)}\to 1$, where μ is Haar measure, then the measures $\mu_n(A)=\frac{\mu(A\cap E_n)}{\mu(E_n)}$ converge weakly to Haar measure on $\bar{G}$.*

Proposition 1 and the theorem lead to the questions studied in this paper. Question one asks, does $\int_{\bar{G}}\overline{(U_g f, h)}\,d\mu=(Pf, h)$ hold, where $\overline{(U_g f, h)}$ is a suitable extension of $(U_g f, h)$ to $\bar{G}$ and μ is Haar measure on $\bar{G}$? The theorem says that $(Pf, h)=\lim_n\int(U_g f, h)\,d\mu_n$, where μ_n converges weakly to Haar measure on $\bar{G}$ and Proposition 1 says that the statement is true when $\bar{G}$ is already compact. If the answer is yes, there would be an interesting expression for Pf in terms of the action of U_g on f.

Question two asks for which sequences of integers the mean ergodic theorem holds; i.e., when is it true that $\frac{1}{N}\sum_1^N T^{n_k}f$ converges in mean to the projection of f on the space of elements invariant under T for every unitary T. The theorem says that a sequence is ergodic if and only if the probability measure μ_N giving measure $\frac{1}{N}$ to each integer $n_1, n_2, \dots, n_N$ converges weakly to Haar measure on the Bohr compactification of the integers.

Both questions relate to the study of the measure of sets in the Bohr compactification of the integers.

1. This section is concerned with the first question. It is seen that properties of the spectral resolution of U are crucial.

Proposition 2. *If U has pure point spectrum then $(Pf, h)=\int_{\bar{G}}\overline{(U_g f, h)}\,d\mu$, where $\overline{(U_g f, h)}$ is the unique continuous extension of $(U_g f, h)$ to $\bar{G}$ and μ is Haar measure on $\bar{G}$.*

Proof. By Stone's theorem and the assumption on discrete spectrum $(U_g f, h)=$ $=\int_{\hat{G}}\langle x, g\rangle\,d(E_x f, h)=\sum C_k\langle x_k, g\rangle$, where $\hat{G}$ is the dual group of G and $\sum|C_k|<\infty$. Since $|\langle x_k, g\rangle|=1$, $(U_g f, h)$ is a uniform limit of almost periodic functions and hence is almost periodic. It follows that $(U_g f, h)$ is the restriction of a continuous function on $\bar{G}$. By the theorem

$$(Pf, h) = \lim_G \int (U_g f, h)\, d\mu_n = \lim_{\bar{G}} \int \overline{(U_g f, h)}\, d\mu_n = \int_{\bar{G}} \overline{(U_g f, h)}\, d\mu.$$

In fact, if a function $\varphi(g)$ is merely bounded and continuous a.e. $d\mu$ on $\bar{G}$,

$$\int \varphi(g)\, d\mu_n \to \int \varphi(g)\, d\mu.$$

Lemma. *Let X be a normal topological space and μ a finite regular measure on X. If $\mu_n \to \mu$ weakly and if φ is bounded and continuous a.e. $d\mu$, then $\int \varphi\, d\mu_n \to \int \varphi\, d\mu$.*

As the proof of this lemma is somewhat technical and independent of the rest of the paper it will be relegated to an appendix.

For the remainder of the paper attention is limited to unitary groups generated by a single operator.

Proposition 3. *If the maximal spectral type of T has no continuous singular part then*

$$\int_{\bar{Z}} \overline{(T^n f, h)}\, d\mu = (Pf, h),$$

where $\overline{(T^n f, h)}$ is a continuous a.e. $d\mu$ extension of $(T^n f, h)$.

Proof.

$$(T^n f, h) = \int_0^{2\pi} e^{int}\, d(E_t f, h) = \int_0^{2\pi} e^{int} \varrho(t)\, dt + \Sigma C_k e^{int_k},$$

where $\int_0^{2\pi} |\varrho(t)|\, dt < \infty$ and $\Sigma |C_k| < \infty$.

As in Proposition 2, $\sum C_k e^{int_k}$ has a unique continuous extension to $\bar{Z}$. By the Riemann—Lebesgue lemma $\int e^{int_k} \varrho(t)\, dt \to 0$ as $n \to \infty$. It follows that if

$$\overline{(T^n f, h)} = \begin{cases} (T^n f, h) & \text{on } Z, \\ \overline{\Sigma C_k e^{int_k}} & \text{on } \bar{Z} - Z, \end{cases}$$

then $\overline{(T^n f, h)}$ is continuous except on Z itself. That is, if $n_k \to g \in \bar{Z} - Z$ then $\int e^{in_k t} \varrho(t) \to 0$ so that $(T^{n_k} f, h) \to \overline{\sum C_k e^{int_k}(g)}$. But Z has measure 0 in $\bar{Z}$ so that $\overline{(T^n f, h)}$ is continuous a.e. $d\mu$. By the theorem and lemma,

$$(Pf, h) = \lim_k \int (T^n f, h)\, d\mu_k = \lim_k \int \overline{(T^n f, h)}\, d\mu_k = \int \overline{(T^n f, h)}\, d\mu.$$

Finally T with continuous singular spectrum must be considered. It is no longer true that $(T^n f, h) \to 0$. However, $(T^n f, h)$ does approach 0 except on a sequence of density 0. An increasing sequence n_k has density 0 if $\frac{\#\{n_k<n\}}{n}\to 0$. (If A is a set then $\#(A)$ is the cardinality of A.) If it could be shown that such sequences have closure of Haar measure 0 in $\bar{Z}$ then by a similar argument to that in Proposition 3 it could be shown that $\int \overline{(T^n f, h)}\,d\mu = (Pf, h)$, were $\overline{(T^n f, h)}$ is defined as in Proposition 3.

This leads to the problem of determining when the Haar measure of the closure of sets of integers in $\bar{Z}$ is zero.

2. Question one leads to the question of which sequences have closure of measure 0 in $\bar{Z}$. Question two asks which sequences induce measures converging weakly to Haar measure on $\bar{Z}$. Such scquences must be dense in $\bar{Z}$. Otherwise there is an open set Θ in $\bar{Z}$ containing no elements of the sequence. By Urysohn's lemma there is a non-negative continuous function φ with support inside Θ such that $\int \varphi\,d\mu > 0$. But $\int \varphi\,d\mu_n \equiv 0$. It will thus be of interest to find conditions merely for denseness of sequences of integers in $\bar{Z}$.

Proposition 4. *Cosets of the subgroup* $H=\{0, \pm m, \pm 2m, \ldots\}$ *have disjoint closures in* $\bar{Z}$ *and each has measure* $1/m$.

Proof. A neighborhood of g_0 in $\bar{Z}$ is defined by $\{g \,|\, |\langle t_i, g\rangle - \langle t_i, g_0\rangle| < \varepsilon\}$ where $\varepsilon > 0$ and $0 \leqq t_i < 2\pi$. Consider the character corresponding to $t = \frac{2\pi}{m}$. Then if $g \in k+H$, $\langle t, g\rangle = e^{2\pi ik/m}$ while if $g' \in k'+H$, $\langle t, g'\rangle = e^{2\pi ik'/m}$. If $g_0 \in \overline{k+H}$ then $\langle t, g_0\rangle = e^{2\pi ik/m}$ while if $g_0 \in \overline{k'+H}$, $\langle t, g_0\rangle = e^{2\pi ik'/m}$. Since $\bigcup_{k=1}^{m} (k+H)$ is dense in $\bar{Z}$ and $\overline{k+H}$ and $\overline{k'+H}$ are translates of one another, $\mu(\overline{k+H}) = 1/m$, $k=1, 2, \ldots, m$.

Corollary. *The following sequences have closure of measure* 0 *in* $\bar{Z}$.

(i) $n!$,

(ii) a^n, *where* a *is an integer,*

(iii) p_n, *the sequence of primes,*

(iv) n^k, *where* k *is a fixed integer* $\geqq 2$.

Proof. Since each integer has measure 0 and the topology is Hausdorff, a finite number of elements in the sequence can be neglected.

(i) For $n \geqq m$, $n! \equiv 0 \mod m$. Hence $\mu(\overline{\{n!\}}) \leqq 1/m$ since $\{n! \,|\, n \geqq m\}$ is a subset of $\{0, \pm m, \pm 2m, \ldots\}$. But m is arbitrary. Hence $\mu(\overline{\{n!\}}) = 0$.

(ii) For $n \geqq m$, $a^n \equiv 0 \mod a^m$. Thus $\mu(\overline{\{a^n\}}) \leqq 1/a^m$. Again m is arbitrary so $\mu(\overline{\{a^n\}}) = 0$.

(iii) Consider the set of residues of all primes modulo m. For a given prime p either $p \leqq m$ or $p = km + r$, where $k \geqq 1$. In the latter case r is relatively prime to m. Otherwise p would be divisible by the greatest common divisor of m and r. By the prime number theorem the number of primes less than or equal to m divided by m goes to 0. The number of integer r less than m and relatively prime to m is just the Euler Φ function of m. $\Phi(m) = m\prod_i\left(1-\frac{1}{q_i}\right)$, where $m=q_1^{\alpha_1}\dots q_s^{\alpha_s}$ is the prime factorization of m. $\frac{\Phi(m)}{m} = \prod_i\left(1-\frac{1}{q_i}\right)$ which can be made arbitrarily small since $\sum 1/p_i=\infty$, where p_i is the sequence of primes (LEVEQUE [5], p. 100). Thus $\mu(\overline{\{p_i\}}) \leqq \frac{\Phi(m)}{m}+o(m)$. Since m can be any product of primes, $\mu(\overline{\{p_i\}})=0$.

(iv) By Dirichlet's Theorem (LEVEQUE [5], p. 76) there is an infinite number of primes of the form $p = kn+1$ as n goes through the integers.

For such primes p, $p-1$ is divisible by k. As a consequence of Theorem 4-14 (LEVEQUE [4], p. 58) there are $\frac{p-1}{k}+1$ residue classes occupied by the residues of kth powers mod p. The fraction of classes occupied is $\frac{1}{p}\left(\frac{p-1}{k}+1\right) \leqq \frac{1}{k}+\frac{1}{p} \leqq \frac{5}{6}$ if $k\geqq 2$ and $p>3$. For a fixed k choose an infinite sequence of primes $p_m>0$ of the above form. By the Chinese Remainder Theorem

$$Z_{p_1 p_2 \dots p_m} \cong Z_{p_1}\times Z_{p_2}\times\cdots\times Z_{p_m}$$

via the map $x\to(x \bmod p_1, x \bmod p_2, \dots, x \bmod p_m)$.

Thus the number of residue classes occupied by kth powers modulo $p_1\dots p_m$ is $\prod_{i=1}^{m}\left(\frac{p_m-1}{k}+1\right)$. The fraction occupied is less than $\left(\frac{5}{6}\right)^m$ which can be made arbitrarily small by choosing m large enough. Hence $\mu(\overline{\{n^k\}})=0$.

If all sequences of density 0 had closures of measure 0 in $\bar{Z}$ question one would be answered. Unfortunately this is not the case as is shown in the next proposition. The question is still open as to whether the sequence where the Fourier transform of a continuous singular measure fails to go to zero can be of this type, namely, have closure of positive measure.

Proposition 5. *Consider a set S of integers of the form C_n+k, $k = 0, 1, 2, \dots$ $\dots, n-1$, where C_n increases and $\frac{n(n+1)}{C_n} \to 0$. As a sequence, S has density 0 but $\bar{S}=\bar{Z}$.*

Proof. The number of elements of S less than C_n is $\frac{n(n+1)}{2}$. Thus on the subsequence C_n

$$\frac{\#\{\text{elemens of } S \text{ which are less than } C_n\}}{C_n} = \frac{n(n+1)}{2C_n} \to 0.$$

Since there are relatively few terms of the sequence between C_n and C_{n+1}, if n is large enough the oscillation in the density between C_n and C_{n+1} goes to zero. Hence the sequence of elements of S has density 0.

However, it is easy to check that the sequence of sets E_n of the first n elements of S satisfies the conditions of the corollary to the theorem in the introduction. That is, for any k, $\frac{\#(E_n \cap E_n + k)}{n} \to 1$. (The fraction of elements of E_n with a kth successor approaches one.) By the corollary the measures $\mu_n(A) = \frac{\#(A \cap E_n)}{n}$ converge weakly to Haar measure on $\bar{Z}$ and by the argument at the beginning of this section, S must be dense in $\bar{Z}$.

Corollary. *Mean ergodic theorems hold for some sequences of density* 0.

Proof. This follows from the theorem and the proof of this proposition.

The sequence described in Proposition 5 is not only dense in $\bar{Z}$, its induced sequence of measures converges weakly to Haar measure on $\bar{Z}$. The remainder of the paper considers conditions for denseness alone.

For this part of the work a generalized Kronecker Theorem is needed. As stated in RUDIN [6], p. 98, G is a locally compact Abelian group. For $x \in G$, put $S(x) = T$ if x has infinite order; if x has order q, put $S(x)$ = the qth roots of unity.

Theorem. *Suppose E is a finite independent (in the group theoretic sense) set in G, f is a function on E such that $f(x) \in S(x)$ for all $x \in E$ and $\varepsilon > 0$. Then there exists a $\gamma \in \Gamma$ such that*

$$|\langle x, \gamma \rangle - f(x)| < \varepsilon \quad (x \in E).$$

A concrete Kronecker theorem is in KATZNELSON [4], p. 60.

For our purposes G is the unit circle in C and $\Gamma = Z$. x has infinite order if it is of the form $2\pi\alpha$, where α is irrational and x has finite order if x is of the form $2\pi\frac{k}{m}$. The abstract Kronecker theorem gives a necessary condition for denseness in the Bohr group. Namely,

Corollary. *In order that a sequence of integers be dense in $\bar{Z}$ it is necessary that for every finite independent set E in T, and every f such that $f(x) \in S(x)$, and $\varepsilon > 0$ there exist an n_k in the sequence such that $|f(x) - e^{in_kx}| < \varepsilon$ for $x \in E$.*

Proof. There is some integer n such that $|f(x) - e^{inx}| < \varepsilon/2$ for $x \in E$. To approximate this integer n in the topology on the Bohr group there must be an n_k such that $|e^{inx} - e^{in_kx}| < \varepsilon/2$ for $x \in E$. By the triangle inequality $|f(x) - e^{in_kx}| < \varepsilon$ for $x \in E$.

It also follows from the Kronecker theorem that covering every residue class of every integer is not sufficient for density in $\bar{Z}$.

Corollary. *There exists a sequence n_k with elements in each residue class of every integer with $\mu(\{\overline{n_k}\})=0$, where μ is Haar measure on $\bar{Z}$.*

Proof. Take a fixed irrational number α and an arbitrary integer m. By the Kronecker theorem, given any residue class j of m and any $\varepsilon>0$ there is an integer n with $e^{\frac{2\pi in}{m}}=e^{\frac{2\pi ij}{m}}$ and $|e^{2\pi i\alpha n}-1|<\varepsilon$. By varying j, m and ε a sequence n_k can be selected going through every residue class of every integer with $e^{2\pi i\alpha n_k}$ converging to 1.

Such a sequence $\{n_k\}$ must have closure of measure 0. To see this note that if S is a set of integers such that $|e^{2\pi i\alpha n}-1|<\varepsilon$ for $n\in S$, then for any g in $\bar{S}$, $|\langle 2\pi\alpha, g\rangle-1|<\varepsilon$. Since $e^{2\pi i\alpha k}$ is dense in the circle as k goes through the integers there exist $\left[\frac{2\pi}{\varepsilon}\right]$ translates of S with disjoint closures, where $\left[\frac{2\pi}{\varepsilon}\right]$ is the greatest integer less than $\frac{2\pi}{\varepsilon}$. Hence by the same argument as Proposition 4 $\mu(\bar{S})\leqq 1/\left[\frac{2\pi}{\varepsilon}\right]$.

For the constructed sequence n_k we can neglect a finite number of terms to show $\mu(\{\overline{n_k}\})\leqq 1/\left[\frac{2\pi}{\varepsilon}\right]$ for all $\varepsilon>0$. Hence $\mu(\{\overline{n_k}\})=0$.

Let n_k be a sequence of integers and E_m the set of the first m of them. E_m+k is the shift of the set E_m by k. Let

(i) be the statement $\lim_{m=\infty}\frac{\#\{E_m\cap E_m+k\}}{m}=1$ for all k.

(ii) be the statement $\mu_m(A)=\frac{\#(A\cap E_m)}{m}$ converges weakly to Haar measure on $\bar{Z}$.

(iii) be the statement that if $x=2\pi\alpha$, α irrational then e^{ixn_k} is uniformly distributed on the unit circle and if $x=2\pi r$, r a primitive qth root of unity then e^{ixn_k} is uniformly distributed on the qth roots of unity.

(iv) be the statement that the set $\{n_k\}$ is dense in $\bar{Z}$, and

(v) be the statement

If E is a finite independent set on the circle and $f(x)$ is a function on E of absolute value one such that if x is a primitive qth root of unity $f(x)$ is a qth roots of unity then $\forall\varepsilon>0$, there is an n_k such that

$$|f(x)-e^{ixn_k}|<\varepsilon.$$

Putting several results together we get

Theorem 1. (i)$\Rightarrow$(ii)$\Leftrightarrow$(iii)$\Rightarrow$(iv)$\Rightarrow$(v).

Proof. (i)$\Rightarrow$(ii) is in Blum—Eisenberg [2] and is stated in the introduction. (ii)$\Rightarrow$(iii) is part of the theorem in the introduction. (iii)$\Rightarrow$(iv) is from the argument at the beginning of section 2. (iv)$\Rightarrow$(v) is the corollary to the abstract Kronecker Theorem.

What is amazing about this result is that it says if the statement (i) is true then the sequence can be used to approximate functions in the Kronecker sense. (ii), (iii) and (iv) and (v) each seem very difficult to verify themselves, but (i) gives many sequences satisfying (ii) to (v). In fact, no sequences are known which satisfy (ii) but not (i).

Appendix

Lemma. *Let X be a normal topological space and μ a finite regular measure on X. If $\mu_n \to \mu$ weakly and if f is bounded and continuous a.e. $d\mu$, then $\int f d\mu_n \to \int f d\mu$.*

Proof. Take an open set Θ in X. There is a closed set $C \subset \Theta$ with $\mu(\Theta - C) \leqq C$. By Urysohn's lemma there is a continuous function r with $r=1$ on C, $|r| \leqq 1$, and $r=\Theta$ on Θ^C. $\mu_n(\Theta) \geqq \int r\,d\mu_n \to \int r\,d\mu \geqq \mu(\Theta)-\varepsilon$. Hence $\underline{\lim}\ \mu_n(\Theta) \geqq \mu(\Theta)$.

The set $A=\{a|\mu\{x|f(x)=a\}>0\}$ is countable since μ is finite. Approximate $\int f d\mu_n$ and $\int f d\mu$ by $\sum a_i \mu_n\{a_i<f<b_i\}$ and $\sum a_i \mu\{a_i<f<b_i\}$, respectively, where $|a_i - b_i| < \varepsilon$ and a_i and b_i do not belong to the countable set A.

Let $\{a_i<f<b_i\}=C_i$. If $x \in \bar{C}_i - C_i$ then either $f(x)=a_i$ or b_i or x is a point of discontinuity of f. Thus $\mu(\bar{C}_i - C_i) = 0$.

Let $x \in C_i - C_i^0$. Then $f(x) \in (a_i, b_i)$ while for each neighborhood of x there is a y with $f(y) \notin (a_i, b_i)$. Hence x is a discontinuity point of f and $\mu(C_i - C_i^0) = 0$. Thus $\mu(\bar{C}_i - C_i^0) = 0$, and $\mu(\bar{C}_i) = \mu(C_i) = \mu(C_i^0)$. But

$$\overline{\lim}\ \mu_n(\bar{C}_i) \leqq \mu(\bar{C}_i) = \mu(C_i^0) \leqq \underline{\lim}\ \mu_n(C_i^0).$$

Thus $\lim \mu_n(C_i) = \mu(C_i)$.

Hence $\sum a_i \mu_n(C_i) \to \sum a_i \mu(C_i)$. Since $\int f\,d\mu_n \sim \sum a_i \mu_n(C_i)$ and $\int f\,d\mu \sim$ $\sim \sum a_i \mu(C_i)$, it must be that $\int f\,d\mu_n \to \int f\,d\mu$.

We would like to thank G. EFROYMSON for some helpful conversations.

References

[1] P. BILLINGSLEY, *Convergence of Probability Measures*, John Wiley and Sons (New York, 1968).

[2] J. R. BLUM and B. EISENBERG, Mean Ergodic Theorems for Unitary Groups (in preparation).

[3] E. HEWITT and K. ROSS, *Abstract Harmonic Analysis*, Springer-Verlag (Berlin, 1963).

[4] Y. KATZNELSON, *An Introduction to Harmonic Analysis*, John Wiley and Sons (New York, 1968).

[5] W. J. LEVEQUE, *Topics in Number Theory*, Addison Wesley (Reading, Mass., 1956).

[6] W. RUDIN, *Fourier Analysis on Groups*, Interscience Publishers (New York, 1967).

(Received February 22, 1972)

On the Banach—Steinhaus theorem and approximation in locally convex spaces

By P. L. BUTZER, K. SCHERER and U. WESTPHAL in Aachen (FRG)

Dedicated to Professor B. Sz.-Nagy on his 60th birthday on July 29, 1973, in friendship and high esteem

1. Introduction

The Banach—Steinhaus theorem essentially states that a family of bounded operators is convergent on a whole space if and only if the operators are uniformly bounded as well as convergent on a dense subspace. It is the purpose of this note to extend part of the results of P. L. BUTZER—K. SCHERER [3], namely to give necessary and sufficient conditions upon a family of operators such that they tend to some limiting operator with a given order of approximation. This can be interpreted as the Banach—Steinhaus theorem equipped with a *rate* of convergence. The results are stated for locally convex spaces. They yield applications to weighted approximation, error estimates for quadrature formulae and the mean ergodic theorem. It is to be noted that all three applications are of a quite different structure.

2. The Banach—Steinhaus theorem with rate

Let X and Y be locally convex Hausdorff spaces with topologies generated by the families of filtrating seminorms $\{p\}$, $\{q\}$, respectively.

Let T_ϱ, $\varrho \geqq 0$, T be bounded mappings defined on X into Y such that $T_\varrho - T$ is sublinear for each $\varrho \geqq 0$, i.e.

$$q[(T_\varrho - T)(f_1+f_2)] \leqq q[(T_\varrho - T)f_1] + q[(T_\varrho - T)f_2]$$

(1)

$$q[(T_\varrho - T)(af)] = q[a(T_\varrho - T)f]$$

for each $q \in \{q\}$ and $f_1, f_2, f \in X$, $a \in \mathbf{R}$. Provided X is barrelled, the theorem of Banach—Steinhaus states: the family $\{T_\varrho f;\ \varrho \geqq 0\}$ converges to Tf in the topology of Y for each $f\in X$, i.e. for each $q\in\{q\}$ one has

(2) $$\lim_{\varrho\to\infty} q[T_\varrho f-Tf]=0 \qquad (\forall f\in X)$$

if and only if

(3, i) $\{T_\varrho;\ \varrho\geqq 0\}$ is uniformly bounded, i.e. to each $q\in\{q\}$ there exists $p\in\{p\}$ and a constant $M>0$ such that

$$\sup_{\varrho\geqq 0} q[(T_\varrho-T)f]\leqq Mp(f) \qquad (\forall f\in X),$$

and

(3, ii) $\{T_\varrho f;\ \varrho\geqq 0\}$ converges to Tf in the topology of Y for each $f\in A$, A being a total set in X.

For the Banach—Steinhaus theorem, see H. G. Garnir—M. De Wilde—J. Schmets [8, p. 453], N. Bourbaki [1, p. 27], H. H. Schaefer [14, p. 86].

In order to study the rate of convergence of the given family, it is useful to introduce a quantity in place of the classical modulus of continuity, namely a modification of the K-functional. It is defined for $t>0$, $f\in X$, $p\in\{p\}$ and $\bar p\in\{\bar p\}$ by

(4) $$K(t,f;X,A)_{p,\bar p}=\inf_{g\in A}\{p(f-g)+t\bar p(g)\},$$

where $(A,\{\bar p\})$ is a subspace of $(X,\{p\})$.

Theorem 1. *Let $(X,\{p\})$, $(A,\{\bar p\})$, $(Y,\{q\})$ be locally convex spaces with $A\subset X$. Let T_ϱ, $\varrho\geqq 0$, and T be bounded operators mapping X into Y such that $T_\varrho-T$ is sublinear for each $\varrho\geqq 0$. Then to each $q\in\{q\}$ there exist $p\in\{p\}$ and $\bar p\in\{\bar p\}$ such that*

(5) $$q[(T_\varrho-T)f]\leqq C\varphi(\varrho)K(\psi(\varrho)[\varphi(\varrho)]^{-1},f;X,A)_{p,\bar p} \qquad (\forall f\in X),$$

where $\varphi(\varrho)$ and $\psi(\varrho)$ are positive functions of ϱ, if and only if

(6, i) $$q[(T_\varrho-T)f]\leqq M\varphi(\varrho)p(f) \qquad (\forall f\in X)$$

and

(6, ii) $$q[(T_\varrho-T)f]\leqq D\psi(\varrho)\bar p(f) \qquad (\forall f\in A),$$

where C, D and M are constants independent of ϱ and f.

Proof. To establish the implication (5)⇒(6), first note that

(7, i, ii) $$K(t,f;X,A)_{p,\bar p}\leqq\begin{cases}p(f) & \forall f\in X,\\ t\bar p(f) & \forall f\in A.\end{cases}$$

Then (5) implies by (7, i, ii) upon setting $t=\psi(\varrho)[\varphi(\varrho)]^{-1}$

$$q[(T_\varrho-T)f]\leqq\begin{cases}C\varphi(\varrho)p(f), & \forall f\in X\\ C\varphi(\varrho)\psi(\varrho)[\varphi(\varrho)]^{-1}\bar{p}(f), & \forall f\in A.\end{cases}$$

This yields (6, i) and (6, ii) with $M=D=C$.

To establish the converse, in view of the sublinearity of $T_\varrho-T$ one has by (6, i, ii) for each $g\in A$

$$q[(T_\varrho-T)f]\leqq q[(T_\varrho-T)(f-g)]+q[(T_\varrho-T)g]\leqq$$
$$\leqq M\varphi(\varrho)p(f-g)+D\psi(\varrho)\bar{p}(g)\leqq$$
$$\leqq\max(M,D)\varphi(\varrho)\{p(f-g)+\psi(\varrho)[\varphi(\varrho)]^{-1}\bar{p}(g)\}.$$

Taking the infimum over all $g\in A$ one has that for all $f\in X$

$$q[(T_\varrho-T)f]\leqq\max(M,D)\,\varphi(\varrho)\,K(\psi(\varrho)[\varphi(\varrho)]^{-1},f;X,A)_{p,\bar{p}}.$$

This proves the theorem.

The sufficient direction of Theorem 1 in case X is a Banach space with $X=Y$, $T=I$, may be found in P. L. BUTZER—K. SCHERER [3]. In this case, for $\psi(\varrho)\to 0$ as $\varrho\to\infty$, condition (6, ii) is referred to as a Jackson-type inequality. In this respect note that P. O. RUNCK [13] has actually given necessary and sufficient conditions upon T_ϱ such that a Jackson-type inequality is satisfied.

In the foregoing theorem the constants C, D, M were independent of f and ϱ. In the following deeper and more theoretical version the corresponding constants C and D may depend upon the element f.

Theorem 2. *Let $(X,\{p\})$, $(A,\{\bar{p}\})$, $(Y,\{q\})$ be locally convex Hausdorff spaces such that A is continuously embedded in X, i.e. to each $p\in\{p\}$ there is $\bar{p}\in\{\bar{p}\}$ and $c>0$ with $p(f)\leqq c\bar{p}(f)$ for all $f\in A$. In addition, let X as well as A be barrelled. If T_ϱ, $\varrho\geqq 0$, and T are bounded operators mapping X into Y such that $T_\varrho-T$ is sublinear for each $\varrho\geqq 0$, then the following two assertions are equivalent:*

(8) *to each $q\in\{q\}$ there is $p\in\{p\}$ and $\bar{p}\in\{\bar{p}\}$ such that* $(\delta>0)$

$$q[(T_\varrho-T)f]=O[K(\varrho^{-\delta},f;X,A)_{p,\bar{p}}]\quad(\forall f\in X),$$

(9) *to each $q\in\{q\}$ there is $p\in\{p\}$ and $M>0$ such that*

(9, i) $$\sup_{\varrho\to 0}q[T_\varrho-T)f]\leqq Mp(f)\quad(\forall f\in X).$$

(9, ii) $$q[(T_\varrho-T)f]=O(\varrho^{-\delta})\quad(\forall f\in A).$$

Proof. (8)⇒(9): The estimate (8) together with (7, i, ii) implies ($t=\varrho^{-\delta}$)

$$q[(T_\varrho - T)f] = \begin{cases} O(1), & \forall f \in X \\ O(\varrho^{-\delta}), & \forall f \in A. \end{cases}$$

The second assertion is the required (9, ii). To obtain (9, i), apply the uniform boundedness theorem (= necessary condition of classical Banach—Steinhaus theorem) to $q[(T_\varrho - T)f] = O(1)$, all $f \in X$, noting that X is barrelled.

(9)⇒(8): A being barrelled, (9, ii) implies by the uniform boundedness principle that there exists $\bar{p} \in \{\bar{p}\}$ and $D>0$ such that condition (6, ii) of Theorem 1 holds. (6, i) being valid here by assumption, one may therefore apply Thm. 1 with $\varphi(\varrho)=1$ and $\psi(\varrho)=\varrho^{-\delta}$.

Concerning the structure of Theorem 2 in comparison with the Banach—Steinhaus theorem, the assertions (2), (3, ii) on convergence per se are replaced by the assertions (8), (9, ii) involving an order of convergence. Indeed, if A is dense in X $\lim_{t\to 0+} K(t,f;X,A)_{p,\bar{p}}=0$. This is the situation in the applications to follow.

3. Weighted approximation

The first application will be concerned with weighted approximation; it will turn out to be an actual example of approximation in a locally convex space. Here the space Y will be seen to be equal to the locally convex space X and the limit operator will be the identity. The corresponding problem was first considered by J. Kemper—R. J. Nessel [10] using classical methods.

Let E be the space of functions given on the reals $\mathbf{R}$ which are either uniformly continuous and bounded on $\mathbf{R}$ or measurable and pth power ($1 \leqq p < \infty$) integrable on $\mathbf{R}$, and let E be normed in the usual fashion. Let E_{loc} be those functions which are either continuous on $\mathbf{R}$ or p th power integrable on each compact subset of $\mathbf{R}$.

Let

$$X = \{f \in E_{\mathrm{loc}};\quad \|e^{-\beta x^2} f(x)\|_E < +\infty, \quad \forall \beta > 0\};$$

it is a locally convex Hausdorff space with respect to the family of norms

$$p_\beta(f) = \|e^{-\beta x^2} f(x)\|_E \qquad (\forall f \in E).$$

Let

$$A = \{f \in X;\, f, f' \text{ loc. abs. continuous}, f'' \in X\}$$

and

$$\bar{p}_\beta(f) = \|e^{-\beta x^2} f''(x)\|_E \qquad (\forall f \in A)$$

be a family of seminorms on A.

It is our purpose to consider the Weierstrass integral

$$(W_\varrho f)(x) = \frac{\varrho}{2\sqrt{\pi}} \int_{-\infty}^{\infty} f(x-u) \exp\left(-\frac{\varrho^2 u^2}{4}\right) du$$

for $\varrho > 0, f \in X$. This defines a family of operators W_ϱ on X into itself which converge in the topology of X towards the identity operator, i.e., for each $\beta > 0$ and each $f \in X$

$$\lim_{\varrho \to \infty} \|e^{-\beta x^2} [(W_\varrho f)(x) - f(x)]\|_E = 0.$$

Now $W_\varrho - I$ satisfies the hypotheses of Thm. 1. Indeed, taking $\varrho_0 > 0$ arbitrary fixed, it is easy to verify, using [10], that for each $\beta > 0$

$$\|e^{-\beta x^2} [(W_\varrho f)(x) - f(x)]\|_E \leqq (1+\sqrt{2}) \|e^{-\eta x^2} f(x)\|_E \qquad (\varrho \geqq \varrho_0; \forall f \in X),$$

where $\eta = 1/2 \min (\beta, \varrho_0^2/8)$. Thus (6, i) is satisfied with $M = 1+\sqrt{2}$ and $\varphi(\varrho) = 1$. Likewise with (6, ii); indeed, for each $\beta > 0$ (see [10])

$$\|e^{-\beta x^2} [(W_\varrho f)(x) - f(x)]\|_E \leqq 4\sqrt{2}\, \varrho^{-2} \|e^{-\eta x^2} f''(x)\|_E \qquad (\varrho \geqq \varrho_0; f \in A).$$

Thus one may apply Thm. 1, (6, i, ii)$\Rightarrow$(5), to get for each $\beta > 0$

$$\|e^{-\beta x^2} [(W_\varrho f)(x) - f(x)]\|_E \leqq 4\sqrt{2}\, K(\varrho^{-2}, f; X, A)_{\eta,\eta}.$$

The following lemma is of importance (compare [2, p. 192] in the case of semigroup operators)

Lemma. *Under the preceding hypotheses we have*

$$K(t^2, f; X, A)_{\eta,\beta} \leqq \tfrac{3}{2} \omega_2(t, f; X)_\zeta \qquad (t > 0; f \in X),$$

where $\zeta = \min (\eta, \beta)$ *and*

$$\omega_2(t, f; X)_\zeta = \sup_{|s| \leqq t} \|e^{-\zeta x^2} [f(x+s) + f(x-s) - 2f(x)]\|_E.$$

Proof. It is obvious that

$$f(x) = -\frac{1}{2t^2} \int_{-t/2}^{t/2} \int_{-t/2}^{t/2} [f(x+\tau_1+\tau_2) + f(x-\tau_1-\tau_2) - 2f(x)]\, d\tau_1\, d\tau_2 + g_t(x)$$

where

$$g_t(x) = \frac{1}{t^2} \int_{-t/2}^{t/2} \int_{-t/2}^{t/2} f(x+\tau_1+\tau_2)\, d\tau_1\, d\tau_2.$$

This yields, first of all,

$$p_\eta(f - g_t) \leqq \tfrac{1}{2} \omega_2(t, f; X)_\eta.$$

Furthermore, since $g_t''(x) = t^{-2}[f(x+t)+f(x-t)-2f(x)]$, we have

$$\bar{p}_\beta(g_t) \leqq t^{-2}\omega^2(t,f;X)_\beta .$$

Combining the results, the desired inequality follows immediately.

Proposition 1. *For each $\beta>0$ there is $\eta(=1/2\min(\beta, \varrho_0^2/8))$ such that*

$$\|e^{-\beta x^2}[(W_\varrho f)(x)-f(x)]\|_E \leqq 6\sqrt{2}\,\omega_2(\varrho^{-2},f;X)_\eta .$$

In particular, if $\omega_2(t,f;X)_\eta = O(t^\alpha)$ for each $\eta>0$, where $0<\alpha\leqq 2$, i.e. $f\in \mathrm{Lip}_2(\alpha;X)$, then

$$\|e^{-\beta x^2}[(W_\varrho f)(x)-f(x)]\|_E = O(\varrho^{-\alpha}) \qquad (\forall \beta>0).$$

The latter result, a direct approximation theorem, as well as its converse, is already to be found in [10].

A more interesting related application would be the approximation of an operator $T: L^p(\mathbf{R}) \to L^q(\mathbf{R})$ by bounded operators $T_\varrho: L^p(\mathbf{R}) \to L^q(\mathbf{R})$, $\varrho>0$, by some modulus of continuity, where both T and T_ϱ are singular integrals of Fourier convolution type, i.e.

$$(Tf)(x) = \int_{-\infty}^{\infty} f(x-u)\,\chi(u)\,du, \quad (T_\varrho f)(x) = \int_{-\infty}^{\infty} f(x-u)\,\chi_\varrho(u)\,du,$$

with $\chi, \chi_\varrho \in L^r(\mathbf{R})$, $p^{-1}+r^{-1} \geqq 1$ and $q^{-1} = p^{-1}+r^{-1}-1$. An open problem here would be to express conditions (6, i, ii) or (9, i, ii) of Thm. 1 or 2 equivalently in terms of the kernels χ, χ_ϱ themselves. Whereas condition (i), namely $\chi-\chi_\varrho$ being in $L^r(\mathbf{R})$ is satisfied by assumption, (ii) would be the difficult one. A solution would deliver conditions which are not only sufficient for an estimate by some modulus of continuity but also necessary.

4. Error estimates for quadrature formulae

Our general theorem enables one to deduce estimates for numerical integration formulae as was pointed out to us by Dr. H. Esser.

For $f\in C^\mu[a,b]$, the space of μ-times ($\mu=0, 1, 2, \ldots$) continuously differentiable functions on $[a,b]$, let us set (compare V. I. Krylov [11])

$$(10) \qquad Qf = \int_a^b f(x)\,dx, \quad Q_n^\mu f = \sum_{i=1}^n A_{i,n} f(x_{i,n}) + \sum_{\nu=1}^{\mu} \sum_{i-1}^{n} B_{i,n}^\nu f^{(\nu)}(x_{i,n}^\nu),$$

with given nodes $x_{i,n}, x_{i,n}^\nu \in [a,b]$ and weights $A_{i,n}, B_{i,n}^\nu$. Then Q and Q_n^μ define linear functionals on $C^\mu[a,b]$. In order to obtain an error estimate of Qf by $Q_n^\mu f$ for large n, one assumes that the quadrature formula $Q_n^\mu f \approx Qf$ is exact for polynomials p_m of fixed degree $m(\geqq\mu)$, i.e. $Q_n^\mu p_m = Qp_m$.

To apply Thm. 1 we take $X=C^{\mu}[a,b]$ and $A=C^{m+1}[a,b]$, $m\geqq\mu$, equipped with seminorms

$$p(f)=|f|_{C^{\mu}}\equiv\sup_x|f^{(\mu)}(x)|\qquad(f\in C^{\mu})$$

and

$$\bar{p}(f)=|f|_{C^{m+1}}\equiv\sup_x|f^{(m+1)}(x)|\qquad(f\in C^{m+1}),$$

respectively. In the setting of this example conditions (6, i; ii) of Thm. 1 may be rewritten as

$$|Q_n^{\mu}-Q|_{[C^{\mu},\mathbf{R}^1]}\leqq Mn^{-\mu},\quad |Q_n^{\mu}-Q|_{[C^{m+1},\mathbf{R}^1]}\leqq Dn^{-m-1}$$

with ϱ the discrete n, $\varphi(n)=n^{-\mu}$, $\psi(n)=n^{-m-1}$ and

$$|Q_n^{\mu}-Q|_{[C^l,\mathbf{R}^1]}=\sup_{\substack{f\in C^l\\ f\neq 0}}\frac{|Q_n^{\mu}f-Qf|_{\mathbf{R}^1}}{|f|_{C^l}}$$

for $l=\mu$ and $l=m+1$, respectively.

Now, in case $l\geqq 1$, these quantities may be computed with the aid of the theorem of Peano asserting that

(11) $$Q_n^{\mu}f-Qf=\int_a^b f^{(l)}(t)\,\chi_{n,l-1}^{\mu}(t)\,dt\qquad(f\in C^l[a,b];\ l=\mu,\ m+1,\ l\geqq 1),$$

where

(12) $$\chi_{n,l-1}^{\mu}(t)=\frac{1}{(l-1)!}(Q_n^{\mu}-Q)_x(x-t)_+^{l-1}$$

and

$$(x-t)_+^{l-1}=\begin{cases}(x-t)^{l-1}, & x\geqq t,\\ 0\quad, & x<t,\end{cases}$$

the index x in (12) meaning that the functional $Q_n^{\mu}-Q$ is applied to $(x-t)_+^{l-1}$ with respect to x. From (11) we obtain

$$|Q_n^{\mu}-Q|_{[C^l,\mathbf{R}^1]}=\int_a^b|\chi_{n,l-1}^{\mu}(t)|\,dt\qquad(l=\mu,m+1,l\geqq 1).$$

In case $l=0$, i.e. $\mu=0$, there holds

$$|Q_n^0-Q|_{[C,\mathbf{R}^1]}=|Q_n^0|_{[C,\mathbf{R}^1]}+|Q|_{[C,\mathbf{R}^1]}=(b-a)+\sum_{i=1}^n|A_{i,n}|.$$

Concerning (5) of Thm. 1 we may estimate the K-functional

$$K(t^{m+1-\mu},f;\ C^{\mu}[a,b],\ C^{m+1}[a,b])$$

by the $(m+1-\mu)$-th modulus of continuity (cf. H. JOHNEN [9]):

$$\omega_{m+1-\mu}(t;f^{(\mu)}) \leqq m_1 K(t^{m+1-\mu}, f;\ C^\mu[a,b],\ C^{m+1}[a,b]) \leqq$$
$$\leqq m_2\omega_{m+1-\mu}(t;f^{(\mu)}),$$

where $(m+1-\mu=r)$

$$\omega_r(t;f^{(\mu)}) = \sup_{|s|\leqq t}\left\{\sup_{x,x+rs\in[a,b]}\left|\sum_{k=0}^{r}(-1)^{r-k}\binom{r}{k}f^{(\mu)}(x+ks)\right|\right\}.$$

Combining these results one obtains

Theorem 3. *Under the above definitions*

$$|Q_n^\mu f-\int_a^b f(x)\,dx| \leqq C_{\mu,m}\, n^{-\mu}\,\omega_{m+1-\mu}(n^{-1};f^{(\mu)}) \qquad (\forall f\in C^\mu[a,b])$$

holds if and only if

$$\text{(i)}\quad \left.\begin{matrix}(\mu\geqq 1):\displaystyle\int_a^b|\chi_{n,\mu-1}^\mu(t)|\,dt\\ (\mu=0):\displaystyle\sum_{i=1}^n|A_{i,n}|\end{matrix}\right\} = O(n^{-\mu}), \qquad \text{(ii)}\quad \int_a^b|\chi_{n,m}^\mu(t)|\,dt = O(n^{-m-1}).$$

Let us note that (i) and (ii) may be verified for many examples, for instance in case $\mu=0$ for the composite Newton—Cotes formulae; cf. P. J. DAVIS—P. RABINOWITZ [6].

For such examples our result would yield error estimates for the quadrature formula $Q_n^0 f\approx Qf$ which are entirely free of derivatives. The determination of the best possible constants $C_{0,m}$ is another problem.

Derivative-free error estimates, at least in the case of functions which are analytic, were originally investigated by G. HÄMMERLIN [8a, b]. Thm. 3 may be interpreted as a result in ESSER [7] now equipped with rate. See also [7] for literature on the subject.

5. Mean ergodic theorem

This application gives part of the results obtained by P. L. BUTZER, D. LEVIATAN and U. WESTPHAL in [4, 5, 12], where the mean ergodic theorem was studied with respect to the rate of its convergence.

Let $\sigma_n^\alpha(T)$ be the Cesàro-means of order $\alpha\geqq 1$ of the iterates of a bounded linear operator T from a Banach space X (norm $\|\cdot\|$) into itself, i.e.

$$\sigma_n^\alpha(T) = \binom{n+\alpha}{n}^{-1}\sum_{i=0}^{n}\binom{n-i+\alpha-1}{n-i}T^i \quad (\alpha\geqq 1,\ n=0,1,2,\ldots).$$

If $\|T^n\|_{[X,X]} \leqq M_0$, $n=0, 1, 2, \dots$, then the mean ergodic theorem asserts

$$\lim_{n\to\infty} \|\sigma_n^\alpha(T)f - Pf\| = 0 \qquad (\forall f \in X_0),$$

where $X_0 = N(I-T) \oplus \overline{R(I-T)}$, $N(I-T)$ denoting the null space and $\overline{R(I-T)}$ the closure of the range of $(I-T)$, and P is the projection of X_0 on $N(I-T)$ parallel to $\overline{R(I-T)}$. If $T_0 = T/X_0$, define a linear operator B with domain $D(B) = N(I-T) \oplus \oplus R(I-T_0)$ and range in X_0 by $Bf=g$, where $g \in X_0$ is uniquely determined by $(I-P)f = (I-T_0)g$ and $Pg=0$.

We may then apply Thm. 1 to X_0 normed by $p(f) = \|f\|$ and $D(B)$ with seminorm $\bar{p}(f) = \|Bf\|$. Indeed, since the following inequalities are valid (compare [5, 12])

(i) $$\|\sigma_n^\alpha(T)f - Pf\| \leqq M_0(M_0+1)\|f\| \qquad (\forall f \in X_0),$$

(ii) $$\|\sigma_n^\alpha(T)f - Pf\| \leqq \frac{\alpha}{n+1}(M_0+1)\|Bf\| \qquad (\forall f \in D(B)),$$

one concludes that

$$\|\sigma_n^\alpha(T)f - Pf\| \leqq CK(n^{-1}, f; X_0, D(B)) \qquad (\forall f \in X_0).$$

Defining a generalized Lipschitz class by

$$\mathrm{Lip}\,(\delta; X_0) = \{f \in X_0;\, K(t, f; X_0, D(B)) = O(t^\delta)\},$$

one has

Proposition 2. *If* $f \in \mathrm{Lip}\,(\delta; X_0)$, $0 < \delta \leqq 1$, *then*

$$\|\sigma_n^\alpha(T)f - Pf\| = O(n^{-\delta}).$$

References

[1] N. Bourbaki, *Eléments de Mathématique*, Livre V: *Espaces vectoriels topologiques*, Chap. III: *Espaces d'applications linéaires continues*, Hermann (Paris, 1955).

[2] P. L. Butzer—H. Berens, *Semi-Groups of Operators and Approximation*, Springer (Berlin—Heidelberg—New York, 1967).

[3] P. L. Butzer—K. Scherer, Approximation theorems for sequences of commutative operators in Banach spaces, *Proc. Conf. Constructive Function Theory*, Varna, Bulgaria, 1970. Ed. B. Penkov—D. Vačov (Sofia, 1972), 137—145.

[4] P. L. Butzer—U. Westphal, The mean ergodic theorem and saturation, *Indiana Univ. Math. J.*, **20** (1971), 1163—1174.

[5] P. L. Butzer—U. Westphal, Ein Operatorenkalkül für das approximationstheoretische Verhalten des Ergodensatzes im Mittel, *Proc. Conf. Linear Operators and Approximation*, Oberwolfach, 1971. Ed. P. L. Butzer—B. Sz.-Nagy—J.-P. Kahane, ISNM Vol. 20, Birkhäuser (Basel, 1972).

[6] P. J. Davis—P. Rabinowitz, *Numerical Integration*, Blaisdell (Waltham, Mass., 1967).

[7] K.-H. ESSER, *Neue Konvergenzsätze und Konstruktionsprinzipien zur Aufstellung von Quadraturverfahren* (Dissertation, RWTH Aachen, 1970).

[8] H. G. GARNIR—M. DE WILDE—J. SCHMETS, *Analyse Fonctionnelle*. Tome I: *Théorie générale*, Birkhäuser (Basel 1968).

[8a] G. HAMMERLIN, Über ableitungsfreie Schranken für Quadraturfehler, *Num. Math.*, **5** (1963), 225—233.

[8b] G. HÄMMERLIN, Über ableitungsfreie Schranken für Quadraturfehler. II. Ergänzungen und Möglichkeiten zur Verbesserung, *Num. Math.*, **7** (1965), 232—237.

[9] H. JOHNEN, Inequalities connected with moduli of smoothness, *Mat. Vesnik*, **9 (24)** (1972), 289—303.

[10] J. KEMPER—R. J. NESSEL, Gewichtete Approximation durch variationsvermindernde Operatoren vom Faltungstyp, *Forschungsberichte des Landes Nordrhein-Westfalen* (in print).

[11] V. I. KRYLOV, *Approximate Calculation of Integrals*, Macmillan (New York, 1962).

[12] D. LEVIATAN—U. WESTPHAL, On the mean ergodic theorem and approximation, *Mathematica* (Cluj) (in print).

[13] P. O. RUNCK, Über Konvergenzgeschwindigkeit linearer Operatoren in Banach-Räumen, *Proc. Conf. Approximation Theory*, Oberwolfach, 1963. Ed. P. L. Butzer—J. Korevaar, ISNM Vol 5, Birkhäuser (Basel, 1964).

[14] H. H. SCHAEFER, *Topological Vector Spaces*, Macmillan (New York, 1966).

[15] M. DE WILDE, Ultrabornological spaces and the closed graph theorem, *Bull. Roy. Sci. Liège*, **40** (1971), 116—118.

LEHRSTUHL A FÜR MATHEMATIK
TECHNOLOGICAL UNIVERSITY
AACHEN, GERMANY

(Received October 6, 1972)

Translation invariant transformations of integration spaces

By J. L. B. COOPER in London (Great Britain)*)

Dedicated to Professor Béla Szőkefalvi-Nagy on his sixtieth birthday

1. Introduction

A number of writers (e.g. [1—5]) have dealt with the existence and properties of linear transformations between function spaces obeying various functional equations. In almost all cases these equations are of a type that I have termed "appropriate" in my article [1]; this term will now be defined.

Let $A(X)$ be a space of functions defined on a set X. A linear transformation W on $A(X)$ to itself is called appropriate if for each x in X and f in $A(X)$ the value of $Wf(x)$ depends exclusively on the value of f at some point in X, say Vx, or (in the case of spaces of functions defined only up to sets of measure zero) if a similar statement is true in the limit for functionals on $A(X)$ whose support tends to x. Equivalently, W is an appropriate transformation if $Wf(x)=Q(x)f(Vx)$. An appropriate group is a representation of a group by a group of appropriate transformations.

A linear operator T from a space $A(X)$ to a space $B(U)$ of functions on X and U respectively is said to obey an appropriate functional equation if it is an intertwining operator between appropriate groups of transformations on $A(X)$ and $B(U)$: that is to say, if there is a group G represented by appropriate groups $W(g)$ and $W^*(g)$ of transformations on $A(X)$ and $B(U)$ and if T obeys

$$TW(g) = W^*(g)T, \qquad g\in G. \tag{1.1}$$

In [1] I have shown that if X and U are intervals of the real line, and G is the additive group of reals, then after possible splitting of X and U into intervals invariant in the groups $V(g)$ and $V^*(g)$ and changes of variable in X and U all ap-

*) My thanks are due to the Mathematics Department, University of Toronto for an appointment as Distinguished Visiting Scientist during the tenure of which this work was partly carried out.

propriate functional equations can be reduced to four canonical forms, of which the most important are:

I. a $$T\left\{\frac{p(x+h)}{p(x)} f(x+h)\right\}(u) = \frac{q(u+h)}{q(u)} \{Tf(x)\}(u+h),\ u,\ h \in R,$$

II. a $$T\left\{\frac{p(x+h)}{p(x)} f(x+h)\right\}(u) = e^{hq(u)} \{Tf(x)\}(u),\ u \in E,\ h \in R$$

for some set $E \subset R$.

If we replace $A(X)$ by $pA(X)$ and $B(U)$ by $qB(U)$ then I.a becomes

I. $$[Tf(x+h)](u) = [Tf(x)](u+h),\ u,\ h \in R$$

and if we make only the first of these transformations I.a. becomes

II. $$[Tf(x+h)](u) = e^{hq(u)} (Tf(x)](u),\ u \in E,\ h \in R.$$

If X and U are sets in R^k, $k > 1$, and G is the group of translations of R^k, then the situation is considerably more complicated and there can be many more essentially distinct forms of appropriate functional equations. The problem of classifying these may be of interest; but it is complicated by the fact that there can be, for example, periodic representations of the translation group. It remains of interest to study equations I and II; I in particular has received considerable attention, notably in the work of HÖRMANDER [5], in which the existence of solutions mapping a space L_p onto a space L_q with respect to Lebesgue measure on R^k is studied.

The reduction of the equation Ia to the form I for transforms over integration spaces involves a change in the measure on the space; and it is therefore of interest to investigate the equation for spaces $L_p(\mu)$ and $L_q(\nu)$ with general Radon measures.

Writing $\tau(h)f(x) = f(x-h)$, the equation I becomes $T\tau(h) = \tau(h)T$. In general there is a difficulty in interpreting this equation; for the translation operator $\tau(h)$ does not necessarily map $L_p(\mu)$ to itself, and if it does not the meaning of the equation is unclear. We therefore start our investigation by studying what properties on μ ensure that $\tau(h)$ is always defined, and also some other properties of $\tau(h)$ that simplify the structure of μ. The following sections then give conditions that are necessary for the existence of nonzero T satisfying the equation I, and also discuss some properties of the solutions.

2. Conditions for existence of $\tau(h)$

Notation. If μ is a positive Radon measure on R^k, $\mathscr{L}_p(\mu)$ is the space of all μ measurable functions $f(x)$ such that $|f(x)|^p$ is μ summable, and $L_p(\mu)$ is the corresponding Banach space of functions modulo null functions. We write m for Lebesgue measure, $m(dx)=dx$, and if $\mu=\lambda m$, that is $\mu(dx)=\lambda(x)dx$, we write $L(p,\lambda)$ for $L_p(\mu)$.

Theorem 1. *Let μ be a positive Radon measure in R^k. The following conditions on μ, for any p, $1\leqq p<\infty$, are equivalent:*

(a) *If $f\in\mathscr{L}_p(\mu)$ then $\tau(h)f\in\mathscr{L}_p(\mu)$ for all $h\in R^k$;*

(b) *if f and g are in the same equivalence class in $\mathscr{L}_p(\mu)$ so are $\tau(h)f$ and $\tau(h)g$ for any $h\in R^k$;*

(c) *for any h, $\tau(h)$ takes $L_p(\mu)$ into itself;*

(d) *$\tau(h)$ is a continuous map of $L_p(\mu)$ to itself for any h;*

(e) *there is a positive Lebesgue measurable function $\lambda(x)$, bounded with $\lambda(x)^{-1}$ over any compact set of values of x, such that $\lambda(x)dx=\mu(dx)$ and $\|\tau(h)\|_p^p=\sup\dfrac{\lambda(x+h)}{\lambda(x)}$ is bounded over any compact set of values of h.*

(b) and (c) are clearly equivalent, and imply (a). Now let f and g be equivalent in $\mathscr{L}_p(\mu)$ and let $r(x)=0$ if $f(x)=g(x)$, $r(x)=\infty$ otherwise; then r is a μ null function, and so in $\mathscr{L}_p(\mu)$. If $\tau(h)f(x)\neq\tau(h)g(x)$, then $\tau(h)r(x)=\infty$; thus if (a) holds, the set with $\tau(h)r(x)=\infty$ is a null set an so $\tau(h)f(x)=\tau(h)g(x)$ almost everywhere: thus (a) implies (b) and (c).

Let us now write $\tau(h)\mu=\mu_h$, that is

$$\int f(x)\,\mu_h(dx) = \int f(x+h)\,\mu(dx).$$

Our arguments show that (a), (b) or (c) imply that μ is quasi-invariant [7]: μ null sets translate into μ null sets and μ_h is absolutely continuous with respect to μ, so that there is for each h a function $\varphi(x,h)$ nonnegative and μ summable over any set of finite μ measure as a function of x such that $\mu_h(dx)=\varphi(x,h)\mu(dx)$; since μ is absolutely continuous with respect to μ_h it follows that $1/\varphi(x,h)$ is also summable over any set with finite μ_h and so finite μ measure. For any f and h $\int|f(x)|^p\mu(dx)$ is finite if and only if $\int|f(x)|^p\varphi(x,h)\mu(dx)$ is finite. Let $\varphi_n(x,h)=\min(\varphi(x,h),n)$; then the map $f(x)\to\varphi_n(x,h)f(x)$ is bounded on $L_p(\mu)$ to itself for any fixed n and h and the set $\{\varphi_n(x,h)f(x);\ n=1,2,\dots\}$ is bounded in $L_p(\mu)$ for each f; by the Banach—Steinhaus Theorem it is uniformly bounded, and so $\varphi(x,h)$ is bounded for each h. This proves that $\tau(h)$ is a bounded transformation of $L_p(\mu)$ to itself; if $K(h)=\text{ess sup}\,\{\varphi(x,h):x\in R^k\}$ then $\|\tau(h)\|=K(h)^{1/p}$, $\log\|\tau(h)\|=p^{-1}\log K(h)=p^{-1}L(h)$

say, and since $\tau(h)\tau(h') = \tau(h+h')$ it follows that $L(h)$ is everywhere finite, measurable and subadditive. We now show that $L(h)$ is bounded over any compact set of h. If it is not bounded above, then there is a convergent sequence (h_n) for which $L(h_n) \to \infty$; since, for any h, $L(h+h_n) \geqq L(h_n) - L(-h)$, we can find such a sequence convergent to any assigned h; and we suppose that h has each coordinate h^j positive. For any such h let $C(h) = \{x: 0 \leqq x^j \leqq h^j,\ j=1,2,\dots,k\}$. Since $L(h_n) \leqq L(x) + L(h_n - x)$, either $L(x) > \frac{1}{2}L(h_n)$ or $L(h-x) \geqq \frac{1}{2}L(h_n)$ holds for any given x; and so $m\{x: L(x) \geqq L(h_n);\ x \in C(h_n)\} \geqq \frac{1}{2}mC(h_n)$. We can choose the h_n so that $h_n^j \sim 0$ for all j and n, and $L(h_n) > n$; then the set $\{x: L(x) > \frac{1}{2}n,\ x \in C(h_n)\}$ has for each n Lebesgue measure greater than $\frac{1}{2}mC(h_n) \to \frac{1}{2}mC(h)$ and hence $m\{x: L(x) = \infty,\ x \in C(h)\} \geqq \frac{1}{2} \geqq mC(h)$; but this set is empty, so we have a contradiction. If L is not bounded below on a set C, it is not bounded above on $-C$; hence L is bounded on every compact.

Now by the Lebesgue—Vitali Theorem (e.g. [6], Vol. I, Theorem III. 12. 6)

$$\lim_{r \to 0} \frac{B(a,r)}{mB(a,r)} = \lambda(a) \text{ say}$$

exists for almost all a and is finite. Choose a in $B(0,r)$ so that this holds. Any set E in $B(0,r)$ can be covered by a finite number of translates of a ball $B = B(a,r)$ with $\mu(B) < (\lambda(a)+1)m(B)$, and hence $\mu(E) < M(\lambda(a)+1)m(E)$, where $\log M$ is the upper bound of $L(h)$ in $B(0, 2r)$. Thus μ is absolutely continuous with respect to Lebesgue measure $m: \mu(dx) = \lambda(x)\,dx$ with λ bounded over any compact. Since then $\mu_h(dx) = \lambda(x+h)\,dx$, $\varphi(x,h) = \dfrac{\lambda(x+h)}{\lambda(x)}$ $\|\tau(h)\|^p = \sup \dfrac{\lambda(x+h)}{\lambda(x)}$ is bounded over every compact, and so is $\|\tau(h)\|^{-1}$.

This proves that (a) implies (d) and (e). On the other hand, it is easy to see that (d) or (e) imply (a).

It is essential for the truth of this theorem that $p < \infty$. If $p = \infty$, then for (a) to hold μ must be quasi-invariant, and if this is the case then $\|\tau(h)f\| = \|f\|$ for any f and h. We can conclude again that $\mu(dx) = \lambda(x)\,dx$ with λ locally summable, but not that λ is necessarily bounded over a compact or restricted in growth.

In future we write $L(p, \lambda)$ for $L_p(\mu)$ when $\mu(dx) = \lambda(x)\,dx$. We define $l(x) = l(\lambda, x) = \log \lambda(x)$, $L(h) = L(\lambda, h) = \operatorname{ess\,sup}\,[l(x+h) - l(x)]$.

Theorem 2. *Let μ obey the conditions of Theorem 1 for some p, $1 \leqq p < \infty$. Then $L(\lambda, h)$ is subadditive and for any h*

$$F(\lambda, h) = \lim_{\alpha \to \infty} \frac{L(\lambda, \alpha h)}{\alpha} \tag{2.1}$$

exists. $F(h) = F(\lambda, h)$ is a continuous convex positive homogeneous function, everywhere finite: $F(\lambda, h) \geqq -F(\lambda, -h)$.

If $\|\tau(h)\|$ is continuous in h, then $\lambda(x)$ is continuous and $l(x)$ and $L(h)$ are uniformly continuous, and $L(h)$ is continuous. The limit in (2. 1) exists uniformly in h over the sphere $\|h\|=1$.

Theorem 1 shows that $L(h)$ exists and is finite everywhere; it is obviously subadditive. Consequently

$$\frac{L((\alpha+\beta)h)}{\alpha+\beta} \leqq \frac{\alpha}{\alpha+\beta}\frac{L(\alpha h)}{\alpha}+\frac{\beta}{\alpha+\beta}\frac{L(\beta h)}{\beta}$$

for any α, β and so $L(\alpha h)/\alpha$ decreases with α for positive α, and this proves that $F(h)$ exists and is less than ∞. Since $L(\alpha h)+L(-\alpha h) \geqq L(0) = 0$, it follows that $F(h)+F(-h) \geqq 0$ and hence $F(h) \geqq -F(-h) \geqq -\infty$. Clearly $F(\beta h)=\lim L(\alpha\beta h)/\alpha= =\beta F(h)$ for any positive β and $F(h+k) \leqq \lim L(\alpha h)/\alpha+\lim L(\beta h)/\beta = F(h)+F(k)$. It follows that F is positive homogeneous and subadditive, hence that it is convex and so continuous.

Since $p \log \lambda(h) = \operatorname{ess\,sup} \{l(x+h)-l(x)\}$ continuity of $\lambda(h)$ implies that the righthand side tends to 0 as $h\to 0$, that is

$$\text{(2. 2)} \qquad \operatorname{ess\,sup} [\lambda(x+h)/\lambda(x)] \to l \text{ as } h \to 0$$

$\lambda(x)$ is everywhere equal to $\lim_{r\to 0} \mu B(x, r)/mB(x, r)$ that is, to

$$\lim \int_{\|x-y\|<r} \lambda(y)\, dy/mB(x, r).$$

The corresponding integral for $\lambda(x+h)$ has y replaced with $y+h$ and so since $\lambda(x)$ is bounded over any compact and because of (2. 2), λ is continuous. $l(x)$ is then also continuous and so $\sup [l(x+h)-l(x)] = p \log \|\tau(h)\| \to 0$ as $h\to 0$, that is, $l(x)$ is uniformly continuous. Uniform continuity of $L(h)$ follows immediately: if $\delta>0$ is such that $l(x+h)-l(x) < \varepsilon$ when $\|h\|<\delta$ then $|L(h)-L(h')| < \varepsilon$ if $|h-h'| < \delta$.

If the limit in (2. 1) does not exist uniformly over $\|h\|=1$, then for some $\varepsilon>0$ we can select a sequence (h_n), convergent to a point h on the sphere, and indeed such that $\|n(h-h_n)\| \to 0$ as $n\to\infty$, so that for all n,

$$L(nh_n) > n[F(h_n)+3\varepsilon].$$

For each n we can choose x_n such that $l(x_n+nh_n)-l(x_n) > L(nh_n)-\varepsilon$, and then

$$n[F(h_n)+3\varepsilon] < l(x_n+nh_n)-l(x_n) = l(x_n+nh_n)-l(x_n+nh)+l(x_n+nh)-l(x) \leqq$$
$$\leqq L(nh_n-nh)+L(nh) < L(nh_n-nh)+n[F(h)+\varepsilon].$$

For some n_0, $L(nh_n-nh) < \varepsilon$ and $|F(h_n)-F(h)| < \varepsilon$ if $n>n_0$, and then the lefthand side is greater than $n(F(h)+2\varepsilon)$ and the righthand side is less than

$\varepsilon + n[F(h)+\varepsilon]$, so that the inequality cannot hold for all n. This proves the uniform convergence in (2. 1).

Theorem 3. *Let $F(x)$ be the positive homogeneous convex function defined by* (2. 1). *Then*

(a) *for any h and any $\varepsilon > 0$ there is an α_0 so that if $\alpha > \alpha_0$*

$$\exp\left[-\alpha\big(F(-h)+\varepsilon\big)\right] < \lambda(\alpha h) < \exp\left\{\alpha\big(F(h)+\varepsilon\big)\right\};$$

(b) *if $\|\tau(h)\|$ is continuous in h then this holds uniformly over $\|h\|=1$, that is, there is an α_0 so that if $\|x\| > \alpha_0$*

$$\exp\left[-\big(F(x)+\varepsilon\|x\|\big)\right] < \lambda(x) < \exp\big(F(x)+\varepsilon\|x\|\big).$$

By definition, $-L(-\alpha h) < l(\alpha h) - l(0) < L(\alpha h)$ for all positive α and all h, so that for large enough α

$$-\alpha[F(-h)+\varepsilon] < l(\alpha h)-l(0) < \alpha[F(h)+\varepsilon]$$

and the first statement follows; the second is a consequence of this and of the uniform convergence of $L(\alpha h)$ to $F(h)$ over the unit sphere.

The problem of the existence of intertwining operators between representations of $\tau(h)$ on integration spaces is clearly linked with the topological properties of the group $\tau(h)$ acting on these spaces: and more precise results can be found if the behaviour of $\tau(h)$ is more closely specified.

For any $R > 0$ let us write $L(p, \lambda, R)$ for the set of functions in $L(p, \lambda)$ whose supports are in $B(0, R)$. We examine conditions under which the action of $\tau(h)$ is closely approximated by its action on $L(p, \lambda, R)$.

If, for each $\varepsilon > 0$, there is an R such that, for any h

$$\sup\left\{\frac{\|\tau(h)f\|}{\|\tau(h)\|\,\|f\|} : f \in L(p, \lambda, R)\right\} \geqq (1-\varepsilon), \tag{2.3}$$

then we say that τ is compactly approximated: and if there is an R so that

$$\sup\left\{\frac{\|\tau f\|}{\|\tau(-h)f\|\,\|\tau(h)\|} : f \in L(p, \lambda, R)\right\} \geqq (1-\varepsilon) \tag{2.4}$$

then τ is inversely compactly approximated.

Theorem 4. *If τ is compactly approximated then for any $\varepsilon > 0$ there is for any h an $\alpha_0(h, \varepsilon)$ and a constant A so that if $\alpha > \alpha_0$*

$$Ae^{F(\alpha h)} \leqq \lambda(\alpha h) \leqq e^{F(\alpha h + \varepsilon\alpha)},$$

and if $\|\tau(h)\|$ *is continuous in h, there is an* $\alpha_0(\varepsilon)$ *so that if* $\|x\|>\alpha_0$

$$Ae^{F(x)}\leqq\lambda(x)\leqq e^{F(x)+\varepsilon\|x\|}.$$

If τ *is inversely compactly approximated, then these equations become*

$$e^{-(F(-\alpha h)+\varepsilon\alpha)}\leqq\chi(ah)\leqq Ae^{-F(-\alpha h)},$$

$$e^{-(F(-x)+\varepsilon\|x\|)}\leqq\lambda(x)Ae^{-F(-x)},$$

respectively.

It is easy to see that $\sup\left\{\frac{\|\tau(h)f\|}{\|f\|}:f\in L(p,\lambda,R)\right\}$ is equal to

$$\sup\{[\lambda(x+h)/\lambda(x)]^{1/p};\ \|x\|\leqq R\}$$

and so that τ is compactly approximated if and only if there is for each $\varepsilon>0$ an R such that, for all h

$$L(h)\geqq\sup\{l(x+h)-l(x):\|x\|<R\}\geqq L(h)-\varepsilon.$$

In that case one has on the one hand that $l(\alpha h)-l(0)\leqq L(\alpha h)$ and on the other that

$$l(\alpha h)=l(\alpha h)-l(x+\alpha h)+l(x)+l(x+\alpha h)-l(x)\leqq$$

$$\leqq-L(-x)+l(x)+l(x+\alpha h)-l(x)\leqq\sup\{l(x)-L(-x):\|x\|<R\}+L(\alpha h)+\varepsilon=$$

$$=A'+L(\alpha h)-\varepsilon,$$

say. Now for $\alpha>\alpha_0(h,\varepsilon)$, $\alpha(F(h)+\varepsilon)\geqq L(\alpha h)\geqq\alpha F(h)$ so that $F(h)+A'/\alpha\leqq$ $\leqq l(\alpha h)/\alpha\leqq l(0)/\alpha+F(h)+\varepsilon$, and so

$$Ae^{F(\alpha h)}\leqq\lambda(\alpha h)\leqq e^{F(\alpha h)+\varepsilon\alpha}.$$

The second inequalities follow from the uniform convergence of $L(\alpha h)/\alpha$ if $\|\tau(h)\|$ is continuous.

The last two inequalities follow by a similar argument, based on the observation that

$$\sup\left\{\frac{\|\tau f\|}{\|\tau(-h)f\|}:f\in L(p,\lambda,R)\right\}=\sup\{[\lambda(x)/\lambda(x-h)]^{1/p}:\|x\|<R\}.$$

The importance of these results to our later arguments is that they give conditions under which the growth of $\lambda(x)$ as $\|x\|\to\infty$ is regular. The approximations to $\lambda(x)$ in these formulae give examples of the E functions defined in the following.

Definition. $E(\lambda, h)$ is an upper E function for λ if $\lambda(x+h)/\lambda(x)E(\lambda, h)$ is bounded for all x and h and if for all x

$$\limsup \frac{\lambda(x+h)}{\lambda(x)E(\lambda, h)} \leqq 1;$$

as h tends to infinity along any ray;

$E(\lambda, h)$ is a lower E function for λ if

$$\liminf \frac{\lambda(x+h)}{\lambda(x)E(\lambda, h)} \geqq 1$$

as h tends to infinity along any ray.

Theorem 5. *If $E(\lambda, h)$ is an upper E function for λ, then for any $f \neq 0$*

$$\limsup \frac{\|\tau(h)f\|}{\|f\|\, E(\lambda, h)^{1/p}} \leqq 1, \qquad \limsup \frac{\|f+\tau(h)f\|}{\|f\|\,(1+E(\lambda, h))^{1/p}} \leqq 1$$

$$\limsup \frac{\|\tau(h)f+\tau(-h)f\|}{\|f\|\,(E(\lambda, h)+E(\lambda, -h)^{1/p})} \geqq 1,$$

as h tends to infinity along any ray. If $E(\lambda, h)$ is a lower E function then the same inequalities hold with signs reversed and lim sup *replaced by* lim inf.

For any h we have

$$\frac{\|\tau(rh)f\|^p}{E(\lambda, rh)} = \int |f(x)|^p \frac{\lambda(x+rh)}{\lambda(x)E(\lambda, rh)} \lambda(x)\, dx \leqq \int |f(x)|^p S(x, h, R)\, dx$$

when $r > R$, if $S(x, h, R) = \sup\{\lambda(x+rh)/\lambda(x)E(\lambda, rh) : r > R\}$. Since this is bounded and has a limit not greater than 1 as $R \to \infty$,

$$\limsup_{r\to\infty} \frac{\|\tau(rh)f\|^p}{E(\lambda, rh)} \leqq \|f\|^p$$

and this proves the first statement. To prove the second and third, we note that for any $\varepsilon > 0$ we can find f_1 such that $f_2 = f - f_1$ has norm less than ε and f_1 has compact support. Then if h is large enough the supports of f_1, $\tau(h)f_1$ and $\tau(-h)f_1$ are disjoint, so that

$$\|f_1 + \tau(h)f_1\|^p = \|f_1\|^p + \|\tau(h)f_1\|^p,$$

$$\|\tau(h)f_1 + \tau(-h)f_1\|^p = \|\tau(h)f_1\|^p + \|\tau(-h)f_1\|^p.$$

Thus, for h sufficiently large in any direction

$$\|f+\tau(h)f\| \leqq \|f_1+\tau(h)f_1\|+\|f_2\|+\|\tau(h)f_2\| \leqq$$
$$\leqq \|f\|\left(1+E(\lambda,h)\right)^{1/p}(1+\varepsilon)+\varepsilon+KE(\lambda,h)^{1/p}\varepsilon$$
$$\leqq \|f\|\left(1+E(\lambda,h)\right)^{1/p}(1+\varepsilon)+K\varepsilon\left(1+E(\lambda,h)\right)^{1/p},$$

where $K=\sup\{\lambda(x+h)/E(\lambda,h)\lambda(x)\}$, and this leads to the second inequality. The third inequality follows in the same way.

The inequalities for a lower E function follow similarly; but the proof relies on Fatou's lemma, and does not need boundedness.

3. Existence of continuous translation invariant operators

We now consider conditions on $\lambda(x)$ and $\mu(x)$ that are necessary for there to be a nonzero continuous linear T mapping $L(p,\lambda)$ to $L(q,\mu)$ and obeying

(3.1) $$[Tf(x+h)](u)=[Tf(x)](u+h),\ u,h\in R.$$

Our first results depend on the following general theorem, which includes many of the special criteria that have been used in such problems.

Theorem 6. *Let S be a directed set and for all s in S let $V(s)$ and $W(s)$ be bounded linear operators on normed spaces A and B respectively, mapping each space to itself, and let $v(s)$ and $w(s)$ be positive valued functions such that, for all $f\in A$ and $g\in B$*

(3.2) $$\limsup_s \frac{\|V(s)f\|_A}{v(S)\|f\|_A}\leqq 1,\qquad \liminf_s \frac{\|W(s)g\|_B}{w(s)\|g\|_B}\geqq 1.$$

Then if there is a continuous nonzero linear $T\colon A\to B$ such that $TV(s)=W(s)T$ for all $s\in S$ we must have

(3.3) $$\liminf_s \frac{v(s)}{w(s)}\geqq 1.$$

These statements remain true for a general set S if lim sup *and* lim inf *are replaced by* sup *and* inf *respectively throughout.*

If a T obeying the conditions exists, then for any f and $\varepsilon>0$ there is an element $s(f,\varepsilon)$ after which

$$\frac{\|TV(s)f\|_B}{w(s)\|T_1\|_B}\geqq 1-\varepsilon,\qquad \frac{\|V(s)\|_A}{v(s)\|f\|_A}\geqq 1+\varepsilon,$$

so that

$$\frac{\|Tf\|_B}{\|f\|_A}\leqq \frac{\|TV(s)f\|_B}{V(s)\|f\|_V}\,\frac{v(s)}{w(s)}\,\frac{1+\varepsilon}{1-\varepsilon}\leqq T\,\frac{1+\varepsilon}{1-\varepsilon}\,\frac{v(s)}{w(s)}.$$

We can choose f so that the lefthand side is greater than $1-\varepsilon\|T\|$ and this leads to (3. 3).

With Theorem 5, this leads to the following criteria.

Theorem 7. *Let* $E(\lambda, h)$, $E(\mu, h)$ *be, respectively, upper and lower E functions for* λ *and* μ. *In order that a nonzero continuous solution of* (3. 1) *exist it is necessary that*

$$\text{(3. 4)} \qquad \liminf \frac{E(\lambda, h)^{1/p}}{E(\mu, h)^{1/q}} \geqq 1, \quad \liminf \frac{(1+E(\lambda, h))^{1/p}}{(1+E(\mu, h))^{1/q}} \geqq 1,$$

$$\text{(3. 5)} \qquad \liminf \frac{(E(\lambda, h)+E(\mu, -h))^{1/p}}{(E(\mu, h)+E(\mu, -h))^{1/q}} \geqq 1$$

as h tends to infinity in any direction.

These result follow on taking $V(h)$ and $W(h)$ to be $\tau(h)$, $1+\tau(h)$, $\tau(h)+\tau(-h)$ respectively, and then applying Theorems 5 and 6.

Important cases are those with $\lambda(x)=e^{a\|x\|}$ (or $(1+\|x\|)^a$) and $\lambda(x)= =e^{b\|x\|}$(or $(1+\|x\|)^b$). $E(\lambda, h)$ and $E(\mu, h)$ can then be taken to be $\lambda(h)$ and $\mu(h)$ respectively; and we find that for a solution it is necessary that $qa-pb \geqq 0$, using the first inequality in (3. 4). If $a=b=0$ the first inequality gives no result: the second then shows that we must have $p \leqq q$, a result due to Hörmander ([5], Theorem 1. 1). If $qa-pb=0$, (3. 5) shows that we must have $p \leqq q$.

Sufficiency of the conditions. The conditions given in Theorem 7 are not usually sufficient for the existence of solutions. Somewhat stronger conditions are sufficient if $p \geqq q$, as the following theorem shows.

Theorem 8. *In order that the identity be a continuous imbedding of* $L(p, \lambda)$ *into* $L(q, \mu)$ *it is necessary and sufficient that* $p \geqq q$ *and that* $\lambda^{-q}\mu^p \in L^{1/(p-q)}$.

For the imbedding to be continuous we must have

$$\text{(3. 6)} \qquad \left(\int f^q \mu d\chi\right)^{1/q} \leqq K\left(\int f^p \lambda d\chi\right)^{1/p}$$

for all f and some fixed constant K.

Take $f(x)=\|x\|^{-\alpha}\chi_R(x)$, where χ_R is the characteristic function of the ball $B(0, R)$. Then if $p<k/\alpha<q$ the lefthand side of (3. 6) is infinite, and the righthand side finite; hence for (3. 6) to hold it is necessary that $p \geqq q$. Now take $f=(\mu/\lambda)^{1/(p-q)}\chi_R$. Substituting in (3. 6) gives

$$\int_{\|x\|<R} (\lambda^{-q}\mu^p)^{1/(q-p)}\, dx \leqq K^{\frac{pq}{p-q}},$$

driving the necessity of the second condition.

On the other hand, if the conditions hold, then by Hölder's inequality, for any $f \in L(p, \lambda)$

$$\left(\int f^q \mu dx\right)^{1/q} \leqq \left(\int f^p \lambda dx\right)^{1/p} \left(\int \mu^{p/(p-q)} \lambda^{-q/(p-q)} dx\right)^{\frac{p-q}{p}},$$

so that the conditions are sufficient.

If $\lambda(x) = e^{a\|x\|}$, $\mu(x) = e^{b\|x\|}$, the condition becomes $qa - pb > 0$. If $\lambda(x) = (1+\|x\|)^a$, $\mu(x) = (1+\|x\|)^b$, they become $qa - pb > k(p-q)$.

The study of the sufficiency of the conditions in other cases depends on examining more complicated transforms such as the Hilbert transforms or Riesz potentials.

If T is a bounded translation invariant operator from $L(p, \lambda)$ to $L(q, \mu)$ then there is a unique distribution $k \in D'$ such that

$$Tu = k * u, \quad u \in D.$$

If $\lambda(x) = O\|x\|^m$ as $\|x\| \to \infty$, for some m, then k is in S'.

Here D is Schwartz' space of infinitely differentiable functions with compact support, S his space of functions with the seminorms $\sup \{\|x\|^r |D^j f(x)|\}$, and D', S' the corresponding duals. The argument is close to that of Hörmander, Theorem 1. 2; modified to allow for the fact that S need not be in $L(p, \lambda)$.

For any $u \in D$ and any differential operator D^j, $TD^j u = D^j Tu$. Since $D^j u \in L(p, \lambda)$, $TD^j u \in L(q, \lambda)$ and so by the Sobolev imbedding theorem (cf. [5] Lemma 1. 1) $Tu(x)$ is continuous after correction on a set of measure zero and (after correction)

$$Tu(0) \leqq c \sum_{|j| \leqq k} \Big(\int_{\|x\|<1} |D^j u|^p dx\Big)^{1/p} \leqq c' \sum_{j \leqq k} \|D^j u\|_{p,\lambda}$$

for some constants C, C'. If u tends to 0 in the distribution sense, so does $\|D^j u\|_{p,\lambda}$ and hence $u \to Tu(0) \in D'$; it follows, that for some distribution k, $Tu(0) = (k * u)(0)$ and hence, by translation invariance, that $Tu = k * u$.

If, for some m, $\lambda(x) = O(\|x\|^m)$ then if $u \in S$, $D^j u \in L(p, \lambda)$ for any j, and the argument above goes through with S and S' replacing D and D'.

4. Scope of a transformation

The map $u_g: f \to \langle f, g\rangle = \int fg\, dx$ is an element of the dual of $L(p, \lambda)$ if $g\lambda^{-1}$ is in $L(p', \lambda)$ and its norm is the norm of $g\lambda^{-1}$ in that space, that is $\|u(g)\| = \left(\int |g|^{p'} \lambda^{-p'/p} dx\right)^{1/p'}$. Writing $p'' = 1/(p-1) = p'/p$, we see that g is an element of $L(p', \lambda^{-p''})$ and that $\|u(g)\|$ is the norm of g as an element of that space.

Now let T map $L(p, \lambda)$ to $L(q, \mu)$. The dual of the latter space is represented by elements $h \in L(q', \mu^{-q'})$, and for such an h $\langle Tf, h\rangle$ is continuous in f, so that there is a $T'h$ in $L(p', \lambda^{-p'})$ such that $\langle Tf, h\rangle = \langle f, T'h\rangle$.

T is represented by a distribution in D': $Tf = kf$ if $f \in D$. We have already pointed out that T is in S' if S is contained in $L(p, \lambda)$, that is to say, if $\lambda(x)$ is of not more than polynomial growth as $\|x\| \to \infty$. The result we have just proved enables us to show that this also holds if $\mu(x)^{-1}$ is of not more than polynomial growth.

If f and h are in D then

$$\langle Tf, h\rangle = \langle k*f, h\rangle = \langle f, \check{k}*h\rangle$$

where $\check{k}(x) = k(-x)$. It follows that $h \to \check{k}*h$ maps $L(q', \mu^{-q'})$ to $L(p', \lambda^{-p'})$ and if $\mu^{-q'}$ is of not more than polynomial growth $\check{k}$ and so k is in S'.

We sum up and extend these results in the following theorem.

Theorem 9. *Let T be a continuous transformation linear from $L(p_0, \lambda_0)$ to $L(q_0, \mu_0)$ with norm M_0 and from $L(p_1, \lambda_1)$ to $L(q, \mu_1)$ with norm M_1. For $0 \leqq t \leqq 1$ let*

$$\frac{1}{p_t} = \frac{1-t}{p_0} + \frac{t}{p_1}, \quad \frac{1}{q_t} = \frac{1-t}{q_0} + \frac{t}{q_1}, \quad \lambda_t^{1/p_t} = \lambda_0^{(1-t)/p_0} \lambda_1^{t/p_1}$$

$$\mu_t^{1/p_t} = \mu_0^{(1-t)/p_0} \mu_t^{1/p_1}.$$

Then T is a continuous translation invariant map from $L(p_t, \lambda_t)$ to $L(q_t, \mu_t)$ and from $L(q_t', \mu^{-q_t''})$ to $L(p_t', \lambda^{-p_t''})$ with norm not greater than $M_0^{1-t} M_1^t$.

If, for any t, λ_t and μ_t or their reciprocals grow at infinity not faster than a polynomial, then T is represented by a convolution with a distribution k in S'.

This theorem follows from the previous arguments together with the theorem of Stein and Weiss on interpolation of operators with change of measure [9]. The importance of the last result is that it extends the range of transformations to which the Fourier transform methods of the next section apply.

5. Fourier transforms of solutions

Explicit characterizations of transforms can be obtained by Fourier transform methods for transforms acting between spaces on both of which $\|\tau(h)\|$ is a continuous function of h and τ is either compactly approximated or inversely compactly approximated on both. Theorem 4 shows that in that case $\lambda(x)$ and $\mu(x)$ are approximated by functions of the form $\exp \pm F(x)$ up to multiples of order $\exp \varepsilon \|x\|$. We shall suppose in what follows that λ and μ are exactly of these forms. The results that follow apply without change if λ and μ are of this form up to multiplication by functions that are bounded above and below by positive constants. In the general case the arguments need some modifications, but will go through if the hypotheses in Theorem 10 that $C(\mu)/q$ is contained in $C(\lambda)/p$ is replaced by the hypothesis that $C(\mu)/q$ is in the interior of $C(\lambda)/p$.

Let us suppose that $\lambda(x)=\exp F(\lambda, x)$, $\mu(x)=\exp F(\lambda, x)$, where $F(\lambda, x)$ and $F(\mu, x)$ are as in section 2. These functions are positive homogeneous and convex, and hence are the support functions of closed convex sets $C(\lambda)$, $C(\mu)$ respectively: $F(\lambda, x)=\sup\{v\cdot x\colon v\in C(\lambda)\}$.

If $f\in L(p, \lambda)$ then its Fourier transform for $w = u+iv$ is given by

(6.1)
$$(2\pi)^{\frac{1}{2}k}\hat{f}(w) = \int f(x)\,e^{-iw\cdot x}\,dx = \int f(x)\,e^{v\cdot x-iu\cdot x}\,dx$$

Now

$$\int |f(x)|^p\,e^{pv\cdot x}\,dx = \int |f|^{pe^{v\cdot x}-F(\lambda,x)}\lambda(x)\,dx$$

so that $f(x)e^{v\cdot x}$ is in L_p if $pv.x\leqq F(\lambda, x)$ for all x, that is, if v is in $C(\lambda)/p$. If v is an interior point of $C(\lambda)/p$ let the distance of v from the exterior $C(\lambda)/p$ be $a>0$, so that $pv\cdot x-F(\lambda, x)\leqq a$ for all x; then $f(x)e^{v\cdot x}\in L$, for

$$\int |f(x)|\,e^{v\cdot x}\,dx = \int |f(x)|\,e^{v\cdot x-F(\lambda,x)/p}\lambda(x)^{1/p}\,dx \leqq$$
$$\leqq \|f\|_{p,\lambda}\Big[\int e^{-p'a\|x\|}\,dx\Big]^{1/p'} = C\,a^{-k/p'},$$

by Hölder's inequality, with C a constant depending only on k and p.

$\hat{f}(w)$ therefore exists if $v\in C(\lambda)/p$ whenever $p\leqq 2$ and, as a function of u, $\hat{f}(u+iv)$ is the Fourier transform of a function in L_p, that is, is in $\hat{L}_p$. In general $\hat{f}(w)$ is analytic for v in the interior of $C(\lambda)/p$ and is the Fourier transform of a function in L when considered as a function of u, and has a supremum for fixed v of the order of $a^{-k/p'}$ with a the distance of v from the boundary of $C(\lambda)/p$.

Now let $T\colon f\to k*f$ be a map from $L(p, \lambda)$ to $L(q, \mu)$; then the adjoint T' can be represented as a map $f\to k*f$ from $L(q', \mu^{-q''})$ to $L(p', \lambda^{-p''})$ and k is a distribution in S', e_w, where $e_w(x)=e^{-iw\cdot x}$, is in $L(q', \mu^{-q''})$ if $q'v.x-q''F(\mu, x)\leqq 0$ for all x, that is provided that $qv\in C(\mu)$, and then $(k*e_w)(x) = e^{-iw\cdot x}(2\pi)^{\frac{1}{2}k}\hat{k}(w)$ is in $L(p', \lambda^{-p''})$. This implies that $\hat{k}(w)$ exists and that $v\in C(\lambda)/p$. Thus for a nonzero map T to exist we must have that $C(\lambda)/q\subset C(\mu)/p$ and $\hat{k}(w)$ must be analytic for v in the interior of $C(\mu)/w$.

The problem of maps from spaces with measures of form $\exp(-F(\lambda, x))$ to those with measures of similar form reduces to the one just discussed: this is, in the notation above, the question of a map from $L(p, \lambda^{-1})$ to $L(q, \mu^{-1})$ and this is equivalent to a map from $L(q', \mu^{q''})$ to $L(p', \lambda^{p''})$; the condition for this is that $p''C(\lambda)/p'\subset$ $\subset q''C(\mu)/q'$, that is, that $C(\lambda)/p\subset C(\mu)/q$.

We now investigate conditions for a multiplier function to give such a map.

Definition. For $a \geqq 0$ let $M^q_{p,a}$ be the set of all $m(u)$ such that $m(u)\hat{f}(u)$ is in L_q whenever f is in $L(p, e^{a\|x\|})$. If $m(u)\hat{f}(u)=\hat{g}(u)$ then we write $M^q_{p,a}(m)$ (or $M^q_{p,a}(m(u))$ when it is necessary to specify the variable in m) for the norm of the map $f \to g$ on $L(p, e^{a\|x\|})$, to L_q.

Since $L(p, e^{a\|x\|})$ decreases as a increases, and since the norm of a fixed element increases with a, $M^q_{p,a}$ increases with a, and the norm of a fixed m decreases with a. These monotonicities are strict. Thus in particular if $p>q$, $M^q_{p,a}$ is empty if $a=0$, but contains the unit function if $a>0$.

Theorem 10. *Let $\lambda(x)=\exp F(\lambda, x)$, $\mu(x)=\exp F(\mu, x)$, where $F(\lambda, x)$, $F(\mu, x)$ are support functions of closed convex sets $C(\lambda)$, $C(\mu)$. Let $pF(\lambda, x) \geqq qF(\mu, x)$ for all x. Let $m(w)$ be analytic if $v \in C(\mu)/q$ and for each v in $C(\mu)/q$ let $m(u+iv)$ be in $M^q_{p,\delta}$ with $M^q_{p,\delta}(m(u+iv)) \leqq K\delta^{-\gamma}$ where K, γ are constants independent of v, $\gamma<k$, and δ is the distance of v from the boundary of $C(\lambda)/p$. Let $C(\mu)/q$ be not entirely contained in the boundary of $C(\lambda)/p$. Then if $q \leqq 2$ the map $f \to g$, where $\hat{g}(w)=m(w)\hat{f}(w)$ for $v \in C(\mu)/q$ is a bounded translation invariant map from $L(p, \lambda)$ to $L(q, \mu)$.*

If $2<q<\infty$ the same holds provided that $m(w)$ is uniformly bounded for v in $C(\mu)$.

According to the hypotheses there is for each v in $C(\mu)/q$ an element $g(v, x) \in L_q$ such that

(6. 3) $$(2\pi)^{\frac{1}{2}k}\hat{g}(u+iv) = \int g(v, x)\, e^{-iu\cdot x}\, dx,$$

and

(6. 4) $$\|g(v, \cdot)\|_q \leqq \delta^{-\gamma} K\left(\int |f(x)|^p e^{\delta\|x\|}\, dx\right)^{1/p}.$$

Our first aim is to prove that this g is essentially independent of v, that is, $g(v, x)= =g(x)e^{v\cdot x}$ for some g in $L(q, \mu)$.

Choose a, b, v in $C(\mu)/q$ so that $a_r \leqq v_r \leqq b_r$ for all r; we may as well suppose the inequalities strict, since there is nothing to prove unless some are strict and we can ignore the coordinates for which a_r, b_r, v_r are equal. Letting $w_r = u_r + iv_r$, for large enough T Cauchy's theorem gives

(6. 5) $$\hat{g}(w) = \int_{C(T)} \frac{\hat{g}(z)\, dz}{P(z, w)},$$

where $P(z, w) = (2\pi i)^{-k}\Pi(z_r - w_r)$, and where $C(T)=\Pi C_r(T)$ with $C_r(T)$ the rectangle with vertices at $ia_r \pm T$, $ib_r \pm T$ described positively. Calling the integral on the right $I(T)$ it follows that

$$\hat{g}(w) = \int_T^{T+1} I(t)\, dt = J(T)+R(T),$$

where $J(T)$ is the part of the integral in (6. 4) over the product of the horizontal sides of the rectangles, and $R(T)$ is a sum of terms of the forms

$$\int_T^{T+1} dt \int_{ia+T}^{ib+T} \frac{\hat{g}(z)}{P(z, w)}\, dz \quad \text{and} \quad \int_T^{T+1} dt \int_T^{t} \frac{\hat{g}(ia+x)}{P(ia+x, w)}\, dx$$

together with others obtained by replacing T and $T+1$ by $-T$, $-T-1$ and a by b. The terms of the first type are not greater in modulus than

(6. 6)

$$\int_a^b dy \int_T^{T+1} \frac{|g(t+iy)|}{|P(t+iy, w)|}\, dt \leqq \int_a^b dy \left(\int_T^{T+1} \hat{g}(t+iy)|^{q'} dt \right)^{1/q'} \left(\int_T^{T+1} \frac{dt}{|P(t+iy, w)|^q} \right)^{1/q} \leqq$$

$$\leqq AT^{-k} \int_a^b dy \left(\int_T^{T+1} |\hat{g}(t+iy)|^{q'} dt \right)^{1/q'},$$

for some constant A. If $q \leqq 2$ the inner integral is not greater than $C_q \| g(y, \cdot)\|_q$ where C_q is the norm of the Fourier transform as a map from L_q to $L_{q'}$. Moreover, $\hat{g}(u+iv) = m(u+iv)\hat{h}(u)$ where $k(x) = f(x)e^{v \cdot x}$ and then

$$\int |h(x)|^p e^{\delta \|x\|} dx = \int |f(x)|^p e^{pv \cdot x + \delta x} dx \leqq \int |f(x)|^p e^{F(\lambda, x)} dx,$$

because $F(\lambda, x) = \sup \{v.x : v \in C(\lambda)\} \geqq pv.x + \delta \|x\|$ since all points within δ of v are in $C(\lambda)$. Then by (6. 4) $\|g(v, \cdot)\|_q \leqq K\delta^{-\gamma} \|f\|_p$.

We can suppose without loss of generality that at most one point of $[a, b]$, say a, lies in the boundary of $C(\lambda)/p$. The integral in (6. 6) is of the order of $\int_a^b \delta^{-\gamma} dy$ where δ is the distance of y from the boundary of $C(\lambda)/p$. If a is not in the boundary, this is bounded; if a is in the boundary, it is of the order of the integral $\int_0 r^{k-1-\gamma} dr$ with $r = \|y-a\|$, and is again bounded if $\gamma < k$ as required by hypothesis. Hence the term (6. 6) is $O(T^{-k})$ as $T \to \infty$.

This argument fails if $q < 2$; however, in that case under the strengthened hypothesis that $m(u+iv)$ is bounded we can consider the side terms in the integral (6. 5) directly. According to (6. 2), assuming again that a is in the boundary of $C(\lambda)/p$, $|f(u+iv)|$ is bounded by a constant multiple of $\delta^{-k/p'}$ and, arguing much as above, the integral is of the order of T^{-k} multiplied by the integral of $\int_0 r^{k(1-1/p')-1} dr$ and since $p'>1$ this is bounded. Once again the side terms tend to 0 and we can write $\hat{g}(w) = I_a - I_b$ where

$$(2\pi)^{\frac{1}{2}k} I_a = \int\limits_{y=a} \frac{dz}{P(z,w)} \int\limits_{-\infty}^{\infty} g(a,\xi) e^{-ix\cdot\xi} d\xi =$$

$$= \int\limits_{-\infty}^{\infty} g(a,\xi)\,d\xi \int\limits_{y=a} \frac{e^{-ix\cdot\xi}}{P(z,w)}\,dz = \int\limits_{0}^{\infty} g(a,\xi) e^{-a\cdot\xi - i\xi\cdot w} d\xi,$$

and similarly

$$(2\pi)^{\frac{1}{2}k} I_b = -\int\limits_{-\infty}^{0} g(b,\xi)\, e^{-b\cdot\xi - i\xi\cdot w} d\xi.$$

On comparison with (6. 3), it follows that $g(x) = g(v,x)e^{-v\cdot x}$ is independent of v and then

$$\left(\int |g(x)|^q \mu(dx)\right)^{1/q} = \left(\int |g(v,x)|^q e^{-qv\cdot x + F(\mu,x)} dx\right)^{1/q} \leqq$$

$$\leqq \left(\int |g(v,x)|^q dx\right)^{1/q} \leqq K\delta^{-\gamma} f_p,$$

when the point v is chosen arbitrarily in the interior of $C(\lambda)/p$. Thus the map $f \to g$ is continuous on $L(p,\lambda)$ to $L(q,\mu)$.

Note that the conclusions of the theorem are valid if $q \leqq 2$ even if $C(\mu)/q$ is entirely contained in the boundary of $C(\lambda)/p$ provided that γ can be taken to be zero.

Sufficient conditions for a map generated by m to be continuous $L(p,\lambda^{-1}) \to L(q,\mu^{-1})$ follow from this theorem, by using the duality arguments. These are that $C(\lambda)/p \subset C(\mu)/q$, that $C(\lambda)/p$ is not completely contained in the boundary of $C(\mu)/q$, that if $p \geqq 2$ $m(u+iv)$ is in $M^{p'}_{q',\delta}$ and that $M^{p'}_{q',\delta}(m(u+iv)) < K\delta^{-\gamma}$ where δ is the distance from v to the boundary of $C(\mu)/q$ and K, γ are as before. If $p<2$ we need in addition that $m(u+iv)$ is uniformly bounded for $v \in C(\mu)/q$.

These conditions are unaltered by translations of $C(\lambda)/p$, $C(\mu)/q$ through the same displacement. This is a particular case of the following observation:

If the kernel k generates a continuous transformation from $L(p,\lambda)$ to $L(q,\mu)$ then the kernel $k(x)e^{a\cdot x}$ generates a continuous transformation from $L(p,\lambda e^{-pa\cdot x})$ to $L(q,\mu e^{-qa\cdot x})$.

The effect of the change in measures involved in this statement is to alter $F(\lambda,x)$ to $F(\lambda,x) - pa\cdot x$ and $F(\mu,x)$ to $F(\mu,x) - qa\cdot x$ and so to displace both $C(\lambda)/p$ and $C(\mu)/q$ by $-a$.

We give some applications of these arguments to particular kernels.

a) The one-dimensional Hilbert transform has $k(x) = 1/\pi x$, $\hat{k}(u) = \operatorname{sgn} u$. This has no analytic extension, so that H cannot map any $L(p,\lambda)$ continuously to $L(q,\mu)$ if $C(\mu)$ contains any nonzero v: that is to say, if μ is in the class we are considering in this section, it must be constant.

b) The Riesz potentials are the transforms R_α with kernels $\|x\|^{\alpha-k}$, $0<\alpha<k$, apart from a constant multiple; the Fourier transform $m(u)$ is a constant multiple of $\|u\|^{-\alpha}$. This has no analytic extension and the same conclusions apply as for the Hilbert transform.

c) Let $e(x)=e^{-\|x\|}$. Then $\hat{e}(w) = (1+\sum w_r^2)^{-k}$ is analytic for $\|v\|<1$, and uniformly bounded on any region $\|v\| < 1-\delta$, if $\delta>0$. Conditions sufficient in order that the map T,

$$Tf(x) = \int e^{-|x-y|} f(y)\,dy,$$

act continuously $L(p, \lambda) \to L(q, \mu)$ are:

A. $F(\mu, x)<q\|x\|$ for all x and B: if $p \leqq q \leqq 2$, $pF(\mu, x) \leqq (qF\lambda, x)$ and if $p>q$, $pF(\mu, x)<qF(\lambda, x)$ for all x.

Dually, the conditions that T act continuously from $L(p, \mu^{-1})$ to $L(q, \lambda^{-1})$ are

A′. $F(\lambda, x)<p\|x\|$; B' if $q \geqq p \geqq 2$, $qF(\lambda, x) \leqq pF(\mu, x)$ for all x, if $p>q$ $qF(\lambda, x)<pF(\mu, x)$ for all x.

d) For the kernel $k(x)=e^{-\|x\|^2}$ the conditions are the conditions B and B' above.

e) If $q<p$ and $C(\mu)/q$ is contained in the interior of $C(\lambda)/p$ then the identity is a continuous map $L(p, \lambda) \to L(q, \mu)$.

This follows easily enough from the theorem, or directly from Theorem 8, for if the conditions hold then there is an ε neighbourhood of $C(\mu)/q$ in $C(\lambda)/p$ and so $F(\lambda, x)/p - F(\mu, x)/q > \varepsilon\|x\|$ for all x, so that the conditions of Theorem 8 hold.

The theorem also allows us to state some cases in which the class of translation invariant maps from $L(p, \lambda)$ to $L(q, \mu)$ is vacuous, that is, consists only of the zero map. Among these are the following:

a) $C(\mu)/q$ is not in $C(\lambda)/p$, that is, $pF(\mu, x)>qF(\lambda, x)$ for some x.

b) $p>q$ and $C(\mu)/q$ contains a boundary point of $C(\lambda)/p$, that is $pF(\mu, x)= =qF(\lambda, x)$ for some x.

For suppose that there is a v common to the boundaries; for this v, $\delta=0$, and if $m(w)$ generates a map $L(p, \lambda) \to L(q, \mu)$ then $m(u+iv) \in M_p^q$ and this, as we have seen, is vacuous if $p>q$.

On the other hand, the class of maps is never vacuous if $C(\mu)/q$ is in the interior of $C(\lambda)/p$; for then if $p>q$ the identity is a continuous map $L(p, \lambda) \to L(q, \mu)$ and if $p \leqq q$ the function $m(w)=e^{-w^2}$ is a multiplier generating a nontrivial continuous map.

Lastly, with λ and μ of the same forms, there remains the question of the existence of map $L(p, \lambda) \to L(q, \mu^{-1})$ and $L(p, \lambda^{-1}) \to L(q, \mu)$. Here the functions in one space or the other are very large at infinity, and do not have Fourier transforms: nor do the functions involved in the dual problem in the other space. The problem can be considered by other methods, and some results may be obtained by using generalized Fourier transforms. However, the following result is obvious, as consequences of previous theorems.

For the map $L(p, \lambda^{-1}) \to L(q, \mu)$ the comparison functions of Theorem 7 are $E(\lambda^{-1}, h) = -\exp(-F(\lambda, h))$, $E(\mu, h) = \exp F(\mu, h)$ so that $L(p, \lambda^{-1}) \to L(q, \mu)$ is vacuous unless λ and μ are 1.

The question of maps from $L(p, \lambda)$ to $L(q, \mu^{-1})$ reduces to that of maps from $L(p, \lambda^{e^{pa\cdot x}})$ to $L(q, \mu^{-1} e^{qa\cdot x})$ according to the argument above, where a is any point. If a is common to $C(\lambda)/p$ and $-C(\mu)/q$, 0 is a common point of the corresponding regions after displacement: for the equivalent F's $F(\lambda, x) \geqq 0$, $F(\mu, x) \geqq 0$ for all x. Assuming this to be the case, we have the continuous imbeddings $L(p, \lambda) \subset L_p$, $L_q \subset L(q, \mu^{-1})$, and so any continuous map $L_p \to L_q$ induces one from $L(p, \lambda)$ to $L(q, \mu^{-1})$. If $p \leqq q$ there are always such maps that are not zero. If $p > q$, the identity is a map from $L(p, \lambda)$ to $L(q, \mu^{-1})$ if a is interior to $C(\lambda)/p$.

References

[1] J. L. B. Cooper, Functional equations for linear transformations, *Proc. London Math. Soc.*, (3) **20** (1970), 1—32.

[2] M. Plancherel, Quelques remarques à propos d'une note de G. H. Hardy: The resultant of two Fourier kernels, *Proc. Cambridge Phil. Soc.*, **33** (1937), 413—418.

[3] H. Kőber, On certain linear operations and relations between them, *Proc. London Math. Soc.*, (3) **11** (1961), 434—456.

[4] H. Kőber, On functional equations and bounded linear transformations, *Proc. London Math. Soc.*, (3) **14** (1964), 495—519.

[5] L. Hörmander, Estimates for translation invariant operators in L^p spaces, *Acta Math.*, **104** (1960), 93—140.

[6] T. A. Timan, Singular integral transformations in weight spaces L_p, *Dokl. Akad. Nauk SSSR*, **184**, (1969). Translation: *Soviet Math. Dokl.*, **10** (1969), 78—82.

[7] N. Bourbaki, *Intégration*, Chapter 7, Hermann (Paris, 1963).

[8] N. Dunford and J. T. Schwartz, *Linear Operators*. Part I, Interscience (New York, 1958).

[9] E. M. Stein and G. Weiss, Interpolation of operators with change of measure, *Trans. Amer. Math. Soc.*, **87** (1958), 159—172.

CHELSEA COLLEGE,
MANRESA ROAD,
LONDON, SW3 6LX.

(Received December 15, 1972)

On weak convergence of randomly selected partial sums

By M. CSÖRGŐ[1]) in Ottawa (Canada) and S. CSÖRGŐ in Szeged

In honour of Professor Béla Szőkefalvi-Nagy on his sixtieth birthday

1. Introduction. Let $\xi_1, \xi_2, \ldots$ be a sequence of random variables (r.v.'s) defined on a probability space $(\Omega, \mathscr{B}, P)$ and suppose that the partial sums $S_n = \xi_1+\xi_2+ +\cdots+\xi_n$ obey the central limit theorem, say with the positive norming factors a_n, so that the distribution of $a_n^{-1} S_n$ is asymptotically the unit normal. Now let ν_n be a sequence of positive integer valued r.v.'s defined on the same probability space. Beginning with the early work of ANSCOMBE (1952), [1], several authors have dealt with the convergence problem of $a_{\nu_n}^{-1} S_{\nu_n}$ (see e.g. [15], [13], [5] and [6]) and, in general, with the following problem: if we are given that a sequence of r.v.'s already satisfies an asymptotic law, then under what conditions should the same sequence, but indexed by ν_n, satisfy the same law (see [17], [12], [8] and [9]). On the other hand, these results have established the background for studying the problem of weak convergence of randomly selected partial sum processes on function spaces, and this work has begun with BILLINGSLEY ([12], 1962). This paper is going to deal with this latter approach, trying to provide a general procedure.

2. Weak convergence of randomly selected partial sum type processes on the space D

Let $D=D[0, 1]$ be the space of functions with discontinuities only of the first kind. Under Prohorov's metric [14] or under Skorokhod's metric [18] with Billingsley's modification of it [3] D is a complete and separable metric space. Let $\mathscr{D}$ be the σ-algebra generated by the open sets of D. If for each $n\geqq 0$, X_n is a measurable mapping from $(\Omega, \mathscr{B})$ to $(D, \mathscr{D})$, that is, in Billingsley's terminology (which will be followed throughout, [3]), X_n is a random function of D, and $\mathscr{P}_n$ denotes the induced image law of X_n on $(D, \mathscr{D})$, then we say X_n converges in distribution to X_0 with the

[1]) Work supported by a National Research Council of Canada Grant at Carleton University, Ottawa.

induced image law $\mathscr{P}_0$, written $X_n \xrightarrow{\mathscr{D}} X_0$, if for all real valued continuous bounded functions g on D $\lim_{n\to\infty} \int_D g\,d\mathscr{P}_n = \int_D g\,d\mathscr{P}_0$ holds. The sequence X_n is called tight if for every positive ε there exists a compact set K in $\mathscr{D}$ such that $\mathscr{P}_n(K) > 1-\varepsilon$, $n=1, 2, \ldots$. The following two theorems (Theorems 15. 2 and 15. 1, [3]), of which the first one characterises the notion of tightness in D and the second one the convergence in distribution (weak convergence), will be used in the sequel.

Theorem A. *The sequence X_n is tight if and only if these two conditions hold:*
(i) *For each positive η there exists a d such that*

$$P\{\sup_t |X_n(t)| > d\} < \eta, \quad n \geqq 1.$$

(ii) *For each positive ε and η, there exist a δ, $0<\delta<1$, and an integer n_0 such that $P\{w'(X_n, \delta)>\varepsilon\}<\eta$, $n\geqq n_0$, where*

$$w'(X_n, \delta) = \inf_{(1,\delta,\{t_i\})} \max_{1\leqq i\leqq r} w_{X_n}([t_{i-1}, t_i)) \tag{1}$$

with $w_{X_n}([t_{i-1}, t_i)) = \sup_{u,v\in[t_{i-1},t_i)} |X_n(u)-X_n(v)|$, *and the infimum extends over all finite sets (t_i) of points satisfying*

$$0 = t_0 < t_1 < \ldots < t_r = 1, \quad t_i - t_{i-1} > \delta, \qquad i = 1, 2, \ldots, r.$$

Theorem B. *For X_n to converge in distribution to X_0 it is necessary and sufficient that the finite dimensional distributions of it should converge to those of X_0 and that X_n should be tight.*

The random functions of D we are going to be concerned with are of the form:

$$X_n(t) = X_n(t, \omega) = a_n^{-1} X(nt, \omega), \quad 0 \leqq t \leqq 1, \quad n \geqq 1, \tag{2}$$

where $X(u, \omega)$, for each fixed ω in Ω, is a right continuous function of u on $[0, \infty)$ having left-hand limits at each point, and for u fixed it is measurable with respect to $(\Omega, \mathscr{B})$; the sequence a_n of positive numbers increases monotonically to ∞ with n, and it is also slowly oscillating in the sense of Karamata. This latter notion means that a_n is of the form $n^\alpha L(n)$ with α positive and $L([cn])/L(n) \to 1$ as $n \to \infty$ for every positive c.

The most immediate examples of the form of (2) are the partial sum processes $a_n^{-1} S_{[nt]}$, $S_0=0$, and several other processes of D can be brought into this form when technicalities of certain proofs so require. As to the latter we mention a fortchoming paper by one of us [10], concerning the weak convergence of the random sample size empirical process.

In this exposition the possibility of deducing the weak convergence of X_{ν_n} from that of X_n of (2) is examined. Towards this end the following lemma is essential.

Lemma. *If X_n of* (2) *is tight and if the sequence of positive integer valued r.v.'s ν_n is such that $\nu_n/n \xrightarrow{P} \nu$, where ν is an arbitrary positive r.v then X_{ν_n} is also tight.*

Proof. The special form of X_n of (2) implies that conditions (i) and (ii) of Theorem A are satisfied if and only if

(i)* For each positive η there exists a d such that

$$P\{\sup_{0 \leqq v \leqq n} |X(v)| > da_n\} < \eta, \quad n \geqq 1.$$

(ii)* For each positive ε and η, there exists a δ, $0<\delta<1$, and an integer n_0 such that

$$P\{\inf_{(n,\delta,\{v_i\})} M_r(n) > \varepsilon a_n\} < \eta, \quad n \geqq n_0,$$

where $M_r(n) = \max_{1 \leqq i \leqq r} \sup_{v,u \in [v_{i-1}, v_i)} |X(u)-X(v)|$, and $\inf_{(n,\delta,\{v_i\})}$ stands for the infimum over the finite sets of points (v_i) satisfying

$$0 = v_0 < v_1 < \dots < v_r = n, \quad v_i - v_{i-1} > n\delta, \quad i = 1, 2, \dots, r.$$

In order to prove our lemma, we simply have to verify conditions (i)* and (ii)* for X_{ν_n}, that is to say (i)* and (ii)* with n replaced by ν_n in them.

First we verify condition (ii)*. Let ε and η be fixed positive numbers and let $\theta = \eta/3$. Choose $0<a<b$ so that $P\{a \leqq \nu < b\} > 1-\theta$. Without loss of generality assume that $\varepsilon < a$, and choose $n_1 = n_1(\varepsilon, \theta)$ so that $P\left\{\left|\frac{\nu_n}{n} - \nu\right| > \varepsilon\right\} < \theta$ for $n \geqq n_1$. For arbitrary δ and $n \geqq n_1$

$$(3) \qquad P\{\inf_{(\nu_n,\delta,\{v_i\})} M_r(\nu_n) > \varepsilon a_{\nu_n}\} \leqq P\{\inf_{(\nu_n,\delta,\{v_i\})} M_r(\nu_n) > \varepsilon a_{\nu_n},$$

$$n(a-\varepsilon) \leqq \nu_n \leqq n(b+\varepsilon)\} + 2\theta.$$

Now for each fixed $\omega \in \Omega$ and ν_n in the indicated range above

$$\inf_{(n(b+\varepsilon),\delta,\{v_i\})} M_r(n(b+\varepsilon)) \geqq \inf_{(\nu_n,\delta,\{v_i\})} M_r(\nu_n),$$

and $a_{[n(a-\varepsilon)]} \leqq a_{\nu_n}$, so the last probability of relation (3) is less than or equal to

$$(4) \qquad P\{\inf_{(n(b+\varepsilon),\delta,\{v_i\})} M_r(n(b+\varepsilon)) > \varepsilon a_{[n(a-\varepsilon)]}\}.$$

Also $a_n = n^\alpha L(n)$ is slowly oscillating and it can be easily computed that $a_{[n(a-\varepsilon)]}/a_{[n(b+\varepsilon)]} \to ((a-\varepsilon)/(b+\varepsilon))^\alpha$ as $n \to \infty$. Thus, if we now choose a positive number ϱ so that $\varepsilon_0 = \varepsilon(((a-\varepsilon)/(b+\varepsilon))^\alpha - \varrho)$ is also positive, then there exists an

$n_2(\geqq n_1)$ so that for $n \geqq n_2$ the probability under (4) is bounded above by

$$P\{\inf_{(n(b+\varepsilon),\delta,\{v_i\})} M_r(n(b+\varepsilon)) > \varepsilon_0 a_{[n(b+\varepsilon)]}\}.$$

Since the sequence X_n is tight, therefore, for ε_0 and θ we can now choose δ and $n_0(\geqq n_2)$ such that the last probability is less than θ which, in turn, implies that the left hand side probability of (3) is less than η.

Turning now to the proof (i)* we let $\varepsilon, \eta, \theta, a, b, \varrho, n_1$ and n_2 be as in the proof of (ii)* above, and putting $d_0 = d(((a-\varepsilon)/(b+\varepsilon))^\alpha - \varrho)$ we get immediately:

$$(5) \qquad P\{\sup_{0 \leqq v \leqq \nu_n} |X(v)| > d a_{\nu_n}\} \leqq P\{\sup_{0 \leqq v \leqq [n(b+\varepsilon)]} |X(v)| > d_0 a_{[n(b+\varepsilon)]}\} + 2\theta,$$

provided n is not less than n_2. Again, since the sequence X_n is tight, for θ we can choose $d = d^*$ so large that d_0 becomes large enough to ensure that the right hand side probability of the inequality of (5) is less than θ for every n. Consequently, for the given η there exist a d^* and n_2 so that

$$(6) \qquad P\{\sup_{0 \leqq v \leqq \nu_n} |X(v)| > d^* a_{\nu_n}\} < \eta, \quad n \geqq n_2.$$

Thus, the only question now whether such a d should also exist which would make relation (6) hold for all n. Since the space D is complete and separable, each single probability measure on $(D, \mathscr{D})$ is tight and so are, therefore, the ones induced by $X_1, X_2, \ldots, X_{n_2-1}$. Now the characterization theorem of the compact subsets of D (Theorem 14. 3, [3]) implies the existence of d_i so that

$$P\{\sup_{0 \leqq v \leqq \nu_i} |X(v)| > d_i a_{\nu_i}\} < \eta, \qquad i = 1, \ldots, n_2 - 1,$$

and relation (6) holds for every $n \geqq 1$ with $d = \max(d^*; d_1, \ldots, d_{n_2-1})$ instead of d^* in it. This completes the proof of Lemma.

Having proved this lemma, our programme now only requires us to be able to deduce the convergence of the finite dimensional distributions of X_{ν_n} from those of X_n. On the bases of recent literature, concerning the limiting distributions of sequences of r.v.'s with random indices, this can be done several ways. We are going to demonstrate two possibilities here which, we believe, are the most important ones available from the point of view of applications. They are based on a recent paper of Guiaşu [12]. As to other ways of possible approach we refer to a forthcoming work of Fischler [11].

For a random function X of D let $T_X = \{t \in [0, 1] : P\{X(t) \neq X(t-)\} = 0\}$.

Theorem 1. *Let X and the sequence X_n be random functions of the space D, X_n having the form as in* (2). *Assume:*

(a) $\nu_n/n \xrightarrow{P} \nu$, *where the sequence ν_n and the r.v. ν are as in Lemma;*

(b) $X_n \xrightarrow{\mathscr{D}} X$;

(c) *For an arbitrary positive integer k, all arbitrary real numbers $c_1, c_2, \dots, (c_k \neq 0)$ and arbitrary time points $t_1, t_2, \dots, t_k \in T_X$, the random variables $Y_n = \sum_{i=1}^{k} c_i X_n(t_i)$ and $Y = \sum_{i=1}^{k} c_i X(t_i)$ satisfy (at every continuity point x of $P\{Y \leqq x\}$)*

$$\lim_{n\to\infty} P\{Y_n \leqq x | A\} = P\{Y \leqq x\},$$

for every $A \varepsilon \mathcal{K}_\nu$, $P\{A\} > 0$, where $\mathcal{K}_\nu$ is the σ-algebra generated by ν;

(d) *For every positive ε and η and every $A \in \mathcal{K}_\nu$, $P\{A\} > 0$, there exist a positive real number $c = c(\varepsilon, \eta)$ and a natural number $n_0 = n_0(\varepsilon, \eta, A)$ such that for every $n \geqq n_0$*

$$P\{\max_{n(1-c)\leqq m \leqq n(1+c)} |X(nt) - X(mt)| > \varepsilon a_n | A\} < \eta,$$

at every fixed $t \in T_X$. Then $X_{\nu_n} \xrightarrow{\mathcal{D}} X$.

Proof. In the light of our Lemma and Theorem B we only have to deal with the convergence of the finite dimensional distributions of X_{ν_n}. If we now observe

$$P\left\{\max_{n(1-c)\leqq m\leqq n(1+c)} \left|\sum_{i=0}^{k} c_i X(nt_i) - \sum_{i=1}^{k} c_i X(mt_i)\right| > \varepsilon a_n \middle| A\right\} \leqq$$

$$\leqq \sum_{i=1}^{k} P\{\max_{n(1-c)\leqq m\leqq n(1+c)} |X(nt_i) - X(mt_i)| > \varepsilon\, a_n/k\, |c_i|\, |A\},$$

then the conditions of Theorem 3 of GUIAŞU [12] are satisfied for the sequence $Y_n = \sum_{i=1}^{k} c_i X_n(t_i) = a_n^{-1} \sum_{i=1}^{k} c_i X(nt_i)$ if one also notes that Theorem 3 of [12] holds for a sequence of random variables Y_n of the form $Y_n = a_n^{-1} Z_n$ with its condition (C6) modified to the extent that in it one writes Z_i and Z_n instead of Y_i and Y_n respectively, and εa_n instead of ε. As a consequence of the Cramér—Wold device ([3], p. 49), our theorem is now proved.

Remarks. Lemma and Theorem 1 remain valid if the sequence ν_n and ν are such that $\nu_n/f(n) \xrightarrow{P} \nu$, where $f(n)$ are constants going to infinity. Incidentally, the Lemma itself would still hold if $a_n = n^\alpha L(n)$ is monotone decreasing or, if α is negative, independently again of a_n being increasing or decreasing. Condition (d) of Theorem 1 with $A = \Omega$ is the classical Anscombe condition [1].

Applications of Theorem 1

1) Let $\nu=\theta$, a positive constant, and $X_n(t)=S_{[nt]}/\sigma\sqrt{n}$ $(\sigma>0)$. Billingsley ([3], Theorem 17. 1) proves that $X_n \xrightarrow{\mathcal{D}} W$ implies $X_{\nu_n} \xrightarrow{\mathcal{D}} W$, where W is the Brownian motion on D. This result is a special case of Theorem 1, for when ν is a constant then $\mathcal{K}_\nu=\{\emptyset, \Omega\}$ and condition (c) with $A=\Omega$ is implied by (b). Also, condition (b) implies tightness of X_n, which, in turn, implies condition (d) with $A=\Omega$. This completes the proof of Billingsley's Theorem 17. 1 and the above procedure also shows that the assumption $X_n \xrightarrow{\mathcal{D}} W$ in his theorem can be replaced by $X_n \xrightarrow{\mathcal{D}} X$, where X is not necessarily the Brownian motion.

2) If the summands $\xi_1, \xi_2, \ldots$ of $S_{[nt]}$ are independent and identically distributed with zero mean and variance σ^2 and $\nu_n/f(n) \xrightarrow{P} \nu$, where ν is a positive r.v., then (again with $X_n(t)=S_{[nt]}/\sigma\sqrt{n}$) $X_{\nu_n} \xrightarrow{\mathcal{D}} W$, (Theorem 17. 2, [3]). This theorem of Billingsley is a generalization of Donsker's theorem, and it is also implied by Theorem 1 as follows: Donsker's theorem says that $X_n \xrightarrow{\mathcal{D}} W$, which implies that X_n is tight and this, in turn, ensures the tightness of X_{ν_n} via Lemma. Condition (c) of Theorem 1 is a mixing condition in the sense of Rényi [16], and for Y_n it can be verified exactly the same way as for one sum of independent, identically distributed r. v.'s. As to condition (d) we refer to Lemma 3 of Blum, Hanson and Rosenblatt [5], which implies that the conditional probability there can be considered only with $A=\Omega$, which then becomes the classical Anscombe condition for sums of independent, identically distributed r.v.'s, and, with this, Theorem 17. 2 of [3] now follows.

For a function $s(t)$ of D let $h(s(t)) = s(t)-ts(1)$. Then, with X_n as in 2) of Applications above and $\sigma^2=1$ we have $h(X_{\nu_n}) \xrightarrow{\mathcal{D}} W^0$, W^0 the Brownian bridge on D. In [7] we indicated a direct proof of this and used it to prove the random sample size Kolmogorov—Smirnov theorems. As already mentioned earlier, the weak convergence of the random sample size empirical process itself will be proved in [10].

The way we have proved Billingsley's Theorem 17. 2, [3], suggests the following version of Theorem 1.

Theorem 2. *Let X and the sequence X_n be random functions of the space D, X_n having the form as in* (2). *Assume conditions* (a), (b), (c) *of Theorem* 1 *and its condition* (d) *with $A=\Omega$. Assume also:*

(e) *For every positive ε and c and every $A\in\mathcal{K}_\nu$, $P\{A\}>0$, we have*

$$\limsup_{n\to\infty} P\{A_n^t|A\} = \limsup_{n\to\infty} P\{A_n^t\}$$

at every fixed $t\in T$, where A_n^t is the event

$$\{\max_{n(1-c)\le m\le n(1+c)} |X(nt)-X(mt)| > \varepsilon a_n\}.$$

Then $X_{\nu_n} \xrightarrow{\mathcal{D}} X$.

Proof. It is sufficient to note that condition (d) with $A=\Omega$ of Theorem 1 and condition (e) together imply condition (d) of Theorem 1. Thus Theorem 2 follows from Theorem 1.

We note that condition (e) of Theorem 2 holds any time each set in the tail σ-field of the sequence X_n has probability 0—1 (Theorem 2, [4]).

Applications of Theorem 2. Let $X_n(t)=S_{[nt]}/\sqrt{n}$, where the summands $\xi_1, \xi_2, \ldots$ of $S_{[nt]}$ have mean zero and variance one, but are not necessarily independent and identically distributed r.v.'s. SREEHARI (Theorems 2. 2 and 3. 1, [19]) proves that if conditions (1) $X_n \xrightarrow{\mathcal{D}} W$, where W is the Brownian motion on D, (2) condition (a) of Theorem 1 holds, (3) the random variables Y_n of condition (c) of Theorem 1 in terms of $X_n(t)=S_{[nt]}/\sqrt{n}$ satisfy

$$P\{Y_n \leqq x|A\}-P\{Y_n \leqq x\} \to 0, \quad \text{as} \quad n \to \infty,$$

for every $A \in \mathcal{K}_\nu$, $P\{A\}>0$, and (4) for every positive ε and c and every $A \in \mathcal{K}_\nu$, $P\{A\}>0$, we have (cf. Remark of [19], p. 437)

$$P\{A_n^t|A\}-P\{A_n^t\} \to 0, \quad \text{as} \quad n \to \infty,$$

where A_n^t is as in condition (e) of Theorem 2 in terms of $X_n(t)=S_{[nt]}/\sqrt{n}$, hold then $X_{\nu_n} \xrightarrow{\mathcal{D}} W$. This result is a special case of Theorem 2, for condition (1) implies that condition (3) is of the form of condition (c) of Theorem 1. Also, condition (4) above implies the form of condition (e) of Theorem 2. Thus Theorems 2. 2 and 3. 1 of Sreehari's paper [19] follow from Theorem 2 and our proof of it also shows that the assumption $X_n \xrightarrow{\mathcal{D}} W$ in his theorems can be replaced by $X_n \xrightarrow{\mathcal{D}} X$, where X is not necessarily the Brownian motion. Examples, satisfying the conditions (3) and (4) above and, therefore, also the relevant conditions of Theorem 2, are given in [19].

References

[1] F. J. ANSCOMBE, Large-sample theory of sequential estimation, *Proc. Cambridge Philos. Soc.*, **48** (1952), 600—607.

[2] P. BILLINGSLEY, Limit theorems for randomly selected partial sums, *Ann. Math. Statist.*, **33** (1962), 85—92.

[3] P. BILLINGSLEY, *Convergence of probability measures* (Wiley, New York, 1968).

[4] D. BLACKWELL and D. A. FREEDMAN, The tail σ-field of a Markov chain and a theorem of Orey, *Ann. Math. Statist.*, **35** (1964), 1290—1295.

[5] J. BLUM, D. HANSON and J. ROSENBLATT, On the central limit theorem for the sum of a random number of independent random variables, *Z. Wahrscheinlichkeitstheorie und verw. Gebiete*, **1** (1963), 389—393.

[6] M. CSÖRGŐ, On the strong law of large numbers and the central limit theorem for martingales, *Trans. Amer. Math. Soc.*, **131** (1968), 259—275.

[7] M. Csörgő and S. Csörgő, An invariance principle for the empirical process with random sample size, *Bull. Amer. Math. Soc.*, **76** (1970), 706—710.

[8] M. Csörgő and R. Fischler, Departure from independence: the strong law standard and random-sum central limit theorems, *Acta Math. Acad. Sci. Hung.*, **21** (1970), 105—114.

[9] S. Csörgő, On limit distributions of sequences of random variables with random indices, *Acta Math. Acad. Sci. Hung.* (to appear).

[10] S. Csörgő, On weak convergence of the empirical process with random sample size, *Acta Sci. Math.* (to appear).

[11] R. Fischler, Weak convergence limit theorems with randomized indices, *Rev. Roumaine Math. Pures Appl.* (to appear).

[12] S. Guiaşu, On the asymptotic distribution of the sequences of random variables with random indices, *Ann. Math. Statist.*, **42** (1971), 2018—2028.

[13] J. Mogyoródi, A central limit theorem for the sum of a random number of random variables, *Publ. Math. Inst. Hung. Acad. Sci.*, **7** (1962), 409—424.

[14] Yu. V. Prohorov, Convergence of random processes and limit theorems in probability theory, *Theor. Probability Appl.*, **1** (1956), 157—214.

[15] A. Rényi, On the central limit theorem for the sum of a random number of random variables, *Acta Math. Acad. Sci. Hung.*, **11** (1960), 383—393.

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

[17] W. Richter, Übertragung von Grenzaussagen für Folgen von Zufälligen Grössen auf Folgen mit Zufälligen Indizes. *Theor. Probability Appl.*, **10** (1965), 82—93.

[18] A. V. Skorokhod, Limit theorems for stochastic processes, *Theor. Probability Appl.*, **1** (1956), 261—290.

[19] M. Sreehari, An invariance principle for random partial sums, *Sankhyā* (A), **30** (1968), 433—442.

(Received September 28, 1972)

Inner dilations of analytic matrix functions and Darlington synthesis*)

By R. G. DOUGLAS**) and J. WILLIAM HELTON
in Stonybrook (N.Y., U.S.A.)

Dedicated to Professor Béla Sz.-Nagy on his 60th birthday

0. Introduction. In this article we prove the following theorem motivated by the problem of Darlington synthesis in electrical network theory.

Theorem. *Let $S(z)$ be an analytic contractive operator-valued function on the unit disk $\mathbf{D}$. If $S(z)$ has a meromorphic pseudo-continuation of bounded type to the exterior $\mathbf{D}_e$ of the unit disk (including the point at infinity), then $S(z)$ has an inner dilation $U(z)$, that is, there exists an analytic operator-valued function $U(z)$ which is unitary-valued on the unit circle $\mathbf{T}$ such that*

$$U(z) = \begin{pmatrix} S(z) & A(z) \\ B(z) & C(z) \end{pmatrix}. \tag{1}$$

If $S(z)$ and $U(z)$ are both matrix-valued functions, or if $U(z)$ is required to have a scalar multiple, then this condition is necessary as well as sufficient.

If $S(z)$ is matrix-valued and has rational functions as entries, then an ordinary meromorphic continuation exists and hence a dilation is always possible in this case. The construction of such an inner function is actually described in most books on network design under the heading of Darlington or Belevitch synthesis (cf. [6]). (The results there are stated for functions on the right half plane since this is the context in which the engineers work.) Our theorem extends this classical result and generalizes a recent result of the electrical engineer DEWILDE ([1], Thm. 4).

Analytic complex-valued functions defined on $\mathbf{D}$ and lying in the Hardy space H^2 which possess meromorphic pseudo-continuations of bounded type to $\mathbf{D}_e$ have been studied in [3], where it is shown that they are precisely the non-cyclic vectors

*) Research supported in part by a grant from the National Science Foundation.
**) Sloan Foundation Fellow.

for the backward shift operator. Several alternate characterizations of this class of functions are given there.

The study of operator functions on D which possess meromorphic pseudo-continuations to D_e would seem to merit further study. We show in passing that inner functions have such a continuation of bounded type if and only if they have a scalar multiple.

1. Meromorphic pseudo-continuation. If $\mathfrak{H}$ is a separable complex Hilbert space, then $L^2_{\mathfrak{H}}$ denotes the Hilbert space of norm square-integrable Lebesgue-measurable $\mathfrak{H}$-valued functions on T (see [9]). We let $H^2_{\mathfrak{H}}$ denote the closed subspace of functions in $L^2_{\mathfrak{H}}$ with zero Fourier coefficients of negative indices. If $\mathfrak{H}_1$ and $\mathfrak{H}_2$ are separable complex Hilbert spaces, then $\mathfrak{L}(\mathfrak{H}_1, \mathfrak{H}_2)$ denotes the Banach space of bounded linear transformations from $\mathfrak{H}_1$ to $\mathfrak{H}_2$. We abbreviate $\mathfrak{L}(\mathfrak{H}, \mathfrak{H})$ as $\mathfrak{L}(\mathfrak{H})$. Moreover, we let $L^\infty_{\mathfrak{L}(\mathfrak{H}_1, \mathfrak{H}_2)}$ denote the Banach space of essentially bounded weakly-measurable $\mathfrak{L}(\mathfrak{H}_1, \mathfrak{H}_2)$-valued functions on T, while $H^\infty_{\mathfrak{L}(\mathfrak{H}_1, \mathfrak{H}_2)}$ denotes the subspace of functions with zero Fourier coefficients of negative indices. We let H^∞ denote $H^\infty_{\mathfrak{L}(\mathbf{C},\mathbf{C})}$.

Functions in the Hardy spaces $H^2_{\mathfrak{H}}$ and $H^\infty_{\mathfrak{L}(\mathfrak{H}_1, \mathfrak{H}_2)}$ can be identified with the boundary values of holomorphic functions on D.

We consider only those vector-valued meromorphic functions which are of "bounded type". In particular, we allow only those meromorphic functions on the exterior D_e (including the point at infinity) of D which can be expressed in the form $\dfrac{f}{\varphi}$ or $\dfrac{F}{\varphi}$, where f and F lie in $H^2_{\mathfrak{H}}(D_e)$ and $H^\infty_{\mathfrak{L}(\mathfrak{H}_1, \mathfrak{H}_2)}(D_e)$, respectively, while φ lies in $H^\infty(D_e)$. Such functions possess radial limits a.e. and the following lemma is an easy exercise.

Lemma 1. *A function F in $L^\infty_{\mathfrak{L}(\mathfrak{H}_1, \mathfrak{H}_2)}$ is the boundary values of a meromorphic function on D_e of bounded type if and only if there exists an inner (unimodular) function φ in H^∞ such that the function G defined by $G(e^{i\theta}) = \varphi(e^{i\theta})F(e^{i\theta})^*$ lies in $H^\infty_{\mathfrak{L}(\mathfrak{H}_1, \mathfrak{H}_2)}$.*

An $\mathfrak{L}(\mathfrak{H}_1, \mathfrak{H}_2)$-valued meromorphic function G of bounded type of D_e is said to be a *pseudo-continuation* of the holomorphic function F in $H^\infty_{\mathfrak{L}(\mathfrak{H}_1, \mathfrak{H}_2)}(D)$ if

$$\lim_{r\to 1^-} F(re^{i\theta}) = \lim_{r\to 1^+} G(re^{i\theta}) \quad \text{a.e.}$$

One can show that such a function, if it exists, is unique. This generalization of the notion of continuation was introduced for scalar functions by Shapiro [8], and additional information can be found in [3].

A function Q in $H^\infty_{\mathfrak{L}(\mathfrak{H}_1, \mathfrak{H}_2)}$ is said to be *outer* if the operator M_Q has dense range, where M_Q is the multiplication operator from $H^2_{\mathfrak{H}_1}$ to $H^2_{\mathfrak{H}_2}$ defined by

$$(M_Q f)(e^{i\theta}) = Q(e^{i\theta})f(e^{i\theta}) \quad \text{for } f \text{ in } \quad H^2_{\mathfrak{H}_1}.$$

A function U in $H^\infty_{\mathfrak{L}(\mathfrak{H})}$ is said to be *inner* if $U(e^{i\theta})$ is a unitary operator on $\mathfrak{H}$ a.e.

Every meromorphic operator-valued function Φ on $\boldsymbol{D}_e$ can be written as the quotient of an analytic operator-valued function Θ by an analytic scalar-valued function φ. One may take φ to have its zeros (of appropriate order) at the poles of Φ. Thus, the question of whether φ can be taken to be bounded depends on the number of poles of Φ. One could probably further generalize the notion of continuation to allow an even weaker relation between the inside and outside functions when radial limits do not exist.

Although we shall make no direct use of the following observation in this note, we want to point out a relation between meromorphic pseudo-continuations and having a scalar multiple. Recall that a contractive function Θ in $H^\infty_{\mathfrak{L}(\mathfrak{H}_1, \mathfrak{H}_2)}$ is said to have a *scalar multiple* if there exist contractive functions δ in H^∞ and Ω in $H^\infty_{\mathfrak{L}(\mathfrak{H}_2, \mathfrak{H}_1)}$ such that

$$\Omega(z)\,\Theta(z) = \delta(z)\,I_{\mathfrak{H}_1} \quad \text{and} \quad \Theta(z)\,\Omega(z) = \delta(z)\,I_{\mathfrak{H}_2} \quad \text{for } z \text{ in } \boldsymbol{D}.$$

If Θ is an inner function, then we obtain $\Theta(e^{i\theta}) = (\Theta(e^{i\theta})^{-1})^* = \left(\dfrac{\Omega(e^{i\theta})}{\delta(e^{i\theta})}\right)^*$, hence $\dfrac{\Omega(1/\bar{z})^*}{\overline{\delta(1/\bar{z})}}$ is a meromorphic pseudo-continuation of Θ to $\boldsymbol{D}_e$, of bounded type. This and the converse we state as

Proposition 2. *If Θ is an inner function in $H^\infty_{\mathfrak{L}(\mathfrak{H})}$, then Θ has a scalar multiple if and only if Θ possesses a meromorphic pseudo-continuation of bounded type to $\boldsymbol{D}_e$.*

Proof. From Lemma 1 it follows that if Θ possesses such a pseudo-continuation, then there exists an inner function δ such that the function Ω defined by $\Omega(z) = \delta(z)\Theta(z)^*$ lies in $H^\infty_{\mathfrak{L}(\mathfrak{H})}$. Thus, Θ has a scalar multiple.

We now appeal to a result from [9, Cor V. 6. 3] or to the Schwarz reflection principle for matrix valued functions to obtain

Proposition 3. *If U is an inner function in $H^\infty_{\mathfrak{L}(C^n)}$, then U possesses a meromorphic pseudo-continuation of bounded type to $\boldsymbol{D}_e$.*

Combining this with a result [9, VI. 5. 1] of Sz.-Nagy and Foiaş shows that a contraction operator lies in class C_0 if and only if its characteristic operator function is inner and has a meromorphic pseudo-continuation of bounded type to $\boldsymbol{D}_e$. An interesting observation of Lax and Ralston is that the scattering matrix in the Lax—Phillips theory [5] has all the above properties except that the continuation is not of bounded type.

We now go on to the main theorem and refer the reader to [4] and [9] for further facts about Hardy spaces and inner and outer functions.

2. Inner dilations and the main theorem. The necessity half of the theorem now follows easily. If $S(z)$ has an inner dilation $U(z)$ with scalar multiple, then it follows from Proposition 2 that $U(z)$ has a meromorphic pseudo-continuation of bounded type to $\boldsymbol{D}_e$ and consequently that $S(z)=PU(z)P$ does.

The sufficiency is less immediate and we begin with some definitions. If R is a contraction-valued function in $H^\infty_{\mathfrak{L}(\mathfrak{H})}$, then an isometric enlargement of R is an isometry-valued function W in $H^\infty_{\mathfrak{L}(\mathfrak{H},\mathfrak{K})}$, where $\mathfrak{K}$ is a superspace of $\mathfrak{H}$, such that $R(e^{i\theta})=QW(e^{i\theta})$ a.e., where Q is the orthogonal projection of $\mathfrak{K}$ onto $\mathfrak{H}$. Some contraction-valued functions have isometric enlargements and some do not. This notion is closely related to that of factorization. A non-negative operator-valued function M in $L^\infty_{\mathfrak{L}(\mathfrak{H})}$ is said to be factorable if there exists F in $H^\infty_{\mathfrak{L}(\mathfrak{H},\mathfrak{G})}$ such that $M(e^{i\theta})=F(e^{i\theta})^*F(e^{i\theta})$ a.e. The precise relation between isometric enlargement and factorization is contained in

Proposition 4. *A contraction-valued function R in $H^\infty_{\mathfrak{L}(\mathfrak{H})}$ has an isometric enlargement if and only if the function M defined by $M(e^{i\theta})=I-R(e^{i\theta})^*R(e^{i\theta})$ is factorable.*

Proof. If R has an isometric enlargement W in $H^\infty_{\mathfrak{L}(\mathfrak{H},\mathfrak{K})}$, then W can be expressed as

$$W(e^{i\theta})=\begin{pmatrix}R(e^{i\theta})\\F(e^{i\theta})\end{pmatrix}$$

where F is in $H^\infty_{\mathfrak{L}(\mathfrak{H},\mathfrak{G})}$ and $\mathfrak{G}=\mathfrak{K}\ominus\mathfrak{H}$. Computing, we have

$$I_{\mathfrak{K}}=W(e^{i\theta})^*W(e^{i\theta})=R(e^{i\theta})^*R(e^{i\theta})+F(e^{i\theta})^*F(e^{i\theta})\quad\text{a.e.}$$

and hence the non-negative operator-valued function $M(e^{i\theta})=I-R(e^{i\theta})^*R(e^{i\theta})$ is factorable. Conversely, if M is factorable, then W can be defined in this manner.

Not every non-negative operator function is factorable and various criteria are known (cf. [2]). We add a further

Proposition 5. *If the non-negative and contractive operator-valued function M in $L^\infty_{\mathfrak{L}(\mathfrak{H})}$ is the boundary value of a meromorphic $\mathfrak{L}(\mathfrak{H})$ valued function of bounded type on $\boldsymbol{D}_e$, then M is factorable as $M(e^{i\theta})=A(e^{i\theta})^*A(e^{i\theta})$ for some outer function A which has a meromorphic pseudo-continuation of bounded type to $\boldsymbol{D}_e$.*

Proof. Since $M\geqq M^2$ one can derive from a standard comparison theorem [2] or ([7], Theorem 1.18) that it suffices to show that M^2 is factorable. By definition there exists a scalar inner function φ such that φM lies in $H^\infty_{\mathfrak{L}(\mathfrak{H})}$. Hence, we have

$$\bigcap_{n\geqq 0}\operatorname{clos}[e^{in\theta}M(e^{i\theta})H^2_{\mathfrak{H}}]=\bar\varphi\{\bigcap_{n\geqq 0}\operatorname{clos}[e^{in\theta}\varphi(e^{i\theta})M(e^{i\theta})H^2_{\mathfrak{H}}]\}\subset$$

$$\subset\bar\varphi\{\bigcap_{n\geqq 0}\operatorname{clos}[e^{in\theta}H^2_{\mathfrak{H}}]\}=\{0\}$$

which shows that M^2 is factorable by ([9], Prop. V. 4. 2).

It remains to show that A has a meromorphic pseudo-continuation of bounded type to D_e. We have $\varphi M = \varphi A^* A$ and since φM lies in $H^\infty_{\mathfrak{L}(\mathfrak{H})}$, the operator φF^* maps $FH^2_{\mathfrak{H}}$ into $H^2_{\mathfrak{H}}$. Since F is outer, $FH^2_{\mathfrak{H}}$ is dense in $H^2_{\mathfrak{G}}$ and therefore φF^* maps $H^2_{\mathfrak{G}}$ into $H^2_{\mathfrak{H}}$. Thus φF^* lies in $H^\infty_{\mathfrak{L}(\mathfrak{G},\mathfrak{H})}$ which completes the proof.

Now we complete the proof of the main theorem. If S is a contraction-valued operator function in $H^\infty_{\mathfrak{L}(\mathfrak{H})}$ with a meromorphic pseudo-continuation of bounded type to D_e, then by Proposition 5, $I - S(e^{i\theta})^* S(e^{i\theta})$ has a factorization which is pseudo-continuable to a meromorphic operator-valued function of bounded type on D_e. Thus, by Proposition 4, S has an isometric enlargement W in $H^\infty_{\mathfrak{L}(\mathfrak{H},\mathfrak{K})}$ which by construction has a meromorphic pseudo-continuation of bounded type to D_e.

If we now consider the function R in $H^\infty_{\mathfrak{L}(\mathfrak{H},\mathfrak{K})}$ defined by $R(e^{i\theta}) = W(e^{-i\theta})^*$, then R is contraction-valued and has the obvious meromorphic pseudo-continuation of bounded type to D_e defined in terms of that for W.

Thus, again applying Propositions 4 and 5 to R, we obtain an isometric enlargement of R to a function V in $H^\infty_{\mathfrak{L}(\mathfrak{K},\mathfrak{K}')}$. An easy argument shows that $V(e^{i\theta})$ is an isometrical isomorphism of $\mathfrak{K}$ onto $\mathfrak{K}'$ a.e. If we define the function U in $H^\infty_{\mathfrak{L}(\mathfrak{K}',\mathfrak{K})}$ such that $U(e^{i\theta}) = V(e^{-i\theta})^*$, then

$$S(e^{i\theta}) = P_{\mathfrak{H}} U(e^{i\theta})_{|\mathfrak{H}} \quad \text{a.e.,}$$

$U(e^{i\theta})$ is an isometrical isomorphism of $\mathfrak{K}'$ onto $\mathfrak{K}$ a.e. and U has a meromorphic pseudo-continuation of bounded type to D_e.

If $\mathfrak{H}$ is finite-dimensional, then both $\mathfrak{K}$ and $\mathfrak{K}'$ are finite-dimensional and have the same dimension. Thus, $\mathfrak{K} \ominus \mathfrak{H}$ and $\mathfrak{K}' \ominus \mathfrak{H}$ can be identified and U is seen to give the promised unitary-valued matrix function dilating S.

If $\mathfrak{H}$ is infinite-dimensional, then it may not be true that $\mathfrak{K} \ominus \mathfrak{H}$ and $\mathfrak{K}' \ominus \mathfrak{H}$ have the same dimension. If one wants to secure a dilation of the promised form, then taking $U \oplus I_{\mathfrak{G}}$, where $\mathfrak{G}$ is an infinite dimensional Hilbert space, enables us to identify $\mathfrak{K} \oplus \mathfrak{G}$ with $\mathfrak{K}' \oplus \mathfrak{G}$ such that $\mathfrak{H}$ is identified with $\mathfrak{H}$. Thus we obtain a unitary-valued dilation for S of the desired form.

We conclude this section with a remark. If the Hilbert space $\mathfrak{H}$ is equipped with a conjugation, denoted by —, then S in $H^\infty_{\mathfrak{L}(\mathfrak{H})}$ is said to be a *real function* if $\bar{S}(e^{i\theta}) = S(e^{-i\theta})$ a.e. It is reasonable to ask that an inner dilation, if one exists, also be real. One can check that this is indeed the case by verifying that each of the operators constructed in the preceding argument is a real operator. In particular, since outer factors are unique, the outer factorization of a real non-negative operator-valued function will consist of real operators.

3. Classical inner dilations. In this section we discuss a more classical approach to obtaining an inner dilation for a matrix-valued analytic function which is closer in spirit to that used by the electrical engineers.

Given S in $H^\infty_{\mathfrak{L}(C^n)}$ satisfying the hypotheses of the theorem we seek to construct a unitary inner function U of the form

$$U(z)=\begin{pmatrix} S(z) & B(z)\\ A(z) & C(z)\end{pmatrix}.$$

First, we choose an outer function A satisfying

$$A(e^{i\theta})^* A(e^{i\theta}) = I - S(e^{i\theta})^* S(e^{i\theta}).$$

Secondly, we choose B_0 such that $B_0^\sim$ defined by $B_0^\sim(e^{i\theta}) = B_0(e^{-i\theta})^*$ is outer and B_0 satisfies

$$B_0(e^{i\theta}) B_0(e^{i\theta})^* = I - S(e^{i\theta}) S(e^{i\theta})^*.$$

Proposition 4 guarantees the existence of both A and B_0. Next we define C_0 such that

$$C_0(e^{i\theta}) = -B_0(e^{i\theta}) R(e^{i\theta})^{-1} S(e^{i\theta})^* A(e^{i\theta}),$$

where $R(e^{i\theta})^{-1}$ is the standard pseudo-inverse*) of $I - S(e^{i\theta})^* S(e^{i\theta})$. One can check that the matrix function U thus defined has unitary boundary values. There is a problem, however; U is not, in general, analytic. However, with a bit of effort one can show that R has a meromorphic pseudo-continuation of bounded type to $\boldsymbol{D}_e$. Consequently, there exists a scalar inner function φ such that φC_0 lies in $H^\infty_{\mathfrak{L}(C^n)}$. If we set $B = \varphi B_0$ and $C = \varphi C_0$, then the matrix function defined by U has the desired properties.

We conclude with a question we have been unable to answer. What functions can occur as the 1—1 entries in a matrix representation for an operator-valued inner function? If the inner function has a scalar factor, then it is necessary and sufficient that the function have a meromorphic pseudo-continuation of bounded type to $\boldsymbol{D}_e$. If we assume the function to have modulus uniformly bounded away from 1, then a modification of the preceding construction argument can be used to provide an answer.

Added in proof. We wish to thank Professor Douglas N. Clark for calling our attention to the fact that results similar to those obtained in this paper are announced in D. Z. Arov, Darlington's method for dissipative systems, *Doklady Akad. Nauk SSSR*, **201** (1971), 559—562 (Russian); *Soviet Phys. Doklady*, **16** (1971), 954—956 (1972) (English translation).

*) The standard pseudo-inverse of a matrix M is the matrix X with the property that XM is the projection onto [null M]$^\perp$ and MX is the pr ojection onto range [M].

Bibliography

[1] P. DEWILDE, Roomy scattering matrix synthesis, *Technical Report* (Berkeley, 1971).
[2] R. G. DOUGLAS, On factoring positive operator functions, *J. Math. Mech.*, **16** (1966), 119—126.
[3] R. G. DOUGLAS, H. S. SHAPIRO, and A. L. SHIELDS, Cyclic vectors and invariant subspaces for the backward shift operator, *Ann. Inst. Fourier*, **20** (1970), 37—76.
[4] H. HELSON, *Lectures on Invariant Subspaces*, Academic Press (New York, 1964).
[5] P. D. LAX and R. S. PHILLIPS, *Scattering Theory*, Academic Press (New York, 1967).
[6] R. W. NEWCOMBE, *Linear Multiport Synthesis*, McGraw Hill (New York, 1966).
[7] M. ROSENBLUM and J. ROVNYAK, The factorization problem for non-negative operator valued functions, *Bull. Amer. Math. Soc.*, **77** (1971), 287—318.
[8] H. S. SHAPIRO, Generalized analytic continuation, *Symposium on Theor. Phys. and Math.*, vol. 8, Plenum Press (New York, 1968), 151—163.
[9] B. SZ.-NAGY and C. FOIAŞ, *Harmonic Analysis of Operators on Hilbert Space*, Akadémiai Kiadó (Budapest, 1970).

(Received September 1, 1972)

Chain conditions on set mappings and free sets

By PAUL ERDŐS and ANDRÁS HAJNAL in Budapest
and ATTILA MÁTÉ in Szeged

Dedicated to Professor Béla Sz.-Nagy on his 60th birthday

1. Introduction

Given an infinite set E, call a function f a *set mapping* (*on* E) if f maps E into $\mathscr{P}(E)$ (the set of all subsets of E) and is such that $x \notin f(x)$ for any $x \in E$. Call two elements x and y of E *independent* (*with respect to* f) if $x \notin f(y)$ and $y \notin f(x)$. Say that a subset X of E is *free* (*with respect to* f) if any two elements of X are independendent. S. RUZIEVICZ [12] conjectured and A. HAJNAL [5] proved the following: if there is a cardinal $\mu < |E|$ (this latter donetes the cardinality of the set E) such that $|f(x)| < \mu$ holds for any $x \in E$, then there is a free set $X \subseteq E$ of cardinallity $|E|$. A well-known example shows that the weaker assumption $|f(x)| < |E|$ does not even guarantee the existence of an independent couple. Still, one can weaken the cardinality assumption on $f(x)$ while ensuring the existence of a large free set by imposing structural restrictions on the range of f. Before we discuss these restrictions, we need a short review of

Notations and terminology. We work within ZFC, i. e. Zermelo-Fraenkel set theory with the Axiom of Choice. We use the usual notations of set theory, although there is one point to be stressed: $\subset$ *always denotes strict inclusion*, i. e.

$$x \subset y \leftrightarrow x \subseteqq y \quad \& \quad x \neq y.$$

As mentioned above, $|x|$ is the cardinallity, and $\mathscr{P}(x)$ is the set of all subsets, of the set x; dom(g) denotes the domain and range(g) the range of the function g. The definition of the *full inverse image* $f^{-1}(x)$ of a set X under the set mapping f will be given in Definition 3. 3.

An ordinal is the set of its predecessors, and cardinals are identified with their initial ordinals. A cardinal μ is *inaccessible* if it is a regular cardinal such that for every cardinal $\nu < \mu$ we have $2^\nu < \mu$. Finally, the *weak cardinal power* $\mu^{\underset{\smile}{\nu}}$ is defined as $\bigcup_{\xi < \nu} \mu^{|\xi|}$.

By *Martin's Axiom* we mean, as usual, Proposition **A** in [10, p. 150] (cf. also [16]), i.e. the following proposition:

For any notion C of forcing that satisfies the countable antichain condition (often called countable chain condition), and for any set F of cardinality $<2^{\aleph_0}$ of dense open subsets of C, there exists an F-generic filter.

As is well known, this proposition is consistent with ZFC$+2^{\aleph_0}>\aleph_1$, provided ZFC itself is consistent (see [16]). Furthermore, it is to be noted that Martin's Axiom implies the regularity of $2^{\aleph_0}$ (see [10, Corollary 2 on p. 164]).

The following concept plays a key role in the discussions below.

Definition 1. 1. Given an ordinal η, we say that the set S satisfies the *η-chain condition (with respect to inclusion)* if there is no sequence $\langle s_\alpha:\alpha<\eta\rangle$ of elements of S such that $s_\alpha\subset s_\beta$ whenever $\alpha<\beta<\eta$.

2. Assumptions on the set mapping and results

Throughout this paper $\varkappa$ will denote a regular cardinal and we shall assume that $E=\varkappa$; this amounts to the same as assuming that the cardinality of E is $\varkappa$. We shall consider a subset S of $\mathcal{P}(\varkappa)$ satisfying one of the two conditions below. These are the conditions we shall usually impose upon the set mapping f with $S=\text{range}\,(f)$.

(A) Every element of S has cardinality $<\varkappa$, and for each subset F of $\varkappa$, the set $\{s\cap F:s\in S\}$ satisfies the $\varkappa$-chain condition (see Definition 1. 1).

The other condition is apparently weaker:

(B) Every element of S has cardinality $<\varkappa$, and, moreover, for any $\tau<\varkappa$ and any decomposition $\varkappa=\bigcup_{\alpha<\tau}E_\alpha$ of $\varkappa$ into mutually disjoint sets E_α of cardinality $\varkappa$, there is an ordinal $\gamma<\tau$ and a set $F\subseteqq E_\gamma$ of cardinality $\varkappa$ such that the set $\{s\cap F:s\in S\}$ satisfies the $\varkappa$-chain condition.

As we mentioned just before, it is clear that (A) implies (B). But the converse is not true:

Lemma 2. 1. (B) *does not imply* (A).

Proof. Split $\varkappa$ into two disjoint sets, each of cardinality $\varkappa$: $X=\{\xi_\alpha:\alpha<\varkappa\}$ and $Y=\{\eta_\alpha:\alpha<\varkappa\}$. Take

$$S=\{\{\xi_\alpha\}\cup\{\eta_\beta:\beta<\alpha\}:\alpha<\varkappa\}.$$

Then it is easy to check that (B) holds but (A) does not. In fact, as for (A), the set $\{s\cap Y:s\in S\}$ does not satisfy the $\varkappa$-chain condition. As for (B), take a sequence

$\langle E_\alpha:\alpha<\tau\rangle$ of sets as described, and take a $\gamma<\tau$ such that $|E_\gamma\cap X|=\varkappa$; then (B) is fulfilled with $F=E_\gamma\cap X$. The proof is complete.

The following condition is an alternative form of (B). The slight change is that here $\bigcup_{\alpha<\tau}E_\alpha$ need only be "almost equal" to $\varkappa$, and we do not require that the sets E_α have cardinality $\varkappa$:

(B′) Every element of S has cardinality $<\varkappa$, and, moreover, for any $\tau<\varkappa$ and any sequence of mutually disjoint subsets E_α, $\alpha<\tau$, of $\varkappa$ such that

$$|\varkappa-\bigcup_{\alpha<\tau}E_\alpha|<\varkappa,$$

there is an ordinal $\gamma<\tau$ and a set $F\subseteqq E_\gamma$ of cardinality $\varkappa$ such that the set $\{s\cap F:s\in S\}$ satisfies the $\varkappa$-chain condition.

Next we prove

Lemma 2. 2. (B) *and* (B′) *are equivalent.*

Proof. It is clear that (B′) implies (B). We show that the converse is also true. To this end assume that (B) holds and, furthermore, let $\langle E_\alpha:\alpha<\tau\rangle$ be such a sequence as is described in (B′). We may suppose that all the sets E_α have cardinality $\varkappa$, as those of cardinality $<\varkappa$ can simply be omitted. Assume first that

$$(*)\qquad |\varkappa-\bigcup_{\alpha<\tau}E_\alpha|\leqq|\tau|$$

holds. Take mutually disjoint sets E'_α such that $\varkappa=\bigcup_{\alpha<\tau}E'_\alpha$ and such that $E_\alpha\subseteqq E'_\alpha$ and $|E'_\alpha-E_\alpha|\leqq 1$ hold for any $\alpha<\tau$. By (B) there is a $\gamma<\varkappa$ and an $F'\subseteqq E'_\gamma$ of cardinality $\varkappa$ such that $\{s\cap F':s\in S\}$ satisfies the $\varkappa$-chain condition. It is then clear that the conclusion of (B′) holds with $F=F'\cap E_\gamma$. This establishes the desired result in case $(*)$ holds. If this is not the case, then start with splitting an arbitrary one of the sets E_α into $|\varkappa-\bigcup_{\alpha<\tau}E_\alpha|$ mutually disjoint sets of cardinality $\varkappa$; then $(*)$ will hold, and the argument above can be used. The proof is complete.

We shall prove that (B) implies the existence of a countably infinite free set. This has essentially been proved by G. Fodor and A. Máté [3, Theorem 2 on p. 4], although under slightly stronger assumptions (condition (B) of that paper requires somewhat more than condition (B) of ours). If $\varkappa$ is inaccessible and weakly compact, then (B) implies the existence of a free set of cardinality $\varkappa$. (A cardinal is *weakly compact* if it is not *strongly incompact*; for the definition see [6, p. 312] or [14, Definition 1. 11 on p. 61]; cf. also Theorem 1.13 in [14, p. 62].) Not even (A) implies, however, the existence of a free set of cardinality $\varkappa$ in the following cases (in cases (i) and (ii) we actually prove somewhat more): (i) for some cardinal λ, $\varkappa=\lambda^+=2^\gamma$; (ii) $\varkappa=2^{\aleph_0}$ and Martin's Axiom holds (see at the end of the Introduction); and (iii) there exists a Souslin $\varkappa$-tree (the definition of Souslin tree is given in the next section).

A theorem of R. B. JENSEN [8, p. 292] says that, assuming the Axiom of Constructibility (see [4]), there exists a Souslin $\varkappa$-tree if and only if $\varkappa$ is not weakly compact. So, this last result in case (iii) and the result mentioned just before imply that, under the assumption of the Axiom of Constructibility, (A) (or (B)) implies the existence of a free set of cardinality $\varkappa$ if and only if $\varkappa$ is weakly compact (in the constructible universe every weakly compact cardinal is inaccessible — see [6, Theorems 2 and 3 on pp. 315—316]). Finally we mention that the results and problems of this paper are related to Problem 73 in [1, p. 46]. P. ERDŐS and A. HAJNAL have recently solved this problem affirmatively. Their proof has not yet been published, only an announcement was made in [2, p. 16].

3. Existence of "large" free sets

The aim of this section is to establish those of our results which confirm that condition (B) described in the preceding section implies the existence of large free sets. The basic tool of these proofs is trees, so here we recall a few concepts concerning them (we refer to [7] as an excellent expository paper on trees; references to other sources are given there).

A partially ordered set $\langle T, \prec\rangle$ is called a *tree* if for any $x \in T$ the set of predecessors of x, $\mathrm{pr}(x)=\mathrm{pr}(x, \langle T, \prec\rangle)=\{y \in T: y \prec x\}$ is wellordered by $\prec$ (we assume that $\prec$ is irreflexive). We sometimes write T instead of $\langle T, \prec\rangle$. A subset linearly ordered by $\prec$ of T is called a *chain (of* or *in T)*, a maximal chain a *branch*, and, furthermore, a (not necessarily proper) lower segment of a branch is said to be a *path*. An *antichain* is a set of elements mutually incomparable in $\prec$ of T. For any $x \in T$, $\mathrm{o}(x)=\mathrm{o}(x, \langle T, \prec\rangle)$ denotes the order type of $\mathrm{pr}(x)$, and for any ordinal α the set $\{x \in T: \mathrm{o}(x)=\alpha\}$ is called the αth *level* of T. The *length* of a tree T is $\bigcup\{\alpha+1:$ the αth level of T is not empty$\}$. An α*-tree* is a tree with length α.

Assume μ is a cardinal. An *Aronszajn μ-tree* is a μ-tree such that each chain and each level has cardinality $<\mu$. A *Souslin μ-tree* is a μ-tree such that each chain and antichain has cardinality $<\mu$. μ is said to have the *Tree Property* if there exists no Aronszajn μ-tree. It is well known that, assuming μ is inaccessible, μ has the tree property if and only if μ is weakly compact (for a proof, see e.g. [14, Theorem 1. 13 on p. 62]). We need some further notions:

Definition 3. 1. A tree $\langle T', \prec'\rangle$ is called a *loose end-extension* of another one, $\langle T, \prec\rangle$, if $T \subseteqq T'$, the restriction of $\prec'$ to T equals $\prec$, and, furthermore, every branch of T' includes a branch of T as a lower segment.

Assume now that we are given a regular cardinal $\varkappa$ and a set mapping f on $\varkappa$. The following concepts depend on and f, although the terms introduced will not stress this explicitly:

Definition 3. 2. A tree $\langle T, \prec\rangle$ such that $T \subseteq \varkappa$ is called *free* if each of its branches is a free set (with respect to f).

Now, for a tree $\langle T, \prec\rangle$ and for a path p of T denote by $\mathrm{ims}(p, T)$ the set of immediate successors in $\prec$ of p. (Note that the empty set is also a path.)

Definition 3. 3. A free tree T is called *regular* if for every nonmaximal path p of T we have $|\mathrm{ims}(p, T)| < \varkappa$ and

$$\bigcap\{f^{-1}(\{\xi\}): \xi \in \mathrm{ims}(p, T)\} = 0,$$

where

$$f^{-1}(X) \overset{\text{def}}{=} \{\xi < \varkappa: X \cap f(\xi) \neq 0\} \qquad (X \subseteq \varkappa).$$

An important consequence of this definition is given by the next lemma. (We need this lemma only for $p=0$, but it does not require any extra effort to establish it for any p.)

Lemma 3. 4. *Assume T is a regular free tree and p is a path in T. Then, with b running over all branches of T, we have*

$$\bigcap\{f^{-1}(b-p): \ p \subseteq b\} = 0.$$

Proof. Given any $\xi < \varkappa$, we are going to show that ξ does not belong to the above intersection. To this end, consider those path p' in T for which $p \subseteq p'$ and

$$\xi \notin f^{-1}(p'-p).$$

Note that p itself is such a path, and, by Zorn's lemma, there is a path that is maximal among those having this property. Assume that p' is already such a maximal one. If p' is a branch, then we are ready. If not, then let $\eta \in \mathrm{ims}(p', T)$ be such that $\xi \notin f^{-1}(\{\eta\})$ (there is such an η by the regularity of T). Then

$$\xi \notin f^{-1}(p \cup \{\eta\} - p),$$

which contradicts the maximality of p'. The proof is complete.

Say that a regular free tree T is less than another one, T', if T' is a loose end-extension of T. It follows easily from Zorn's lemma that, under this partial ordering, there is a maximal regular free tree (note that the empty tree is a regular free tree, and so is the union of a linearly ordered set of regular free trees). Our key result in this section says that a maximal regular free tree cannot be too small provided condition (B) (see the preceding section) holds for $S = \mathrm{range}(f)$:

Theorem 3. 5. *Assume condition* (B) *holds for* $S = \mathrm{range}(f)$. *Let $\langle T, \prec\rangle$ be a regular free tree having less than $\varkappa$ branches and such that $|T| < \varkappa$. Then T has a proper loose end-extension that is also a regular free tree.*

For the proof we need a simple lemma, which occurs in [3] and [11]. It is important for this lemma that we assumed $\varkappa$ to be a regular cardinal.

Lemma 3. 6. *Let H be a set such that each of its elements has cardinality $<\varkappa$ and such that $|\bigcup H|\geqq\varkappa$, and assume that H satisfies the $\varkappa$-chain condition (with respect to inclusion). Then there is a subset X of cardinality $<\varkappa$ of $\bigcup H$ such that $X\subsetneqq h$ holds for any $h\in H$.*

Proof. H can be considered as a set partially ordered by inclusion. By a well-known theorem of F. Hausdorff, there is a maximal linearly ordered subset of H, say K. By another of his theorems, there is a wellordered subset M of K that is cofinal to K. As H satisfies the $\varkappa$-chain condition, we must have $|M|<\varkappa$. Now take an arbitrary element t of $\bigcup H-\bigcup M$, and put $X=\bigcup M\cup\{t\}$. It is clear that this set satisfies the requirements of the lemma.

Now we establish the announced theorem.

Proof of Theorem 3. 5. Let $\langle b_\alpha:\alpha<\tau\rangle$ $(\tau<\varkappa)$ be an enumeration of the branches in T, and put

$$G_\alpha=\varkappa-f^{-1}(b_\alpha)$$

and

$$E_\alpha=G_\alpha-M-\bigcup_{\beta<\alpha}G_\beta \qquad (\alpha<\tau),$$

where

$$M=T\cup\bigcup\{f(\xi):\xi\in T\}.$$

It follows from Lemma 3.4 with $p=0$ that $\bigcup_{\alpha<\tau}E_\alpha=\varkappa-M$. It is clear that here $|M|<\varkappa$, as we assumed both $|T|<\varkappa$ and $|f(\xi)|<\varkappa$ for any $\xi<\varkappa$ (this latter as a part of (B)). So, in view of (B′) (which holds by its equivalence to (B), as established in Lemma 2. 2) we can see that there exists an ordinal $\gamma<\tau$ and a set $F\subseteq E_\gamma$ of cardinality $\varkappa$ such that

$$\{f(\xi)\cap F:\xi<\varkappa\}$$

satisfies the $\varkappa$-chain condition. So, by the lemma just proved, there is a set $X\subseteq F$ of cardinality $<\varkappa$ such that $X\subsetneqq f(\xi)\cap F$ holds for any $\xi<\varkappa$, i.e. such that

$$\bigcap\{f^{-1}(\{\delta\}):\delta\in X\}=0.$$

Make the set $T'=T\cup X$ a tree by stipulating that T' is a loose end-extension of T such that $X=\operatorname{ims}(b_\gamma,T')$. It is clear that these stipulations define T' as a tree unambiguously, and, moreover, that T' is a regular free tree. This completes the proof.

As we mentioned above, there exists a maximal regular free tree. By the theorem just proved, such a tree either must have cardinality $\varkappa$ or it must have at least $\varkappa$ branches. In either case, it cannot have only very short branches; as a branch is a free set, we can thus establish the existence of a large free set. We first prove

Theorem 3. 7. *Assume that $\mu<\varkappa$ is a cardinal such that $\nu^{\underline{\mu}}<\varkappa$ holds for any cardinal $\nu<\varkappa$. Then any maximal regular free tree has a branch of cardinality $\geqq\mu$.*

Proof. Take a maximal regular free tree T, and assume that each branch of T has cardinality $<\mu$. Then, in view of Theorem 3. 5., T must have at least $\varkappa$ branches (indeed, if T has less than $\varkappa$ branches, then we also have: $|T|\leqq$the sum of the cardinalities of all branches of $T<\varkappa$). Let $\eta\leqq\mu$ be the least ordinal such that the tree $T|\eta$ has at least $\varkappa$ branches ($T|\eta$ is, by definition, obtained from T by omitting each of its elements in or above the ηth level). Then each level in $T|\eta$ has cardinality $<\varkappa$. In fact, let $\alpha<\eta$. Then $T|\alpha$ must have less than $\varkappa$ branches by the minimality of η. Since for any path p of T we have $|\text{ims}\,(p, T)|<\varkappa$ (this is stipulated in the definition of a regular free tree), we can conclude from here by the regularity of $\varkappa$ that the αth level in T has cardinality $<\varkappa$.

So there is a cardinal $\nu<\varkappa$ such that each level in $T|\eta$ has cardinality $\leqq\nu$. Therefore, noting that each branch in $T|\eta$ has cardinality $<\mu$, the number of branches in $T|\eta$ is at most

$$\bigcup\{\nu^{|\xi|}:\xi\leqq\eta \quad \& \quad \xi<\mu\}\leqq\nu^{\underline{\mu}}<\varkappa,$$

which is a contradiction, proving the theorem.

From this theorem we can immediately conclude

Theorem 3. 8. *Assume that $\varkappa$ is an infinite regular cardinal and condition* (B) *holds with $S=$*range(f). *Then*

(i) *there exists a free set of cardinality $\aleph_0$;*

(ii) *if μ is a cardinal $<\varkappa$ such that for every cardinal $\nu<\varkappa$ we have $\nu^{\underline{\mu}}<\varkappa$, then there exists a free set of cardinality μ;*

(iii) *if $\varkappa$ is inaccessible and weakly compact, then there exists a free set of cardinality $\varkappa$.*

Proof. (ii) directly follows from the preceding theorem. We establish (iii). As $\varkappa$ is inaccessible in this case, the assumptions of the preceding theorem hold for any cardinal $\mu<\varkappa$; so a maximal regular free tree T must have length $\geqq\varkappa$. As $|\text{ims}\,(p, T)|<\varkappa$ holds for any path p in T (cf. Definition 3. 3.), it follows from the inaccessibility of $\varkappa$ that for any $\alpha<\varkappa$ the αth level in T has cardinality $<\varkappa$. As $\varkappa$ has the tree property (cf. e.g. [14, Theorem 1. 13 on *p.* 62]; note that although not mentioned there, this is also true in case $\varkappa=\aleph_0$ — see [9]), T must have a branch of cardinality $\varkappa$. This being a free set, (iii) is proved. Finally, in case $\varkappa>\aleph_0$ (i) follows from (ii), and in case $\varkappa=\aleph_0$ it follows from (iii) (there is no harm in considering $\aleph_0$ inaccessible). The proof is complete.

4. Nonexistence of "too large" free sets

In many cases we can prove that condition (B) (and even the stronger condition (A)) does not ensure the existence of a free set of cardinality $\varkappa$. But we cannot prove even in the simplest case that there is a cardinal $\mu<\varkappa$ such that (B) does not imply the existence of a free set of cardinality μ. We start with the simple

Theorem 4. 1. *Assume that $\varkappa$ is a regular cardinal such that there exists a Souslin $\varkappa$-tree. Then condition* (A) *with* $S=\operatorname{range}(f)$ *does not imply the existence of a free set of cardinality $\varkappa$.*

Proof. Assume $\langle\varkappa, \prec\rangle$ is a Souslin $\varkappa$-tree, and for any $\xi<\varkappa$ put

$$f(\xi) = \{\alpha<\varkappa\colon \alpha\prec\xi\}\,(=\operatorname{pr}(\xi)).$$

A subset of $\varkappa$ is free with respect to this f exactly if it is an antichain in $\langle\varkappa, \prec\rangle$; so there is no free set of cardinality $\varkappa$. We are going to show that $S=\operatorname{range}(f)$ satisfies condition (A). Assume the contrary, and let F be a subset of $\varkappa$ and $\langle\xi_\alpha\colon\alpha<\varkappa\rangle$ a sequence of ordinals $<\varkappa$ such that

$$f(\xi_\alpha)\cap F\subset f(\xi_\beta)\cap F$$

holds for any $\alpha<\beta<\varkappa$ ($\subset$ indicates strict inclusion). Then it is easy to see that

$$\bigcup_{\alpha<\varkappa}(f(\xi_\alpha)\cap F)$$

is a chain of cardinality $\varkappa$ of $\langle\varkappa, \prec\rangle$. This contradicts the fact that the latter is a Souslin $\varkappa$-tree. The proof is complete.

Next we show that, under the assumption of the Generalized Continuum Hypothesis, condition (A) does not guarantee the existence of a free set of cardinality $\varkappa$ if $\varkappa$ is a successor cardinal. Actually, we prove more:

Theorem 4. 2. *Assume $\varkappa$ and λ are infinite cardinals such that $\varkappa=2^\lambda$ and either* (i) $\varkappa=\lambda^+$, *or* (ii) $\lambda=\aleph_0$ *and Martin's Axiom holds. Then there is a set $S\subseteq\mathscr{P}(\varkappa)$ of cardinality $\varkappa$ satisfying condition* (A) *of Section* 2 *such that for any set $S'\subseteq S$ of cardinality $\varkappa$ we have $|\varkappa-\bigcup S'|<\lambda$.*

An obvious consequence of this is

Corollary 4. 3. *Assume that either* (i) *or* (ii) *of the preceding theorem holds. Then condition* (A) *with* $S=\operatorname{range}(f)$ *does not imply the existence of a free set of cardinality $\varkappa$.*

For the proof of the above theorem we need the following

Lemma 4. 4. *Assume that either* (i) *or* (ii) *of the preceding theorem holds. Let $\eta<\varkappa$ be an ordinal and $\langle A_\xi\colon\xi<\eta\rangle$ a sequence of sets of cardinality λ. Then there is* *a set* $B_\eta \subseteqq \bigcup_{\xi<\eta} A_\xi$ *such that* B_η *meets each* A_ξ, $\xi<\eta$, *but does not include any of them.*

Proof. *Ad* (i). This case, due to F. BERNSTEIN, is well known and simple. We may assume that $\eta \leqq \lambda$; indeed, if this is not the case, the we can rearrange the sequence $\langle A_\xi : \xi<\eta\rangle$. Now define x_ξ and y_ξ by transfinite recursion so that $x_\xi \neq y_\xi$ and

$$x_\xi, y_\xi \in A_\xi - \{x_\alpha, y_\alpha \colon \alpha < \xi\} \qquad (\xi < \eta),$$

and take $B_\eta = \{x_\xi : \xi<\eta\}$.

Ad (ii). Put

$$C = H(\textstyle\bigcup_{\xi<\eta} A_\xi, 2),$$

that is, let C be the set of all functions with values 0 or 1 the domains of which are finite subsets of $\bigcup_{\xi<\eta} A_\xi$. Consider C as partially ordered by inclusion; then, as is well known, C is a notion of forcing satisfying the countable antichain condition (often called countable chain condition; cf. [13, Lemma 10. 3 on p. 372] — Shoenfield's terminology differs from ours, so that in order to agree with it we should order C by reverse inclusion). The set

$$D_\xi = \{p \in C \colon \exists x, y \in A_\xi [x, y \in \operatorname{dom}(p) \ \& \ p(x)=0 \ \& \ p(y)=1]\}$$

is dense open for any $\xi<\eta$; so, by Martin's Axiom, there exists a $\{D_\xi : \xi<\eta\}$-generic filter G. The set

$$B_\eta = \{x \in \operatorname{dom}(\textstyle\bigcup G) \colon (\bigcup G)(x) = 1\}$$

satisfies our requirements (note that $\bigcup G$ is a function the domain of which is included in $\bigcup_{\xi<\eta} A_\xi$). The lemma is proved.

Proof of Theorem 4. 2. We deal with cases (i) and (ii) simultaneously. Let $\langle A_\xi : \xi<\varkappa\rangle$ be an enumeration of all subsets of cardinality λ of $\varkappa$, and for each $\eta<\varkappa$ define B_η as described in the lemma just proved. Put $S = \{B_\eta : \eta<\varkappa\}$. We show that S satisfies (A). It is clear that each element of S has cardinality $<\varkappa$; assume that the rest of (A) does not hold, and let F be a subset of $\varkappa$ such that $\{B_\eta \cap F : \eta<\varkappa\}$ does not satisfy the $\varkappa$-chain condition. Then it is easy to see that there exists a set $I \subseteqq \varkappa$ of cardinality $\varkappa$ such that

$$B_\alpha \cap F \subset B_\beta \cap F$$

holds for any $\alpha, \beta \in I$ with $\alpha<\beta$. Then for any $\alpha \in I$ with $|\alpha \cap I| \geqq \lambda$ we obviously have $|B_\alpha \cap F| \geqq \lambda$; so, for some $\xi<\varkappa$, we have $A_\xi \subseteqq B_\alpha \cap F$. Pick an $\eta \in I$ with $\eta>\alpha, \xi$. Then $A_\xi \nsubseteqq B_\eta$, which contradicts the assumption $A_\xi \subseteqq B_\alpha \cap F \subset B_\eta \cap F$. Thus we have shown that S satisfies (A).

Now take any subset S' of cardinality $\varkappa$ of S. We are about to show that $|\varkappa - \bigcup S'| < \lambda$. Assume the contrary; then there exists a $\xi<\varkappa$ such that $A_\xi \subseteqq \varkappa - \bigcup S'$.

Take a $B_\eta \in S'$ with $\eta > \xi$. Then $A_\xi \cap B_\eta \neq 0$, which is a contradiction. The theorem is proved.

We conclude this paper by pointing out a few problems. As mentioned in Section 2, our discussion is complete as far as the existence of free sets of cardinality $\varkappa$ is concerned in case we assume the Axiom of Constructibility. But without such an assumption many problems remain open. The simples-sounding one is

Problem 1. Assume $\varkappa = \aleph_1$ and $2^{\aleph_0} > \aleph_1$. Does then (A) or (B) with $S = \text{range}(f)$ imply the existence of a free set of cardinality $\varkappa$?

One may try to solve this problem even under the assumption of Martin's Axiom; the answer is unknown to us. Nothing is known about the nonexistence of free sets of a cardinality less than $\varkappa$. E.g. one might ask

Problem 2. Assume $\varkappa = 2^{\aleph_0} > \aleph_1$, and assume that Martin's Axiom holds. Does then (A) or (B) with $S = \text{range}(f)$ imply the existence of a free set of an uncountable cardinality?

It is a well-known result of R. M. Solovay that it is consistent relatively to the existence of a measurable cardinal that $2^{\aleph_0}$ be real-valued measurable (see [15, Theorem 2 and Proposition 1 on pp. 398—399]; cf. also the remark on p. 67 in [14]). The fact that a real-valued measurable cardinal always has the tree property (see [14, Theorem 1. 16 on p. 67]) makes the following problem interesting:

Problem 3. Assume that $\varkappa = 2^{\aleph_0}$, and, furthermore, that $\varkappa$ is real-valued measurable. Does then (A) or (B) with $S = \text{range}(f)$ imply the existence of a free set of cardinality $\varkappa$?

Added in proof. When the paper had already been in print, we obtained the following results, which go a long way in settling Problems 1—3. For an ordinal η, denote by (A_η) the assertion that for the set mapping $f\colon \varkappa \to \mathcal{P}(\varkappa)$ we have $|f(\alpha)| < \varkappa$ whenever $\alpha < \varkappa$, and, for each subset F of $\varkappa$, the set $\{f(\alpha) \cap F\colon \alpha < \varkappa\}$ satisfies the η-chain condition. Then the following propositions are consistent relatively to ZFC: (i) $2^{\aleph_0} = \varkappa =$ anything reasonable, (A_{ω_1}) holds for f, and there is no free set of cardinality $\aleph_1$; (ii) $2^{\aleph_0} = \varkappa$ is real-valued measurable, (A_{ω_1}) holds for f, and there is no free set of cardinality $\aleph_1$. The following propositions are theorems of ZFC: (iii) If $\varkappa = 2^{\aleph_0} = \aleph_2$ and Martin's Axiom holds, then there is an f satisfying $(A_{\omega+1})$ (in fact, $\forall \xi, \eta[\xi < \eta < \omega_2 \to |f(\xi) \cap f(\eta)| < \aleph_0]$) such that there is no free set of cardinality $\aleph_2$; (iv) If $\varkappa = \lambda^+ = 2^\lambda$ and $\text{cf}(\lambda) > \omega$, then there is an f satisfying $(A_\varkappa)$ such that there is no free set of order type $\lambda + \omega$; (v) If $\varkappa = \lambda^+ = 2^\lambda$ and λ is regular, then there is an f such that $(A_{\lambda+1})$ holds (in fact, $\forall \xi, \eta[\xi < \eta < \varkappa \to |f(\xi) \cap f(\eta)| < \lambda)$ and there is no free set of cardinality $\varkappa$.

References

[1] P. Erdős and A. Hajnal, Unsolved problems in set theory, in: *Axiomatic set theory, Proc. of Symposia in Pure Math.* Vol. XIII Part I, ed. by the Amer. Math. Soc., Providence, Rhode Island, 1971; 17—48.

[2] P. Erdős and A. Hajnal, Unsolved and solved problems in set theory, preprint, Calgary, Alberta, Can., 1971.

[3] G. Fodor and A. Máté, On the structure of set mappings and the existence of free sets, *Acta Sci. Math.*, **26** (1965), 1—7.

[4] K. Gödel, *The consistency of the axiom of choice and of the generalized continuum-hypothesis with the axioms of set theory*, Annals of Mathematics Study No. 3, Princeton University Press (Princeton, N. J., 1940).

[5] A. Hajnal, Proof of a conjecture of S. Ruziewicz, *Fund. Math.*, **50** (1961), 123—128.

[6] W. Hanf, Incompactness in languages with infinitely long expressions, *Fund. Math.*, **53** (1964), 309—324.

[7] T. J. Jech, Trees, *J. Symbolic Logic*, **36** (1971), 1—14.

[8] R. B. Jensen, The fine structure of the constructible hierarchy, *Ann. Math. Logic*, **4** (1972), 229—308.

[9] D. König, Über eine Schlußweise aus dem Endlichen ins Unendliche, *Acta Sci. Math.*, **3** (1927), 121—130.

[10] D. A. Martin and R. M. Solovay, Internal Cohen extensions, *Ann. Math. Logic*, **2** (1970), 143—178.

[11] A. Máté, On the cardinality of strongly almost disjoint systems, *Acta Math. Acad. Sci. Hung.*, **18** (1967), 1—3.

[12] S. Ruziewicz, Une généralistation d'un théroeme de M. Sierpinski, *Publications Math. de l'Université de Belgrade*, **5** (1936), 23—27.

[13] J. R. Shoenfield, Unramified forcing, in: *Axiomatic set theory, Proc. of Symposia in Pure Math.* Vol. XIII Part I, ed. by the Amer. Math. Soc., Providence, Rhode Island, 1971; 357—382.

[14] J. H. Silver, Some applications of model theory in set theory, *Ann. Math. Logic*, **3** (1971), 45—110.

[15] R. M. Solovay, Real-valued measurable cardinals, in: *Axiomatic set theory, Proc. of Symposia in Pure Math.* Vol. XIII. Part I, ed. by the Amer. Math. Soc., Providence, Rhode Island, 1971 397—428.

[16] R. M. Solovay and S. Tennenbaum, Iterated Cohen extensions and Souslin's problem (to appear).

MATHEMATICAL INSTITUTE OF THE HUNGARIAN ACADEMY OF SCIENCES, BUDAPEST
BOLYAI INSTITUTE, SZEGED, HUNGARY

(Received August 21, 1972)

Principal sequences and stationary sets

By GÉZA FODOR in Szeged

Dedicated to Professor Béla Sz.-Nagy on his 60th birthday

1. Introduction. Assume we are given two regular cardinals ϱ and $\varkappa$ with $\varrho<\varkappa$, and let us assign to each ordinal $\xi<\varkappa$ that is cofinal to ϱ a principal sequence of order type ϱ, i.e. a sequence of ordinals that is strictly monotonic and continuous and tends to ξ. A natural question to ask whether there are many among these sequences which have the same beginning. This type of question was first raised by B. Rotman. The method we use in answering this problem as far as we can is a generalization of that of ROTMAN [4] and FODOR [3].

2. Notations and terminology. Our considerations below can most naturally be carried out in the framework of Zermelo—Fraenkel set theory with the Axiom of Choice. We use the usual set-theoretical conventions and notations. Thus we consider an ordinal as the set of its predecessors, and a cardinal is identified with its initial ordinal. Ordinals are usually denoted by lower case Greek letters, and the sign + denotes ordinal addition. dom (f) denotes the domain, and $f \restriction x$ the restriction to the set x of the function f. Braces $\{,\}$ are used to define sets, and angle brackets $\langle,\rangle$ to define functions; that is, if f is a function, then by definition we have

$$f=\langle f(x)\colon\ x\in \mathrm{dom}(f)\rangle.$$

A *sequence* is a function whose domain is an ordinal. The word monotone is meant in the wider sense; if we mean strictly monotone, then we say so. A monotone sequence $\langle \xi_\alpha : \alpha<\eta\rangle$ of ordinals tends to ξ if $\xi=\bigcup_{\alpha<\eta}\xi_\alpha$. A principal sequence is a strictly monotone sequence of ordinals that is continuous; here a monotone sequence $\langle \xi_\alpha : \alpha<\eta\rangle$ of ordinals is called continuous if $\xi_\alpha=\bigcup_{\beta<\alpha}\xi_{\beta+1}$ holds for any $\alpha<\eta$.

Given a regular cardinal $\varkappa>\omega$, a set $X\subseteq\varkappa$ is called *stationary* (in $\varkappa$) if it meets every *closed unbounded* subset of $\varkappa$. (Here the topology on $\varkappa$ is that generated by its natural ordering; unbounded means cofinal to $\varkappa$.) A key fact proved by the author is that the nonstationary subsets of $\varkappa$ form a normal ideal on $\varkappa$; that is, if a function f sending ordinals to ordinals is called *regressive* whenever $f(\alpha)<\alpha$ holds for all nonzero α in its domain, then we have (see [2, Satz 2 on p. 141]):

2. 1. Theorem. *Any regressive function the domain of which is a stationary set (in some fixed regular cardinal) is constant on a stationary set.*

3. The next lemma, which is a generalization of results in [3] and [4], will be useful in dealing with our main problem outlined in the introduction. The lemma may be interesting in itself, and it has other applications as well (it enables one to derive special cases of Solovay's decomposition theorem [5] which says that every set stationary in a regular cardinal $\varkappa>\omega$ can be split into $\varkappa$ stationary sets (cf. [3]).

3. 1. Lemma. *Assume ϱ and $\varkappa$ are two infinite regular cardinals, $\varrho<\varkappa$, and S is a set stationary in $\varkappa$ of ordinals less than $\varkappa$ and cofinal to ϱ. Assume, further, that for every $\xi\in S$ $\langle f_\alpha(\xi):\alpha<\varrho\rangle$ is a monotone sequence tending to ξ. Then there exists an ordinal $\nu_0<\varrho$ such that, for every ν with $\nu_0\leqq\nu<\varrho$ there is a set $F(\nu)\subseteqq\varkappa$ of cardinality $\varkappa$ such that the set*

$$\{\xi\in S:f_\nu(\xi)=\gamma\}$$

is a stationary for each $\gamma\in F(\nu)$.

Proof. Call a function f mapping a subset of $\varkappa$ into $\varkappa$ essentially bounded if there exists an ordinal $\alpha<\varkappa$ such that

$$\{\xi\in\operatorname{dom}(f):f(\xi)<\alpha\}$$

is nonstationary. We assert that at least one of the functions f_ν is not essentially bounded. In fact, assuming that α_ν is an essential bound for f_ν, we see that $\alpha=\bigcup_{\nu<\varrho}\alpha_\nu$ is a common essential bound for each f_ν; moreover, it is clear that the set

$$H=\{\xi\in S:\exists\nu<\varrho[f_\nu(\xi)>\alpha]\}\ (\subseteqq\bigcup_{\nu<\varrho}\{\xi\in S:f_\nu(\xi)>\alpha_\nu\})$$

is nonstationary. On the other hand $H=S-\alpha$, as by our assumptions we have $\xi=\bigcup_{\nu<\varrho}f_\nu(\xi)$ for every $\xi\in S$. This contradiction implies that there is indeed a $\nu_0<\varrho$ such that f_{ν_0} is not essentially bounded.

Now take an arbitrary ν with $\nu_0\leqq\nu<\varrho$. Obviously, f_ν is not essentially bounded, as $f_{\nu_0}(\xi)\leqq f_\nu(\xi)$ holds for every $\xi\in S$. Put $F(\nu)=\{\gamma<\varkappa:\{\xi\in S:f_\nu(\xi)=\gamma\}$ is stationary$\}$. We are going to show that $F(\nu)$ is cofinal to $\varkappa$, and so it has cardinality $\varkappa$. Indeed, take any $\alpha<\varkappa$. As f_ν is not essentially bounded, the set

$$X=\{\xi\in S:f_\nu(\xi)\geqq\alpha\}$$

is stationary. So, applying Theorem 2. 1 to the function $f\restriction X$, we see that there is a $\gamma\geqq\alpha$ such that the set

$$\{\xi\in S:f_\nu(\xi)=\gamma\}$$

is stationary. This means that $\gamma \in F(\nu)$, showing that $F(\nu)$ is cofinal to $\varkappa$, as asserted above.

This completes the proof.

4. Now we are in position to establish the main result of these notes:

4. 1. Theorem. *Assume ϱ and $\varkappa$ are two infinite cardinals such that $\varrho < \varkappa$ and $\tau^\sigma < \varkappa$ holds for any two cardinals σ and τ with $\sigma < \varrho$, $\tau < \varkappa$. Let S be a set stationary in $\varkappa$ of ordinals less than $\varkappa$ and cofinal to ϱ. For every $\xi \in S$, let $\langle f_\nu(\xi) : \nu < \varrho \rangle$ be a principal sequence tending to ξ. Then there exists an ordinal $\nu_0 < \varrho$ such that for every ν with $\nu_0 \leqq \nu < \varrho$ there is a set $G(\nu)$ of principal sequences of type ν of ordinals such that the set $\{\xi \in S : \langle f_\mu(\xi) : \mu < \nu \rangle = s\}$ is stationary in $\varkappa$ for each $s \in G(\nu)$.*

Proof. Observe that it is enough to prove this theorem for any ν of form $\eta + 1$ such that $\nu_0 < \nu < \varrho$, for some $\nu_0 < \varrho$. In fact if η is a limit ordinal and $G(\eta + 1)$ has already been defined in a way complying with the requirements of the theorem, then we can take

$$G(\eta) = \{s \restriction \eta : \ s \in G(\eta + 1)\}.$$

One should only note here that the cardinality of $G(\eta)$ is $\varkappa$, for if $s_1, s_2 \in G(\eta + 1)$ and $s_1 \restriction \eta = s_2 \restriction \eta$ then also $s_1 = s_2$ by the continuity of these sequences.

Now take ν_0 to be that of the preceding lemma, and let η be an ordinal with $\nu_0 \leqq \eta < \varrho$. Select an arbitrary $\gamma \in F(\eta)$, this latter set having been defined in the preceding lemma. In view of the assumption that $\tau^\sigma < \varkappa$ holds for any two cardinals $\tau < \varkappa$ and $\sigma < \varrho$ we can see that there are less than $\varkappa$ different ones among the sequences

$$\langle f_\mu(\xi) : \ \mu \leqq \eta \rangle$$

as ξ runs over the elements of the stationary set

$$\{\xi \in S : \ f_\eta(\xi) = \gamma\}.$$

So there is a sequence s_γ such that the set

$$\{\xi \in S : \langle f_\mu(\xi) : \mu \leqq \eta \rangle = s_\gamma\}$$

is stationary. Set

$$G(\eta + 1) = \{s_\gamma : \ \gamma \in F(\eta)\}.$$

The proof is complete.

4. 2. It is not difficult to see that the assumption that $\tau^\sigma < \varkappa$ holds for any two cardinals τ and ϱ with $\tau < \varkappa$, $\sigma < \varrho$ is essential. More exactly, we have

Theorem. *Assume ϱ and $\varkappa$ are two infinite regular cardinals, and $\tau < \varkappa$ and $\sigma < \varrho$ are cardinals such that $\tau^\sigma \geqq \varkappa$. Let S be the set of all ordinals ξ cofinal to ϱ for which $\tau + \sigma \leqq \xi < \varkappa$. Then for every $\xi \in S$, there exists a principal sequences s_ξ tending* *to ξ such that for any sequence s of type $\geqq\sigma$ of ordinals there is at most one $\xi\in S$ for which s is an initial segment of s_ξ.*

Proof. Clearly, one can take $\varkappa$ different principal sequences of type $\sigma+1$ of ordinals $\leqq\tau$; let $\langle t_\xi:\xi\in S\rangle$ be an enumeration of these sequences. For any $\xi\in S$ continue the sequence t_ξ to a principal sequence s_ξ of type ϱ that tends to ξ. The proof is complete.

5. If s_1 and s_2 are two principal sequences of type ϱ, where ϱ is a cardinal, then s_1 and s_2 (or, rather, their ranges) have less than ϱ elements in common. It would, however, be interesting to know the answer to the following

Problem. *Assume ϱ and $\varkappa$ are regular cardinals, $\varrho<\varkappa$. Let S be the set of all ordinals $<\varkappa$ that are cofinal to ϱ. Is then there a principal sequence s_ξ for each $\xi\in S$ such that, for some cardinal $\lambda<\varrho$, the sequences s_η and s_ξ have less than λ elements in common whenever ξ and η are different elements of S.*

By Theorem 4. 1, the answer is in the negative if $\tau^\sigma<\varkappa$ holds for any cardinals $\sigma<\varrho$ and $\tau<\varkappa$. We conjecture, however, that the answer is always in the negative, and remains to be so even if we take S to be any stationary set of ordinals $<\varkappa$ that are cofinal to ϱ.

References

[1] H. BACHMANN, *Transfinite Zahlen,* Ergebnisse der Mathematik und ihrer Grenzgebiete, Band I, 1967.

[2] G. FODOR, Eine Bemerkung zur Theorie der regressiven Funktionen, *Acta Sci. Math.,* 17 (1956), 204—206.

[3] G. FODOR, On stationary sets and regressive functions, *Acta Sci. Math.,* 27 (1966), 105—110.

[4] B. ROTMAN, A note on principal sequences, *Proc. of the Glasgow Math. Association,* 6 (1963—64), 133—135.

[5] R. M. SOLOVAY, Real-valued measurable cardinals, in: Axiomatic Set Theory, *Proceedings of the Symposia in Pure Math.,* Vol. XIII, Part I, Amer. Math. Soc. (Providence, Rhode Island, 1971), 397—428.

BOLYAI INSTITUTE
SZEGED, HUNGARY

(Received September 12, 1972)

Factorisations étranges

Par CIPRIAN FOIAŞ à Bucarest (Roumanie)

Au 60-ième anniversaire de mon Maître, collaborateur et ami, le Professeur Béla Sz.-Nagy

1. L'un des achèvements auxquels a conduit la théorie des dilatations unitaires, initée il y a 20 ans par B. Sz.-Nagy [1], a été la découverte, en 1964, du rôle joué par les factorisations régulières parmi les factorisations de la fonction caractéristique d'une contraction (voir [3] et [4], Ch. VII). Notamment, à tout sous-espace non-banal $\mathfrak{H}_1$, invariant pour une contraction c.n.u. T dans l'espace de Hilbert $\mathfrak{H}$ [1]), il correspond une factorisation *régulière* [2])

$$\Theta_T(\lambda) = \Theta_2(\lambda)\Theta_1(\lambda) \qquad (|\lambda|<1) \tag{1}$$

de la fonction caractéristique de T en fonctions analytiques contractives telles que les parties de pures de Θ_1 et Θ_2 coïncident aux fonctions caractéristiques de la restriction T_1 de T à $\mathfrak{H}_1$ et de la compression T_2 de T à $\mathfrak{H}_2 = \mathfrak{H} \ominus \mathfrak{H}_1$. De plus, à toute factorisation régulière non-banale (1) de Θ_T correspond un sous-espace non-banal $\mathfrak{H}_1$ de $\mathfrak{H}$ invariant pour T, tel que Θ_1 et Θ_2 soient liées à T_1 et T_2 de la manière indiquée auparavant.

Dans la conférence [2] au Congrès International des Mathématiciens de Nice (1970), B. Sz.-Nagy a soulevé la question de savoir s'il existe des factorisations *non-régulières* (1), pour lesquelles il existe néanmoins un sous-espace $\mathfrak{H}_1$ de $\mathfrak{H}$ tel que les parties pures de Θ_1 et Θ_2 coïncident avec les fonctions caractéristiques de la restriction T_1 de T à $\mathfrak{H}_1$ et de la compression T_2 de T à $\mathfrak{H}_2 = \mathfrak{H} \ominus \mathfrak{H}_1$.

A cette occasion, B. Sz.-Nagy a remarqué que si $T \in C_{\cdot 0}$, telles factorisations (que nous appellerons *étranges*) n'existent pas. Ainsi le problème de l'existence des factorisations étranges concerne seulement les fonctions caractéristiques des contractions $T \notin C_{\cdot 0}$, dont on sait qu'elles ont des factorisations (non-banales) régulières (voir [4], Ch. VII, §3 et §5).

1) Tous les espaces seront supposés complexes.

2) Pour cette notion aussi bien que pour toutes les autres, qui ne sont pas définies explicitement, nous renvoyons à [4].

Le but de cette Note est d'apporter quelques réponses partielles au problème des factorisations étranges des fonctions caractéristiques.

2. Proposition 1. *Il existe une fonction caractéristique ayant des factorisations étranges.*

Démonstration. Soit $\mathfrak{E}$ un espace de Hilbert séparable de dimension infinie. Définissons la fonction analytique (en fait constante) contractive $\{\mathfrak{E}, \mathfrak{E}, \Theta(\lambda)\}$ par

$$\Theta(\lambda) = -\frac{1}{\sqrt{2}} I_{\mathfrak{E}} \quad \text{pour tout} \quad |\lambda| < 1, \tag{2}$$

où $I_{\mathfrak{E}}$ désigne l'opérateur identique de $\mathfrak{E}$ et soit T une contraction c.n.u. dont la fonction caractéristique coïncide à la fonction définie par (2). Il est manifeste que T est une somme orthogonale infinie dénombrable de contractions c.n.u. ayant toutes comme fonction caractéristique la fonction numérique constante $\equiv -\frac{1}{2}$. L'opérateur S dans

$$\mathfrak{G} = (H^2 \oplus L^2) \ominus \left\{-\frac{1}{\sqrt{2}} u \oplus \frac{1}{\sqrt{2}} u \colon u \in H^2\right\}$$

défini par

$$S^*(u \oplus v)(e^{it}) = e^{-it}\big(u(e^{it}) - u(0)\big) \oplus e^{-it} v$$

a la fonction caractéristique $\equiv -\frac{1}{2}$ (voir [4], Ch. VI, § 3). Soient dans $\mathfrak{G}$

$$e_n = \begin{cases} 0 \oplus e^{nit} & \text{si } n < 0, \\ \frac{1}{\sqrt{2}} e^{int} \oplus \frac{1}{\sqrt{2}} e^{int} & \text{si } n \geqq 0. \end{cases}$$

Alors $\{e_n\}_{n=-\infty}^{\infty}$ est une base orthonormale dans $\mathfrak{G}$ et

$$Se_n = e_{n+1} \qquad (n \neq -1) \quad \text{et} \quad Se_{-1} = \tfrac{1}{2} e_0. \tag{3}$$

Ainsi la fonction Θ définie par (2) coïncide avec la fonction caractéristique de l'opérateur

$$T = S \oplus S \oplus \cdots \oplus S \cdots$$

(où S est la translation bilatérale pondérée définie par (3)) dans

$$\mathfrak{H} = \mathfrak{G} \oplus \mathfrak{G} \oplus \cdots \oplus \mathfrak{G} \cdots$$

Soit $\mathfrak{G}_1$ les sous-espace de $\mathfrak{G}$ engendré par $\{e_n\}_{n=0}^{\infty}$ et posons

$$\mathfrak{H}_1 = \mathfrak{G}_1 \oplus \mathfrak{G}_1 \oplus \cdots \oplus \mathfrak{G}_1 \oplus \cdots.$$

Alors $\mathfrak{H}_1$ est invariant à T, la restriction T_1 de T à $\mathfrak{H}_1$ est une translation unilatérale de multiplicité infinie tandis que la compression T_2 de T à $\mathfrak{H}_2 = \mathfrak{H} \ominus \mathfrak{H}_1$ est l'adjointe d'une telle translation.

Définissons maintenant les fonctions analytiques (en fait constantes) contractives $\{\mathfrak{E}, \mathfrak{E}\oplus\mathfrak{E}\oplus\mathfrak{E}, \Theta_1(\lambda)\}$ et $\{\mathfrak{E}\oplus\mathfrak{E}\oplus\mathfrak{E}, \mathfrak{E}, \Theta_2(\lambda)\}$ par

$$\text{(4)} \qquad \begin{cases} \Theta_1(\lambda)\,e = \dfrac{1}{\sqrt{2}}e\oplus\dfrac{1}{2}e\oplus\dfrac{1}{2}e & (|\lambda|<1;\ e\in\mathfrak{E}),\\ \Theta_2(\lambda)(e\oplus f\oplus g) = -e & (|\lambda|<1;\ e, f, g\in\mathfrak{E}). \end{cases}$$

Il est manifeste que les fonctions (2), (4) vérifient (1) et que la partie pure Θ_1^0 de Θ_1 est la fonction $\{\{0\}, \{e\oplus f\oplus g: e, f, g\in\mathfrak{E}, \sqrt{2}e+f+g = 0\}, 0\}$, tandis que la partie pure Θ_2^0 de Θ_2 est $\{\{0\}\oplus\mathfrak{E}\oplus\mathfrak{E}, \{0\}, 0\}$. Puisque $\mathfrak{E}\oplus\mathfrak{E}$ a la même dimension que $\mathfrak{E}$ on conclut que Θ_i^0 coïncide avec la fonction caractéristique de T_i ($i=1, 2$). Pour démontrer que la factorisation que nous venons de construire est une factorisation étrange il ne nous reste donc qu'à montrer que cette factorisation n'est pas régulière. Or on a

$$\Delta_1(t) = [I_{\mathfrak{E}} - \Theta_1(e^{it})^*\,\Theta_1(e^{it})]^{1/2} \equiv 0$$

tandis que

$$\Delta_2(t) = [I_{\mathfrak{E}\oplus\mathfrak{E}\oplus\mathfrak{E}} - \Theta_2(e^{it})^*\,\Theta(e^{it})]^{1/2}$$

est la projection orthogonale de $\mathfrak{E}\oplus\mathfrak{E}\oplus\mathfrak{E}$ sur $\{0\}\oplus\mathfrak{E}\oplus\mathfrak{E}$, quel que soit t. Donc

$$\text{(5)} \qquad \overline{\Delta_1\mathfrak{E}}\oplus\overline{\Delta_2(\mathfrak{E}\oplus\mathfrak{E}\oplus\mathfrak{E})} = \{0\}\oplus(\{0\}\oplus\mathfrak{E}\oplus\mathfrak{E})$$

et

$$\text{(6)} \qquad \overline{\Delta_1\mathfrak{E}\oplus\Delta_2\,\Theta_1\mathfrak{E}} = \{0\}\oplus(\{0\}\oplus\{f\oplus f:\ f\in\mathfrak{E}\}).$$

En comparant (5) à (6) on déduit que

$$\overline{\Delta_1(t)E}\oplus\overline{\Delta_2(t)(\mathfrak{E}\oplus\mathfrak{E}\oplus\mathfrak{E})} \neq \overline{\{\Delta_1(t)e\oplus\Delta_2(t)\,\Theta_1(e^{it})e : e\in\mathfrak{E}\}}$$

(pour tout t), d'où le fait que la factorisation que nous venons de considérer n'est pas régulière.

Remarque. L'opérateur T dont la fonction caractéristique coïncide avec la fonction (2), contient une translation unilatérale de multiplicité infinie ainsi que l'adjoint de tel opérateur. Il y a une certaine analogie entre l'exemple ci-dessus et celui donné dans [4], Ch. VII, no. 4. 3, concernant l'existence des diviseurs réguliers qui ne sont pas forts. Toutefois la multiplicité infinie des translations dans l'exemple ci-dessus semble être essentielle (voir la Proposition 2 ci-dessous).

3. Lemma. *Soit*

$$\text{(7)} \qquad A = A_2 A_1$$

une factorisation régulière, ou $A: \mathfrak{A} \to \mathfrak{A}_*$, $A_1: \mathfrak{A} \to \mathfrak{B}$ *et* $A_2: \mathfrak{B} \to \mathfrak{A}_*$, *sont des contractions. Soit*

$$(7') \qquad A = B_2 B_1$$

une autre factorisation telle que

$$B_2 = U_* A_2 V_2, \quad B_1 = V_1 A_1 U$$

où U_*, V_2, V_1, U *sont des opérateurs unitaires (dans les espaces correspondants). Si l'opérateur de défaut* D_A *est de rang fini alors* (7') *est aussi une factorisation régulière.*

Démonstration. Comme

$$(8) \qquad D_{B_2} = V_2^* D_{A_2} V_2, \quad D_{B_1} = U^* D_{A_1} U$$

on a [3])

$$(9) \qquad \dim(\overline{D_{B_1}\mathfrak{A}} \oplus \overline{D_{B_2}\mathfrak{B}}) = \dim(\overline{D_{A_1}\mathfrak{A}} \oplus \overline{D_{A_2}\mathfrak{B}}) = \dim \overline{D_A \mathfrak{A}} < \infty.$$

Soient Z et Z' définis pour $a \in D_A \mathfrak{A}$ par

$$Za = D_{A_1} a \oplus D_{A_2} A_1 a, \quad Z'a = D_{B_1} a \oplus D_{B_2} B_1 a.$$

Les opérateurs Z et Z' se prolongent en des isométries de $\overline{D_A \mathfrak{A}}$ dans $\mathfrak{D} = \overline{D_{A_1}\mathfrak{A}} \oplus \overline{D_{A_2}\mathfrak{B}}$ et $\mathfrak{D}' = \overline{D_{B_1}\mathfrak{A}} \oplus \overline{D_{B_2}\mathfrak{B}}$, selon les cas (voir [4], Ch. VII, § 3). La régularité de (7) veut dire que Z est unitaire. Par suite $W = Z'Z^*$ est une isométrie de $\mathfrak{D}$ dans $\mathfrak{D}'$. En vertu de (9), W doit être unitaire, ce qui oblige Z' d'être aussi unitaire. Ceci signifie que la factorisation (7') est régulière.

4. Proposition 2. *Soit* T *une contraction c.n.u. aux indices de défaut finis. Alors sa fonction caractéristique* Θ_T *n'admet pas des factorisations étranges.*

Démonstration. Soit $\{E^m, E^n, \Theta_T(\lambda)\}$ la fonction caractéristique de T et soit (1) une factorisation étrange de Θ_T; soit de plus $\mathfrak{F}$ l'espace but de $\Theta_1(\lambda)$ ($|\lambda|<1$) et l'espace source de $\Theta_2(\lambda)$ ($|\lambda|<1$). Avec les notations du no. 1, soit

$$T = \begin{pmatrix} T_1 & * \\ 0 & T_2 \end{pmatrix} \text{ dans } \mathfrak{H} = \mathfrak{H}_1 \oplus \mathfrak{H}_2.$$

En vertu de [4], Ch. VII. no. 4. 5, T_1 et T_2 sont aux indices de défaut finis. Comme la partie pure Θ_1^0 et Θ_1 a comme espace source un sous-espace de E^m de dimension $\mathfrak{d}_{T_1}$ et comme espace but un sous-espace de $\mathfrak{F}$ de dimension $\mathfrak{d}_{T_1^*}$ il résulte que

$$\dim \mathfrak{F} = \mathfrak{d}_{T_1^*} + m - \mathfrak{d}_{T_1}.$$

On peut donc supposer que $\mathfrak{F} = E^p$ où $p = \mathfrak{d}_{T_1^*} + m - \mathfrak{d}_{T_1}$. Soit maintenant

$$(10) \qquad \Theta_T = A_2(\lambda) A_1(\lambda) \qquad (|\lambda| < 1)$$

---

[3]) Pour la derniere égalité voir [4], Ch. VII, Prop. 3.2.d.

la factorisation régulière qui correspond au sous-espace invariant $\mathfrak{H}_1$. Par la même raison que ci-dessus on peut supposer que l'espace intermédiaire dans (10) est E^p. Puisque les parties pures de $\Theta_1(\lambda)$ et $A_1(\lambda)$ coïncident avec $\Theta_{T_1}(\lambda)$, et que celles de $\Theta_2(\lambda)$ et $A_2(\lambda)$ coïncident avec $\Theta_{T_2}(\lambda)$, il résulte aisément, du fait que p est fini, que $\Theta_1(\lambda)$ et $A_1(\lambda)$ coïncident, de même que $\Theta_2(\lambda)$ et $A_2(\lambda)$. Donc il existe des opérateurs unitaires U (dans E^m), V_1, V_2 (dans E^p) et U_* (dans E^n) tels que

$$\Theta_1(\lambda)=V_1A_1(\lambda)U,\quad \Theta_2(\lambda)=U_*A_2(\lambda)V_2 \qquad (|\lambda|<1).$$

Passant à la circonférence on a, presque partout en t, que la factorisation

$$\Theta_T(e^{it}) = A_2(e^{it})A_1(e^{it})$$

est régulière (voir [4], Ch. VII, no. 3. 1). Du lemme précédent on déduit alors que la factorisation

$$\Theta_T(e^{it}) = \Theta_2(e^{it})\Theta_1(e^{it}).$$

est aussi régulière, presque partout en t, donc la factorisation (1) est régulière: contradiction!

Par conséquent dans le cas envisagé il n'existe point des factorisations étranges.

5. En analysant la démonstration de la proposition précédente et en améliorant un peu le lemme du no. 3 on aboutit aisément à la conclusion que la Proposition 2 peut être généralisée en la suivante

Proposition 3. *Soit T une contraction c.n.u. telle que le rang de*

$$\Delta(t) = [I-\Theta_T(e^{it})^*\Theta_T(e^{it})]^{1/2}$$

soit fini presque partout en t. Alors Θ_T n'admet pas des factorisations étranges.

Il est possible que la conclusion de cette proposition se conserve sous l'hypothèse plus faible que le rang de $\Delta(t)$ soit fini aux points d'un ensemble de mesure positive.

Bibliographie

[1] B. Sz.-Nagy, Sur les contractions de l'espace de Hilbert, *Acta Sci. Math.*, **15** (1953), 87—92.

[2] ——— Sous-espaces invariants d'un opérateur et factorisation de sa fonction caractéristique, *Actes du Congrès Intern. Math. Nice* (1970), 2, 459—465.

[3] B. Sz.-Nagy et C. Foiaş, Une caractérisation des sous-espaces invariants pour une contraction de l'espace de Hilbert, *C. R. Acad. Sci. Paris*, **258** (1964), 3426—3429.

[4] ——— *Analyse harmonique des opérateurs de l'espace de Hilbert* (Budapest, 1967); voir aussi les traductions russe (Moscou, 1970) et anglaise (Budapest, 1970).

(Reçu le 2 octobre 1972)

On the greatest zero of an orthogonal polynomial. I

By GÉZA FREUD in Budapest

Dedicated to Prof. Béla Szőkefalvi-Nagy on the occasion of his 60th birthday

0. Introduction

Let $w(x)$ $(-\infty<x<\infty)$ be an *even* weight function, and let $\{p_n(w;x)=$ $=\gamma_n(w)x^n+\cdots;\ n=0,1,\ldots\}$ be the sequence of orthonormal polynomials with respect to w, i.e.

(1) $$\int_{-\infty}^{\infty} p_m(w;x)p_n(w;x)\,w(x)\,dx=\begin{cases}0 & (m\neq n),\\ 1 & (m=n).\end{cases}$$

Moreover, let $X_n(w)=x_{1n}(w)$ be the greatest zero of $p_n(w;x)$. In part 1 of the present note we express the order of $X_n(w)$ with the aid of the sequence $\{\gamma_\nu(w)\}$ (see Theorem 1). After deducing some lemmas in part 2, we apply this result in part 3 to the weight

(2) $$w_{\varrho,2k}(x)=|x|^\varrho e^{-x^{2k}}$$

where $\varrho\geqq 0$ and k is a positive integer. We prove the estimate $\gamma_{\nu-1}(w_{\varrho,2k})/\gamma_\nu(w_{\varrho,2k})=$ $=O(\nu^{1/2k})$ which seems far from being trivial and conclude from it that

(3) $$X_n(w_{\varrho,2k})\sim n^{1/2k}\sim\gamma_{n-1}(w_{\varrho,2k})/\gamma_n(w_{\varrho,2k}).$$

The relation (3) has several interesting implications in approximation theory; we hope to return to them soon.

1. An inequality on $X_n(w)$

Theorem 1. *For every even weight function $w(x)$ we have*

(4) $$\max_{1\leqq k\leqq n-1}\frac{\gamma_{k-1}(w)}{\gamma_k(w)}\leqq X_n(w)\leqq 2\max_{1\leqq k\leqq n-1}\frac{\gamma_{k-1}(w)}{\gamma_k(w)}.$$

Remarks. a) Let $w_0(x)=(1-x^2)^{-1/2}$ with support $[-1,1]$. Then the first three orthogonal polynomials are $\frac{1}{\sqrt{\pi}}, \sqrt{\frac{2}{\pi}}\,x, \sqrt{\frac{2}{\pi}}(2x^2-1)$, i.e. $\gamma_0(w_0)/\gamma_1(w_0)=$

$= \frac{1}{\sqrt{2}} = X_2(w_0)$. This example shows that the left-hand part of inequality. (4) is precise.

b) In case $w_{02}(x) = e^{-x^2}$ the orthogonal polynomials $p_n(w)$ are the orthonormal Hermite polynomials $h_n(x)$, so that $\gamma_{n-1}(w_{02})/\gamma_n(w_{02}) = \sqrt{\frac{n}{2}}$ and $X_n(w_{02}) \approx \sqrt{2n}$. This example shows that the factor 2 on the right-hand side of (4) can not be replaced by any smaller constant.

Proof. By a classical result of P. L. CHEBYCHEV (see G. SZEGŐ [2], 7. 7. 2) we have

$$X_n(w) = \max \frac{\int\limits_{-\infty}^{\infty} x[P_{n-1}(x)]^2 w(x)\,dx}{\int\limits_{-\infty}^{\infty} [P_{n-1}(x)]^2 w(x)\,dx}, \tag{5}$$

where $P_{n-1}(x)$ runs over all polynomials of degree $\leqq n-1$. Let us represent $P_{n-1}(x)$ as

$$P_{n-1}(x) = \sum_{j=0}^{n-1} c_j p_j(w; x). \tag{6}$$

We recall that by the recursion formula applied to even w we have

$$x p_j(w; x) = \frac{\gamma_j(w)}{\gamma_{j+1}(w)} p_{j+1}(w; x) + \frac{\gamma_{j-1}(w)}{\gamma_j(w)} p_{j-1}(w; x) \tag{7}$$

(see e.g. G. FREUD [1], § I. 2).

Inserting (6) into (5) and taking (1) and (7) into consideration we obtain

$$X_n(w) = 2 \max \frac{\sum\limits_{k=1}^{n-1} \frac{\gamma_{k-1}(w)}{\gamma_k(w)} c_{k-1} c_k}{\sum\limits_{k=0}^{n-1} c_k^2}, \tag{8}$$

where all the c_k $(k = 0, 1, \ldots, n-1)$ run idependently over the reals. Inserting $c_{j-1} = c_j = 1$ and $c_k = 0$ if $k \neq j-1, j$ into the expression on the right of (8), we obtain

$$X_n(w) \geqq \frac{\gamma_{j-1}(w)}{\gamma_j(w)} \qquad (j = 1, 2, \ldots n-1). \tag{9}$$

In turn, by Cauchy's inequality for every $\{c_k\}$ we have

$$\sum_{k=1}^{n-1} \frac{\gamma_{k-1}(w)}{\gamma_k(w)} c_{k-1} c_k \leqq \max_{1 \leqq k \leqq n-1} \frac{\gamma_{k-1}(w)}{\gamma_k(w)} \sum_{k=0}^{n-1} |c_{k-1} c_k| \leqq$$

$$\leqq \max_{1 \leqq k \leqq n-1} \frac{\gamma_{k-1}(w)}{\gamma_k(w)} \sum_{k=1}^{n-1} c_k^2.$$

The left-hand side of (4) is implied by (9) and the right hand side of (4) is a consequence of (8) and (10), and so Theorem 1 is proved.

2. Lemmata

Let

(11) $$w_{\varrho\beta}(x) = |x|^{\varrho} e^{-|x|^{\beta}} \qquad (-\infty < x < \infty).$$

Lemma 1. *For every $\varrho \geqq 0$ and $\beta > 0$ we have*

(12) $$\frac{\gamma_n(w_{\varrho\beta})}{\gamma_{n-1}(w_{\varrho\beta})} = \frac{\beta}{n + \Delta_n \varrho} \int_{-\infty}^{\infty} p_n(w_{\varrho\beta}; x) p_{n-1}(w_{\varrho\beta}; x) x^{-1} |x|^{\beta} w_{\varrho\beta}(x)\, dx,$$

where

(13) $$\Delta_n = \tfrac{1}{2}[1 + (-1)^{n+1}].$$

Proof. We have

$$\int_{-\infty}^{\infty} p_n'(w_{\varrho\beta}; x) p_{n-1}(w_{\varrho\beta}; x) w_{\varrho\beta}(x)\, dx =$$

$$= \int_{-\infty}^{\infty} [n\gamma_n(w_{\varrho\beta}) x^{n-1} + \cdots] p_{n-1}(w_{\varrho\beta}; x) w_{\varrho\beta}(x)\, dx =$$

$$= \int_{-\infty}^{\infty} \left[n \frac{\gamma_n(w_{\varrho\beta})}{\gamma_{n-1}(w_{\varrho\beta})} p_{n-1}(w_{\varrho\beta}; x) + P_{n-2}(x) \right] p_{n-1}(w_{\varrho\beta}; x) w_{\varrho\beta}(x)\, dx,$$

where P_{n-2} is a polynomial of degree $\leqq n-2$. Applying the orthogonality relations (1), we get

(14) $$n \frac{\gamma_n(w_{\varrho\beta})}{\gamma_{n-1}(w_{\varrho\beta})} = \int_{-\infty}^{\infty} p_n'(w_{\varrho\beta}; x) p_{n-1}(w_{\varrho\beta}; x) w_{\varrho\beta}(x)\, dx.$$

Partial integration gives

$$\int_{-\infty}^{\infty} p_n'(w_{\varrho\beta};x)p_{n-1}(w_{\varrho\beta};x)w_{\varrho\beta}(x)\,dx = -\int_{-\infty}^{\infty} p_n(w_{\varrho\beta};x)[p_{n-1}(w_{\varrho\beta};x)w_{\varrho\beta}(x)]'\,dx =$$

$$(15)\qquad = \beta\int_{-\infty}^{\infty} p_n(w_{\varrho\beta};x)p_{n-1}(w_{\varrho\beta};x)x^{-1}|x|^{\beta}w_{\varrho\beta}(x)\,dx -$$

$$-\varrho\int_{-\infty}^{\infty} p_n(w_{\varrho\beta};x)p_{n-1}(w_{\varrho\beta};x)x^{-1}w_{\varrho\beta}(x)\,dx,$$

since $\int p_n(w)p_n'(w)w\,dx=0$ by (1).

If n is even, $p_{n-1}(w_{\varrho\beta};x)$ is odd, and so $x^{-1}p_{n-1}(w_{\varrho\beta};x)$ is a polynomial of degree $n-2$. Consequently, the second integral on the right of (15) vanishes by (1). In this way, from (15) we obtained

$$(16)\qquad \begin{aligned}&\int_{-\infty}^{\infty} p_n'(w_{\varrho\beta};x)p_{n-1}(w_{\varrho\beta};x)w_{\varrho\beta}(x)\,dx = \\ &= \beta\int_{-\infty}^{\infty} p_n(w_{\varrho\beta};x)p_{n-1}(w_{\varrho\beta};x)x^{-1}|x|^{\beta}w_{\varrho\beta}(x)\,dx. \qquad (n \text{ is even}).\end{aligned}$$

Let now n be odd. Then $p_n(w_{\varrho\beta};x)$ is odd, and so

$$x^{-1}p_n(w_{\varrho\beta};x) = \frac{\gamma_n(w_{\varrho\beta})}{\gamma_{n-1}(w_{\varrho\beta})}p_{n-1}(w_{\varrho\beta};x)+P_{n-2}(x),$$

where $P_{n-2}(x)$ is a polynomial of degree $\leqq n-2$. Using the orthogonality relation (1) we see that

$$(17)\qquad \int_{-\infty}^{\infty} p_n(w_{\varrho\beta};x)p_{n-1}(w_{\varrho\beta};x)x^{-1}w_{\varrho\beta}(x)\,dx = \frac{\gamma_n(w_{\varrho\beta})}{\gamma_{n-1}(w_{\varrho\beta})}. \qquad (n \text{ is odd})$$

From (14), (15), (16), and (17) we see that (12) holds for both even and odd integers n. Q.E.D.

Lemma 2. *For every positive integer k we have*

$$(18)\qquad \left[\frac{\gamma_{n-1}(w_{\varrho,2k})}{\gamma_n(w_{\varrho,2k})}\right]^{2k} \leqq \frac{n+\varrho\Delta_n}{2k}.$$

Remark. For $k=1$, $\varrho=0$ we have equality in (18).

Proof. We infer by induction from the recursion formula (7) that, for every positive integer l and every *even* w, we have

$$x^l p_n(w;x) = \sum_{j=0}^{n+l} A_{n,l,j}(w) p_j(w;x), \tag{19}$$

where *all coefficients* $A_{n,l,j}(w)$ *are nonnegative*.

By (7) we have

$$A_{n,1,n-1}(w) = \frac{\gamma_{n-1}(w)}{\gamma_n(w)}. \tag{20}$$

Moreover, by a repeated application of the recursion formula (7) we obtain

(21)
$$A_{n-1,2,n-1}(w) = \int_{-\infty}^{\infty} x^2 p_{n-1}^2(w;x)\, w(x)\, dx = \frac{\gamma_{n-1}^2(w)}{\gamma_n^2(w)} + \frac{\gamma_{n-2}^2(w)}{\gamma_{n-1}^2(w)} \geqq \frac{\gamma_{n-1}^2(w)}{\gamma_n^2(w)}.$$

Multiplying (19) by x^2 and then applying the special case $l=2$ of the same formula to the right-hand side, we get

$$A_{n,l+2,n-1}(w) \geqq A_{n,l,n-1}(w) A_{n-1,2,n-1}(w). \tag{22}$$

From (21) and (22) we infer by induction that

(23)
$$\int_{-\infty}^{\infty} x^{2s-1} p_n(w;x) p_{n-1}(w;x)\, w(x)\, dx = A_{n,2s-1,n-1}(w) \geqq$$
$$\geqq \left[\frac{\gamma_{n-1}(w)}{\gamma_n(w)}\right]^{2s-1} \qquad (s = 1, 2, \ldots).$$

Let us now insert $\beta=2k$ in (12) and $w=w_{\varrho,2k}$ in (23). Combining the two formulas so obtained we get (18). Q.E.D.

We introduce the moments

$$\mu_r(w) = \int_{-\infty}^{\infty} x^r w(x)\, dx \qquad (r = 0, 1, \ldots). \tag{24}$$

Lemma 3. *For every even* w, *we have*

$$[X_n(w)]^2 \geqq \mu_{2n-2}(w)/\mu_{2n-4}(w). \tag{25}$$

Proof. Denoting by $X_n(w)=x_{1n}>x_{2n}>\cdots>x_{nn}=-X_n(w)$ the zeros of $p_n(w;x)$, by the Gauss—Jacobi quadrature formula we have

$$\mu_{2n-2}(w) = \sum_{j=1}^{n} \lambda_{jn}(w) x_{jn}^{2n-2} \leqq [X_n(w)]^2 \sum_{j=1}^{n} \lambda_{jn}(w) x_{jn}^{2n-4} = [X_n(w)]^2 \mu_{2n-4}(w).$$

Q.E.D.

We can also see that the sign of equality is valid in (25) iff $n=2$.

3. Estimates for $X_n(w_{\varrho,2k})$

Theorem 2. *For every $\varrho \geqq 0$ and $\beta > 0$ we have*

$$\varliminf_{n\to\infty} n^{-1/\beta} X_n(w_{\varrho\beta}) \geqq (2/\beta)^{1/\beta}. \tag{26}$$

Proof. We have

$$\mu_{2r}(w_{\varrho\beta}) = 2\int_0^\infty x^{2r+\varrho} e^{-x^\beta}\,dx = \Gamma\left(\frac{2r+\varrho+1}{\beta}\right) \qquad (r = 0, 1, \ldots). \tag{27}$$

Insert (27) in (25) and apply Stirling's formula to get the desired result.

Theorem 3. *For every $\varrho \geqq 0$ and every positive integer k we have*

$$X_n(w_{\varrho,2k}) \leqq 2(n/2k)^{1/2k}. \tag{28}$$

Remark. We have $X_n(w_{0,2}) \approx \sqrt{2n}$. So the factor on the right of (28) cannot be replaced by any constant smaller than 2.

Proof. This is a consequence of Theorem 1 and Lemma 2. We see from Theorem 2 and Theorem 3 that

$$X_n(w_{\varrho,2k}) \sim n^{1/2k} \tag{29}$$

holds for every $\varrho \geqq 0$ and every positive integer k.

Theorem 4. *For every $\varrho \geqq 0$ and every positive integer k we have*

$$\frac{\gamma_{n-1}(w_{\varrho,2k})}{\gamma_n(w_{\varrho,2k})} \geqq 2^{-2k+1}\left(\frac{n}{2k}\right)^{1/2k}\left(1+\frac{k}{n}\right)^{\frac{2k-1}{2k}}. \tag{30}$$

Remark. From (30) and the combination of (28) and the left hand side of (4) we see that

$$\frac{\gamma_{n-1}(w_{\varrho,2k})}{\gamma_n(w_{\varrho,2k})} \sim n^{1/2k}. \tag{31}$$

Proof. Consider formula (12). The expression $p_n(w_{\varrho,2k}; x)p_{n-1}(w_{\varrho,2k}; x)x^{2k-1}$ is a polynomial of degree $2n+2k-2 < 2(n+k)-1$. Consequently the integral in (12) can be calculated by the Gauss—Jacobi quadrature formula over the zeros of $p_{n+k}(w_\varrho; x)$:

$$\begin{aligned} \frac{n}{2k}\frac{\gamma_n(w_{\varrho,2k})}{\gamma_{n-1}(w_{\varrho,2k})} &\leqq \frac{n+\varrho\Delta_n}{2k}\frac{\gamma_n(w_{\varrho,2k})}{\gamma_{n-1}(w_{\varrho,2k})} = \\ &= \sum_{j=1}^{n+k} \lambda_{j,n+k}(w_{\varrho,2k})\, x_{j,n+k}^{2k-1} p_n(w_{\varrho,2k}; x_{j,n+k}) p_{n-1}(w_{\varrho,2k}; x_{j,n+k}) \leqq \\ &\leqq [X_{n+k}(w_{\varrho,2k})]^{2k-1}\left\{\sum_{j=1}^{n+k}\lambda_{j,n+k}(w_{\varrho,2k}) p_n^2(w_{\varrho,2k}; x_{j,n+k}) \times \right. \\ &\left.\times \sum_{j=1}^{n+k}\lambda_{j,n+k}(w_{\varrho,2k}) p_{n-1}^2(w_{\varrho,2k}; x_{j,n+k})\right\}^{1/2} = [X_{n+k}(w_{\varrho,2k})]^{2k-1}, \end{aligned} \tag{32}$$

since by the quadrature formula we have

$$\sum_{j=1}^{n+k} \lambda_{j,n+k}(w_{\varrho,2k})\, p_r^2(w_{\varrho,2k}; x_{j,n+k}) = \int_{-\infty}^{\infty} p_r^2(w_{\varrho,2k}; x)\, w_{\varrho,2k}(x)\, dx = 1 \quad (r = n-1, n).$$

Inserting estimate (28) into the right-hand side of (32), we obtain the desired estimate (30) after reshuffling the factors. Q.E.D.

Literature

[1] G. FREUD, *Orthogonale Polynome.* Joint edition of Birkhäuser Verl., Deutscher Verl. der Wiss.' and Akadémiai Kiadó (Basel, Berlin, and Budapest, 1969); English translation: Pergamon Press (New York—Toronto—London, 1971).

[2] G. SZEGŐ, *Orthogonal Polynomials,* Amer. Math. Soc. Coll. Publ., vol. 23, 2nd ed. (New York, 1959).

(Received May 25, 1972)

On equational classes of unoids

By F. GÉCSEG and S. SZÉKELY in Szeged

To Professor B. Sz.-Nagy on his 60th birthday

In this paper we present an algorithm to decide for a given finite unoid $\mathfrak{A}$ and an equational class K generated by finitely many finite unoids whether or not $\mathfrak{A}$ is contained by K. This problem has an automata theoretical background. Using this algorithm one can decide for a finite automaton whether it can be given as a homomorphic image of a subautomaton of an A-direct product of smaller automata. (For an automata theoretical terminology, see [1].)

Before stating our theorem, we introduce some notions and notations.

A universal algebra $\mathfrak{A}=\langle A; F\rangle$ is called *unoid* if each operation in F is unary (see А. И. Мальцев [2]). $\mathfrak{A}$ is *finite* if both A and F are finite. Take an arbitrary polynomial xp over F. We say that xp is of *length n* if it has the form $xp=xf_1\ldots f_n(=(\ldots(xf_1)\ldots)f_n)$ $(f_1,\ldots,f_n\in F)$. A polynomial $xp=xf_1\ldots f_n$ is a *subpolynomial* of xq if $xp=xq$ or $xq=xf_1\ldots f_n f_{n+1}\ldots f_m(=xpf_{n+1}\ldots f_m)$ holds, in notation: $xp\subseteqq xq$.

Let X be an arbitrary set and $\mathfrak{X}=\langle X^{(\infty)}; F\rangle$ the free unoid generated by X. By $X^{(n}$ $(n=0,1,\ldots)$ we denote the subset of $X^{(\infty)}$ consisting of all polynomials with length not exceeding n. (Here $X^{(0)}=X$.)

A partition π of $X^{(n)}$ into disjoint subsets is called an *n-congruent partition* of $\mathfrak{X}$ if

(I) for any $x,y\in X^{(n-1)}$ and $f\in F$, $x\equiv y(\pi)$ implies $xf\equiv yf(\pi)$ and
(II) for each $x\in X^{(n)}$ there exists a $y\in X^{(n-1)}$ such that $x\equiv y(\pi)$.

It should be noted that if π is an n-congruent partition of $\mathfrak{X}$ then it can be extended uniquely to a congruent (m-congruent, $m\geqq n$) partition of $\mathfrak{X}$. For the congruent extension of an n-congruent partition π we use the notation π^*. Furthermore, $\mathfrak{X}/\pi^*$ denotes the factor unoid induced by π^*.

Now we are ready to state our

Theorem. *Let $\mathfrak{A}_i=\langle A_i; F\rangle$ $(i=1,\ldots,k)$ and $\mathfrak{A}=\langle A; F\rangle$ be finite unoids. Moreover, let $\langle a_1,\ldots,a_l\rangle$ be a generating system of $\mathfrak{A}$, $X=\langle x_1,\ldots,x_l\rangle$ a set of symbols* *and* $m = \max\langle \bar{\bar{A}}, \bar{\bar{A}}_1, \dots, \bar{\bar{A}}_k\rangle$. *Then* $\mathfrak{A}$ *is contained in the equational class generated by* $\langle \mathfrak{A}_i | i = 1, \dots, k\rangle$ *if and only if there exist m-congruent partitions* $\pi_1^{(m)}, \dots, \pi_r^{(m)}$ *and* $\pi^{(m)}$ *of* $\mathfrak{X} = \langle X^{(\infty)}; F\rangle$ *such that for their extentions* $\pi_1^{(m^r)}, \dots, \pi_r^{(m^r)}$ *and* $\pi^{(m^r)}$ *to* m^r*-congruent partitions the following hold:*

(1) $\pi_1^{(m^r)} \cap \dots \cap \pi_r^{(m^r)} \subseteq \pi^{(m^r)}$,

(2) $\mathfrak{X}/\pi^{(m)*} \cong \mathfrak{A}$ *and each* $\mathfrak{X}/\pi_j^{(m)*}$ $(1 \leqq j \leqq r)$ *is isomorphic to a subunoid of the unoids* $\mathfrak{A}_i$ $(i = 1, \dots, k)$.

Proof. Let us suppose that $\mathfrak{A}$ is contained in the equational class generated by $\langle \mathfrak{A}_i | i = 1, \dots, k\rangle$. Denote by $S(\langle \mathfrak{A}_i | i = 1, \dots, k\rangle)$ the class of all unoids isomorphic to a subunoid of the unoids $\mathfrak{A}_i$ $(i = 1, \dots, k)$. Then there exists a subdirect product $\mathfrak{B} = \langle B; F\rangle$ of unoids from $S(\langle \mathfrak{A}_i | i = 1, \dots, k\rangle)$ such that $\mathfrak{A}$ is a homomorphic image of $\mathfrak{B}$ under a suitable homomorphism φ. Let $\langle b_1, \dots, b_l\rangle$ be a subset of B for which $b_j\varphi = a_j$ $(j = 1, \dots, l)$ hold. It is obvious that the subunoid $\mathfrak{B}'$ of $\mathfrak{B}$ generated by $\langle b_1, \dots, b_l\rangle$ can be given in the subdirect form $\mathfrak{B}' = \mathfrak{B}_1 \times \dots \times \mathfrak{B}_s \times \dots (\mathfrak{B}_1, \dots$ $\dots, \mathfrak{B}_s, \dots \in S(\langle \mathfrak{A}_i | i = 1, \dots, k\rangle))$ and the restriction of φ to $\mathfrak{B}'$ (which is denoted by the same φ) is a homomorphism of $\mathfrak{B}'$ onto $\mathfrak{A}$.

Now take the homomorphism ψ of $\mathfrak{X}$ onto $\mathfrak{B}'$ for which $x_j\psi = b_j$ $(j = 1, \dots, l)$ hold. Then $\psi\varphi$ is a homomorphism of $\mathfrak{X}$ onto $\mathfrak{A}$ such that $x_j(\psi\varphi) = a_j$ $(j = 1, \dots, l)$.

Let us define partitions ϱ_t $(t = 1, \dots, s, \dots)$ on $\mathfrak{B}'$ in the following way:

$$(c_1, \dots, c_t, \dots) \equiv (c_1', \dots, c_t', \dots)(\varrho_t)$$

$((c_1, \dots, c_t, \dots), (c_1', \dots, c_t', \dots) \in B')$ if and only if $c_t = c_t'$. Moreover, take the partition ϱ on $\mathfrak{B}'$ given as follows: $b \equiv b'(\varrho)$ $(b, b' \in B')$ if and only if $b\varphi = b'\varphi$. It can easily be seen that the number of all classes of ϱ_t is equal to the cardinality of B_t. A similar statement is valid for ϱ and A. It is also clear that the intersection of the partitions ϱ_t is the trivial partition ι on $\mathfrak{B}'$.

Now take the following partitions π_t $(t = 1, \dots, s, \dots)$ and π on $X^{(m)}: x \equiv y(\pi_t)$ and $x \equiv y(\pi)$ $(x, y \in X^{(m)})$ if and only if $x\psi \equiv y\psi(\varrho_t)$ and $x\psi \equiv y\psi(\varrho)$, respectively. We show that π_t and π are m-congruent partitions of $\mathfrak{X}$. It can be proved by an easy computation that condition (I) of m-congruence holds for π_t $(t = 1, \dots, s, \dots)$ and π. It remains to be shown that condition (II) is also satisfied by these partitions.

Take an arbitrary polynomial xp from $X^{(m)} \setminus X^{(m-1)}$. Since the number of classes of π_t is less than or equal to m, there are two different subpolynomials xp' and xp'' of xp such that $xp' \subset xp''$ and $xp' \equiv xp''(\pi_t)$. (Here xp' can be x.) Therefore, there exists a polynomial $x'w$ with $xp = xp''w$. Thus $xp \equiv xp'w(\pi_t)$ and the length of $xp'w$ is less than m. The statement concerning π can be proved similarly. But the number of elements of $X^{(m)}$ is finite. Therefore, among the m-congruent partitions π_t $(t = 1, \dots, s, \dots)$ we have only finitely many different. Let us denote them by $\pi_1^{(m)}, \ldots, \pi_r^{(m)}$. Thus we got that the number of all classes of $x_1^{(m)*} \cap \cdots \cap x_r^{(m)*}$ is less than or equal to m^r. Since the restrictions of $x_1^{(m)*}, \ldots, x_r^{(m)*}$ to $X^{(m^r)}$ are the same as the extensions $\pi_1^{(m^r)}, \ldots, \pi_r^{(m^r)}$ of $\pi_1^{(m)}, \ldots, \pi_r^{(m)}$ to $X^{(m^r)}$, we got that $\pi_1^{(m^r)} \cap \cdots \cap x_r^{(m^r)}$ is an m^r-congruent partition. It is obvious that $\pi_1^{(m^r)} \cap \cdots \cap \pi_r^{(m^r)} \subseteqq \pi^{(m^r)}$, $\mathfrak{X}/\pi^{(m)*} \cong \mathfrak{A}$, and $\mathfrak{X}/\pi_h^{(m)*}$ $(h=1, \ldots, r)$ is isomorphic to a unoid in $S(\langle \mathfrak{A}_i | i=1, \ldots, k \rangle)$.

Conversely, let us suppose that the conditions of our theorem are satisfied. Then, as is well known, $\mathfrak{A}(\cong \mathfrak{X}/\pi^{(m)*})$ is a homomorphic image of a subdirect product of $\mathfrak{X}/\pi_1^{(m)*}, \ldots, \mathfrak{X}/\pi_r^{(m)*}$ because $\pi_1^{(m)*} \cap \cdots \cap \pi_r^{(m)*} \subseteqq \pi^{(m)*}$. Therefore, $\mathfrak{A}$ is contained in the equational class generated by $\langle \mathfrak{A}_i | i=1, \ldots, k \rangle$. This ends the proof of the theorem.

We remark that the algorithm given by the theorem above can easily be generalized for equational classes generated by finitely many finite universal algebras of finite type.

References

[1] F. GÉCSEG and I. PEÁK, *Algebraic Theory of Automata* (Budapest, 1972).
[2] А. И. Мальцев, *Алгебраические системы* (Москва, 1970).

(Received September 1, 1972)

Общие теоремы о факторизации оператор-функций относительно контура. I. Голоморфные функции

И. Ц. ГОХБЕРГ и Ю. ЛАЙТЕРЕР (Кишинев, СССР)

Посвящается академику Б. С.-Надь к его шестидесятилетию

Пусть Γ — линия в комплексной полоскости, состоящая из конечного числа непересекающихся замкнутых спрямляемых жордановых кривых. Предполагается, что контур Γ разбивает расширенную плоскость на два открытых множества F^+ и F^-, причем каждая точка Γ является граничной как для F^+, так и для F^-. Кроме того, предполагается, что $\infty \in F^-$ и $0 \in F^+$.

Обозначим через $L(\mathfrak{B})$ алгебру всех линейных ограниченных операторов, действующих в банаховом пространстве $\mathfrak{B}$, и через $GL(\mathfrak{B})$ группу обратимых операторов из $L(\mathfrak{B})$.

Пусть $A: \Gamma \to L(\mathfrak{B})$ — непрерывная оператор-функция. Следуя [1], представление оператор-функции A в виде

$$A = A_- D A_+$$

назовем *факторизацией A относительно контура Γ*, если множители обладают следующими свойствами:

1. оператор-функция $D: \Gamma \to L(\mathfrak{B})$ имеет вид

$$D(\zeta) = \sum_{j=1}^{n} \zeta^{\varkappa_j} P_j + P_0,$$

где операторы P_j $(j=1, 2, \ldots, n)$ являются дизъюнктными одномерными проекторами из $L(\mathfrak{B})$, $P_0 = I - P_1 - P_2 - \cdots - P_n$ и $\varkappa_1 \geqq \varkappa_2 \geqq \cdots \geqq \varkappa_n$ — некоторые целые числа, отличные от нуля;

2. оператор-функции $A_-, A_+: \Gamma \to L(\mathfrak{B})$ допускают продолжения, голоморфные внутри и непрерывные, включая границу, соответственно в $F^+ \cup \Gamma$ и $F^- \cup \Gamma$, причем все значения оператор-функций A_-, A_+ и их продолжений обратимы.

Если оператор-функция $A: \Gamma \to L(\mathfrak{B})$ допускает факторизацию относительно контура Γ, то, как показано в [1], целые числа $\varkappa_1 \geqq \varkappa_2 \geqq \cdots \geqq \varkappa_n$ однозначно определяются оператор-функцией A. Эти числа называются *частными индексами* оператор-функции A, а число $\operatorname{Ind} A$, опеределяемое равенством

$$\operatorname{Ind} A = \sum_{j=1}^{n} \varkappa_j$$

называется *суммарным индексом*.

В настоящей статье устанавливаются критерии возможности факторизации относительно контура для оператор-функций из различных банаховых алгебр.

Рассмотрим сначала случай, когда пространство $\mathfrak{B}$ является гильбертовым пространством ($\mathfrak{B}=\mathfrak{H}$).

Обозначим через $\mathbf{L}_2=\mathbf{L}_2(\Gamma, \mathfrak{H})$ гильбертово пространство сильно измеримых функций $f:\Gamma\to\mathfrak{H}$ со скалярным произведением

$$(f, g)_{\mathbf{L}_2} = \int_{\Gamma} (f(\zeta), g(\zeta))_{\mathfrak{H}} |d\zeta| \qquad (f, g\in \mathbf{L}_2(\Gamma, \mathfrak{H}))$$

Оператор $\mathbf{P}$, определенный в $\mathbf{L}_2(\Gamma, \mathfrak{H})$ равенством

$$(\mathbf{P}\varphi)(\zeta) = \frac{1}{2}\varphi(\zeta)+\frac{1}{2\pi i}\int_{\Gamma}\frac{\varphi(z)}{z-\zeta}dz \qquad (\varphi\in \mathbf{L}_2(\Gamma, \mathfrak{H})). \tag{0.1}$$

является ограниченным проектором*) и вектор-функции из его множества значений $\mathbf{L}_2^+=\operatorname{Im}\mathbf{P}$ допускают голоморфные продолжения в F^+.

Пусть $A:\Gamma\to L(\mathfrak{H})$ — непрерывная оператор-функция. Условимся через $\mathbf{A}$ обозначать линейный ограниченный оператор, действующий в пространстве $\mathbf{L}_2(\Gamma, \mathfrak{H})$ по правилу $(\mathbf{A}f)(\zeta)=A(\zeta)f(\zeta)\,(f\in\mathbf{L}_2, \zeta\in\Gamma)$.

Приведем для примера четыре следствия, вытекающие из основных теорем статьи.

Теорема 0.1. *Пусть оператор-функция $A:\Gamma\to L(\mathfrak{H})$ удовлетворяет условию Гельдера с показателем* α $(0<\alpha<1)$, *т. е.*

$$\sup_{\zeta_1,\zeta_2\in\Gamma;\,\zeta_1\neq\zeta_2} (\|A(\zeta_1)-A(\zeta_2)\|/|\zeta_1-\zeta_2|^{\alpha})<\infty.$$

Для того чтобы оператор-функция A допускала факторизацию относительно контура Γ, необходимо и достаточно, чтобы оператор **PA** *был Ф-оператором*) в пространстве* $\mathbf{L}_2^+$.

Если оператор-функция A допускает факторизацию $A=A_-DA_+$, то множители $A_\pm$ также удовлетворяют условию Гельдера на Γ с показателем α.

---

*) Во введении предполагается, что контур Γ является достаточно гладким.

*) Оператор B насывается Φ-оператором [2], если его множество значений замкнуто и числа $\dim\operatorname{Ker} B$ и $\dim\operatorname{Coker} B$ конечны.

В случае факторизации относительно единичной окружности $\Gamma_0 = \{\zeta : |\zeta| = 1\}$ имеет место следующее предложение.

Теорема 0. 2. *Пусть оператор-функция* $A:\Gamma_0 \to L(\mathfrak{H})$ *разлагается в абсолютно сходящийся ряд Фурье*

$$A(\zeta) = \sum_{j=-\infty}^{\infty} \zeta^j A_j, \qquad \sum_{j=-\infty}^{\infty} \|A_j\| < \infty,$$

где $A_j \in L(\mathfrak{H})$.

Для того чтобы оператор-функция A допускала факторизацию относительно окружности Γ_0, *необходимо и достаточно, чтобы оператор* **PA** *был Ф-оператором в пространстве* $\mathbf{L}_2^+$.

Если оператор-функция A допускает факторизацию $A = A_- D A_+$, *то множители* A_-, A_+ *также разлагаются в абсолютно сходящиеся ряды соответственно по неположительным или неотрицательным степеням* ζ.

Приведем еще две теоремы о факторизации в случае произвольного банахова пространства $\mathfrak{B}$. Обозначим через $H_\alpha(\Gamma, \mathfrak{B})$ $(0<\alpha<1)$ банахово пространство вектор-функций $f:\Gamma \to \mathfrak{B}$ удовлетворяющих условию Гельдера с показателем α и нормой

$$\|f\|_\alpha = \max_{\zeta \in \Gamma} \|f(\zeta)\|_{\mathfrak{B}} + \sup_{\zeta_1, \zeta_2 \in \Gamma; \zeta_1 \neq \zeta_2} \|f(\zeta_1) - f(\zeta_2)\| / |\zeta_1 - \zeta_2|^\alpha$$

Оператор **P**, определенный формулой (0. 1), является линейным ограниченным проектром в пространстве $H_\alpha(\Gamma, \mathfrak{B})$. Обозначим через $H_\alpha^+(\Gamma, \mathfrak{B})$ множество значений оператора **P**.

Теорема 0. 3. *Для того чтобы оператор-функция* $A:\Gamma \to GL(\mathfrak{B})$ *из* $H_\alpha(\Gamma, L(\mathfrak{B}))$ $(0<\alpha<1)$ *допускала факторизауию относительно* Γ, *необходимо и достаточно, чтобы оператор* **PA** *был Ф-оператором в пространстве* $H_\alpha^+(\Gamma, \mathfrak{B})$.

Если A допускает факторизацию $A = A_- D A_+$, *то* $A_\pm \in H_\alpha(\Gamma, \mathfrak{B})$.

Рассмотрим пространство $\mathfrak{W}(\mathfrak{B})$ всех вектор-функций $f:\Gamma_0 \to \mathfrak{B}$ разлагающихся в абсолютно сходящиеся ряды Фурье

$$f(\zeta) = \sum_{j=-\infty}^{\infty} \zeta^j f_j \qquad (f_j \in \mathfrak{B}, |\zeta| = 1)$$

с нормой

$$\|f\|_{\mathfrak{W}} = \sum_{j=-\infty}^{\infty} \|f_j\|_{\mathfrak{B}}$$

Легко видеть, что для проектора **P**, определенного равенством (0. 1), в этом пространстве имеет место равенство

$$(\mathbf{P}f)(\zeta) = \sum_{j=0}^{\infty} \zeta^j f_j.$$

Обозначим через $\mathfrak{W}^+(\mathfrak{B})$ множетство всех значений проектора $\mathbf{P}$.

Теорема 0. 4. *Для того чтобы оператор-функция* $A: \Gamma_0 \to GL(\mathfrak{B})$ *из* $\mathfrak{W}(L(\mathfrak{B}))$ *допускала факторизацию относительно единичной окружности* Γ_0, *необходимо и достаточно, чтобы оператор* $\mathbf{PA}$ *был Φ-оператором в пространстве* $\mathfrak{W}^+(\mathfrak{B})$.

Если A допускает факторизацию $A = A_- D A_+$, *то* $A_\pm \in \mathfrak{W}(L(\mathfrak{B}))$.

Отметим, что при условиях всех четырех теорем, если оператор-функция A допускает факторизацию и $\varkappa_j$ $(j=1, 2, \ldots, n)$ — ее частные индексы, то

$$\dim \operatorname{Ker} \mathbf{PA} = -\sum_{\varkappa_j<0} \varkappa_j, \qquad \dim \operatorname{Coker} \mathbf{PA} = \sum_{\varkappa_j>0} \varkappa_j$$

и, следовательно,

$$\operatorname{Ind}(A, \Gamma) = \operatorname{Ind} \mathbf{PA}.$$

Теоремы 0. 1 и 0. 3 обобщают на бесконечномерный случай известную теорему И. Племели, Н. И. Мусхелишвили и Н. П. Векуа [3, 4] о том, что всякая неособенная матрица-функция с элементами, удовлетворяющими условию Гельдера с показателями α $(0<\alpha<1)$, допускает факторизацию. Отметим, что в случае конечномерного пространства $\mathfrak{H}$ или $\mathfrak{B}$ оператор $\mathbf{PA}$ является Φ-оператором в H_α^+ и $\mathbf{L}_2^+$ для любой неособенной матрицы-функции A.

Теоремы 0. 2 и 0. 4 являются обобщениями теоремы М. Г. Крейна и одного из авторов [5] о факторизации неособенных матриц-функций с элементами, разлагающимися в абсолютно сходящиеся ряды Фурье. Для таких матриц-функций оператор $\mathbf{PA}$ также всегда является Φ-оператором в пространствах $\mathbf{L}_2^+$ и $\mathfrak{W}^+$.

Статья состоит из двух частей. В первой части излагаются результаты, относящиеся к голоморфным оператор-функциям, а во второй — различные обобщения*). Первая часть является основной в идейном отношении. Она состоит из трех параграфов. Первый носит вспомогательный характер. В нем, в частности, формулируются теоремы из статей [6—8], которые играют важную роль в дальнейших доказательствах. Во втором параграфе доказывается основная теорема о факторизации голоморфных оператор-функций. Эта теорема неоднократно используется в доказательствах общих теорем из второй части статьи.

В последнем параграфе исследуются различные возможные обобщения задачи факторизации оператор-функций относительно контура. В этом же параграфе подробно разбирается один поучительный пример.

*) Отметим, что все теоремы, сформулированные во введении, во всей общности доказываются лишь во второй части.

§ 1. Определения и вспомогательные предложения

В этом параграфе формулируются вспомогательные предложения и теоремы, играющие важную роль в доказательстве основных теорем. Начнем с некоторых определений.

1. Пусть по-прежнему Γ — линия в комплексной плоскости, состоящая из конечного числа непересекающихся замкнутых спрямляемых жордановых кривых. Будем предполагать, что контур Γ разбивает расширенную комплексную плоскость на два открытых множества F^+ и F^-, причем каждая точка Γ является граничной для F^+ и F^-. Предположим, что $0\in F^+$ и $\infty\in F^-$. Открытую окрестность K контура Γ назовем *Γ-кольцом*, если 1) ее замыкание $\overline{K}$ состоит из того же числа связных компонент, что и Γ, причем каждая из этих компонент содержит точно одну компоненту контура Γ и гомеоморфна круговому кольцу; 2) граница ∂K состоит из конечного числа замкнутых спрямляемых жордановых кривых; 3) $0\notin\overline{K}$, $\infty\notin\overline{K}$.

Пусть $\mathfrak{B}$ — банахово пространство, $L(\mathfrak{B})$ — алгебра всех линейных ограниченных операторов в $\mathfrak{B}$, а $GL(\mathfrak{B})$ — группа обратимых операторов из $L(\mathfrak{B})$.

Пусть K является Γ-кольцом. Обозначим через $C_\omega(K,\mathfrak{B})$ банахово пространство голоморфных в K и непрерывных на $\overline{K}$ вектор-функций $f:\overline{K}\to\mathfrak{B}$ с нормой

$$\|f\|_{C_\omega(K,\mathfrak{B})}\overset{\text{def}}{=}\max_{\zeta\in K}\|f(\zeta)\|_{\mathfrak{B}}.$$

Через $C_\omega^+(K,\mathfrak{B})$, обозначим подпространство функций из $C_\omega(K,\mathfrak{B})$, допускающих голоморфное продолжение в F^+, а через $C_\omega^-(K,\mathfrak{B})$ — подпространство функций из $C_\omega(K,\mathfrak{B})$, допускающих голоморфные продолжения в F^-, которые обращаются нуль на бесконечности. Легко видеть, что пространство $C_\omega(K,\mathfrak{B})$ распадается в прямую сумму его подпространств $C_\omega^+(K,\mathfrak{B})$ и $C_\omega^-(K,\mathfrak{B})$: $C_\omega(K,\mathfrak{B})=C_\omega^+(K,\mathfrak{B})\dotplus C_\omega^-(K,\mathfrak{B})$. Обозначим через $\mathbf{P}$ проектор, проектирующий пространство $C_\omega(K,\mathfrak{B})$ на $C_\omega^+(K,\mathfrak{B})$ на $C_\omega^+(K,\mathfrak{B})$ параллельно $C_\omega^-(K,\mathfrak{B})$, а через $\mathbf{Q}$ — дополнительный проектор $\mathbf{Q}=\mathbf{I}-\mathbf{P}$. Легко видеть, что проектор $\mathbf{P}$ выражается по формуле (0. 1).

Каждой оператор-функции $A:\overline{K}\to L(\mathfrak{B})$, непрерывной на $\overline{K}$ и голоморфной в K, сопоставим оператор $\mathbf{A}$, действующий в $C_\omega(K,\mathfrak{B})$ по формуле

$$(\mathbf{A}f)(\zeta)=A(\zeta)f(\zeta).$$

Операторы вида $\mathbf{PA}$ и $\mathbf{QA}$ будем рассматривать соответственно в пространствах $C_\omega^+(K,\mathfrak{B})$ и $C_\omega^-(K,\mathfrak{B})$.

Пусть оператор-функция $A:\overline{K}\to GL(\mathfrak{B})$ принадлежит $C_\omega(K,L(\mathfrak{B}))$. Если A допускает факторизацию относительно контура $\Gamma: A=A_-DA_+$, то легко видеть, что $A_+\in C_\omega^+(K,L(\mathfrak{B}))$ и $A_- - A_-(\infty)\in C_\omega^-(K,(\mathfrak{B}))$.

Факторизация относительно контура Γ, в которой средний множитель тождественно равен единице, называется канонической.

2. В следующем параграфе существенно используется

Теорема 1. 1. *Оператор-функция* $A:\bar{K}\to GL(\mathfrak{B})$, *принадлежащая* $C_\omega(K, L(\mathfrak{B}))$, *допускает каноническую факторизацию относительно контура* Γ *в том и только том случае, когда оператор* **PA** *является обратимым в пространстве* $C_\omega^+(K, \mathfrak{B})$.

Эта теорема доказана в статье [8].

Оператор-функция A со значениями из $L(\mathfrak{B})$ называется *конечномероморфной в точке* ζ_0, если она либо голомрфна в ζ_0, либо имеет полюс в точке ζ_0 и в разложении

$$A(\zeta)=\sum_{j=-n}^{\infty}(\zeta-\zeta_0)^j A_j \qquad (A_j\in L(\mathfrak{B}))$$

все операторы A_j $(j=-n, \ldots, -1)$ конечномерны. Оператор-функция A называется *нормальной в точке* ζ_0, если она конечномероморфна в точке ζ_0, A_0 является Φ-оператором и $A(\zeta)\in GL(\mathfrak{B})$ для всех ζ из некоторого проколотого круга $0<|\zeta-\zeta_0|<\varepsilon$. Назовем оператор-функцию A *нормальной в точке* ∞, если оператор-функция $A(\zeta^{-1})$ нормальна в точке 0.

Как и в статье [6], условимся говорить, что непрерывная оператор-функция $A:\Gamma\to GL(\mathfrak{B})$ *вполне нормальна в* $F^+(F^-)$, если она допускает продолжение в $\Gamma\cup F^+(\Gamma\cup{}^-)$, непрерывное на Γ и нормальное во всех точках $F^+(F^-)$.

В доказательствах основных теорем существенную роль играет следующая теорема, доказанная в [6].

Теорема 1. 2. *Если оператор-функция* $A:\Gamma\to GL(\mathfrak{B})$ *допускает представление*

$$A=XY,$$

где X — *вполне нормальная оператор-функция в* F^+, *а* Y — *вполне нормальная оператор-функция в* F^-, *то она допускает факторизацию относительно контура* Γ.

Нам понадобится еще одно предложение, установленное в [9] (см. также [10], лемма 2. 1).

Предложение 1. 1. *Если оператор-функция* $A_\pm$ — *вполне нормальна в* $F^\pm$, *то оператор-функция* $A_\pm^{-1}$ *также вполне нормальна в* $F^\pm$.

3. Пусть K является Γ-кольцом. Обозначим через P проектор, проектирующий $C_\omega(K, L(\mathfrak{B}))$ на $C_\omega^+(K, L(\mathfrak{B}))$ параллельно $C_\omega^-(K, L(\mathfrak{B}))$, а через **Q** — дополнительный проектор $\mathbf{Q}=I-\mathbf{P}$.

Из общей теоремы (см. [11]. гл, , лемма 5. 1.), о факторизации элементов, близких к единичному, в абстрактных банаховых алгебрах вытекает следующая лемма.

Лемма 1. 1. *Любая оператор-функция* $A \in C_\omega(K, L(\mathfrak{B}))$, *удовлетворяющая условию*

$$\|A(\zeta) - I\|_{\mathfrak{B}} < \delta_K \overset{\text{def}}{=} \min\{\|\mathbf{P}\|^{-1}, \|\mathbf{Q}\|^{-1}\} \qquad (\zeta \in \overline{K})$$

допускает каноническую факторизацию относительно Γ.

В дальнейшем используется также следующая аппроксимационная лемма.

Лемма 1. 2. *Пусть* $\mathfrak{A}$ — *одно из пространств* $\mathfrak{B}$ *или* $L(\mathfrak{B})$. *Функции вида*

$$\sum_{j=1}^{n} r_j(\zeta)\, a_j \qquad (a_j \in \mathfrak{A}), \tag{1.1}$$

где r_j — *рациональные функции с полюсами вне* $\overline{K}$, *образуют плотное множество в* $C_\omega(K, \mathfrak{A})$.

В случае $\mathfrak{A} = L(\mathfrak{B})$ эта лемма по существу доказана в [8] (см. доказательство леммы 1. 1 главы 1). Это доказательство остается в силе также в случае $\mathfrak{A} = \mathfrak{B}$.

Лемма 1. 3. *Пусть* Ω *некоторый компакт конечной комплексной плоскости и* $f\colon \Omega \to \mathfrak{B}$ — *голоморфная на* Ω *вектор-функция со свойством* $f(\zeta) \neq 0$ *для всех* $\zeta \in \Omega$.

Тогда существует голоморфная оператор-функция $N\colon \Omega \to GL(\mathfrak{B})$, *такая, что* $N(\zeta)f(\zeta) \equiv x (\zeta \in \Omega)$, *где* x — *некоторый фиксированный вектор из* $\mathfrak{B}$.

Эта лемма доказана в [7] (см. теорему 2. 3). Она также легко выводится из некоторых теорем о голоморфных расслоениях Х. Рёрля [12] и Х. Грауэрта [13] (см. также [14]).

§ 2. Основная теорема о голоморфных оператор-функциях

1. В дальнейшем мы придерживаемся обозначений, введенных в первом параграфе. В частности, через K обозначается Γ — кольцо и через $\mathfrak{B}$ — банахово пространство. Основной в этом параграфе является следующая теорема.

Теорема 2. 1. *Пусть оператор-функция* $A\colon \overline{K} \to GL(\mathfrak{B})$ *голоморфна в* K *и непрерывна на* $\overline{K}$. *Для того чтобы оператор-функция* A *допускала факторизацию относительно* Γ, *необходимо и достаточно, чтобы оператор* **PA** *был* Φ-*оператором в пространстве* $C^+_\omega(K, \mathfrak{B})$.

Доказательство этой теоремы основывается на следующих двух леммах.

Лемма 2.1. *Пусть оператор-функция* $A:\overline{K}\to GL(\mathfrak{B})$ *голоморфна и пусть оператор* **PA** *является Φ-оператором.*

Если $\dim\operatorname{Coker}\mathbf{PA}>0$ *), *то существует голоморфная в* $\overline{K}$ *и вполне нормальная в* F^+ *оператор-функция* B_+, *принимающая всюду в* $\overline{K}$ *обратимые значения, такая, что оператор* $\mathbf{PAB}_+$ *является Φ-оператором в* $C_\omega^+(K, \mathfrak{B})$ *и*

$$\dim\operatorname{Coker}\mathbf{PAB}_+ < \dim\operatorname{Coker}\mathbf{PA}. \tag{2.1}$$

Лемма 2.2. *Пусть оператор-функция* $A:\overline{K}\to GL(\mathfrak{B})$ *голоморфна, и пусть оператор* **PA** *является Φ-оператором. Если* $\operatorname{Im}\mathbf{PA}=C_\omega^+(K,\mathfrak{B})$ *и* $\dim\operatorname{Ker}\mathbf{PA}>0$, *то существует голоморфная в* $\overline{K}$ *и вполне нормальная в* F^- *оператор-функция* B_-, *принимаящая всюду в* $\overline{K}$ *обратимые значения, такая, что*

$$\operatorname{Im}\mathbf{PB}_-\mathbf{A} = C_\omega^+(K,\mathfrak{B}) \tag{2.2}$$

и

$$\dim\operatorname{Ker}\mathbf{PB}_-\mathbf{A} < \dim\operatorname{Ker}\mathbf{PA}. \tag{2.3}$$

2. Доказательство леммы 2.1. Покажем сначала, что существует вектор $x\neq 0$ из $\mathfrak{B}$, такой, что $x\notin\operatorname{Im}\mathbf{PA}$. **)

Допустим противное, т. е. что

$$\mathfrak{B}\subseteqq\operatorname{Im}\mathbf{PA}. \tag{2.4}$$

Тогда для всех $z\in\mathfrak{B}$ и $k=0, 1, 2, \dots$ будем иметь

$$\zeta^k z\in\operatorname{Im}\mathbf{PA}. \tag{2.5}$$

В самом деле, пусть для некоторого целого неотрицательного k функция $\zeta^k z\in\operatorname{Im}\mathbf{PA}$, т. е. существуют вектор-функции $f_+\in C_\omega^+(K,\mathfrak{B})$ и $f_-\in(K,\mathfrak{B})$, удовлетворяющие равенству

$$\zeta^k z+f_-(\zeta) = A(\zeta)f_+(\zeta) \qquad (\zeta\in\overline{K}).$$

Тогда

$$\zeta^{k+1}z+\mathbf{P}[\zeta f_-(\zeta)] = \mathbf{P}[A(\zeta)\zeta f_+(\zeta)]\in\operatorname{Im}\mathbf{PA},$$

Так как

$$\mathbf{P}[\zeta f_-(\zeta)] = \zeta f_-(\zeta)_{\zeta=\infty}\in\operatorname{Im}\mathbf{PA},$$

то $\zeta^{k+1}z\in\operatorname{Im}\ \mathbf{PA}$.

Покажем теперь, что

$$\frac{z}{(\zeta-\alpha)^k}\in\operatorname{Im}\mathbf{PA} \qquad (k=1, 2, \dots) \tag{2.6}$$

---

*) Напомним, что $\operatorname{Coker}\mathbf{PA}\overset{\text{def}}{=}C_\omega^+(K,\mathfrak{B})/\operatorname{Im}\mathbf{PA}$.

**) Здесь и в дальнейшем под x понимаем вектор-функцию на $\overline{K}$ равную тождественно фиксированному вектору x.

для всех $z\in\mathfrak{B}$ и $\alpha\in F^{-}\setminus\overline{K}$ $(\neq 0)$. В силу соотношения $\dim\operatorname{Coker}\mathbf{PA}<\infty$ существуют числа $\beta_1, \ldots, \beta_n$ $(\beta_n\neq 0)$, такие, что

$$\sum_{j=1}^{n}\beta_j\frac{z}{(\zeta-\alpha)^{k_j}}\in\operatorname{Im}\mathbf{PA}.$$

Последнее означает, что

$$f_-(\zeta)+\sum_{j=1}^{n}\beta_j\frac{z}{(\zeta-\alpha)^{k_j}}=A(\zeta)f_+(\zeta),$$

где $f_\pm\in C_\omega^\pm(K,\mathfrak{B})$. Следовательно,

$$\mathbf{P}[(\zeta-\alpha)^{k(n-1)}f_-(\zeta)]+\sum_{j=1}^{n-1}\beta^j(\zeta-\alpha)^{k(n-1-j)}z+\beta_n\frac{z}{(\zeta-\alpha)^k}=$$
$$=\mathbf{P}[A(\zeta)(\zeta-\alpha)^{k(n-1)}f_+(\zeta)]\in\operatorname{Im}\mathbf{PA}.$$

Так как $\beta_n\neq 0$, то отсюда в силу (2. 5) вытекает соотношение (2. 6).

Из (2. 5) и (2. 6) следует, что все функции (1. 1) принадлежат пространству $\operatorname{Im}\mathbf{PA}$. В силу леммы 1. 2 это означает, что $\operatorname{Im}\mathbf{PA}=C_\omega^+(K,\mathfrak{B})$. Последнее противоречит условию $\dim\operatorname{Coker}\mathbf{PA}>0$.

Выберем вектор x из $\mathfrak{B}$ так, чтобы он не принадлежал множеству значений оператора $\mathbf{PA}$. Так как $\dim\operatorname{Coker}\mathbf{PA}<\infty$, то можно подобрать скалярный многочлен

$$\varphi_+(\zeta)=\sum_{j=1}^{n}\zeta^j\alpha_j$$

с $\alpha_n\neq 0$ так, чтобы вектор-функция $\varphi_+(\zeta)$ $(\zeta\in\overline{K})$ принадлежала пространству $\mathbf{PA}$. Это означает, что имеет место равенство

$$A(\zeta)f_+(\zeta)=\varphi_+(\zeta)x+f_-(\zeta), \tag{2.7}$$

где $f_\pm\in C_\omega^\pm(K,\mathfrak{B})$.

Так как оператор-функция A^{-1} голоморфна на замкнутом кольце $\overline{K}$, то из равенства (2. 7) следует, что вектор-функция f_+ голоморфна на замкнутом множестве $F^+\cup\overline{K}$. Пусть $\zeta_1, \ldots, \zeta_k$ — все нули вектор-функции f_+ из множества $F^+\cup\overline{K}$. Положим

$$\tilde{f}_+(\zeta)=\sum_{j=1}^{k}(\zeta-\zeta_j)^{-1}f_+(\zeta). \tag{2. 8}$$

Вектор-функция $\tilde{f}_+$ голоморфна на $F^+\cup\overline{K}$ и $\tilde{f}_+(\zeta)\neq 0$ $(\zeta\in F^+\cup K)$. Следовательно, в силу леммы 1. 3 существует голоморфная операторфункция $N_+:F^+\cup\overline{K}\to GL(\mathfrak{B})$, такая, что

$$N_+^{-1}(\zeta)\tilde{f}_+(\zeta)\equiv y\quad(\zeta\in F^+\cup\overline{K}), \tag{2. 9}$$

где y — некоторый вектор из $\mathfrak{B}$.

Пусть R — проектор, проектирующий пространство $\mathfrak{B}$ на одномерное пространство, порожденное вектором y. Рассмотрим оператор-функцию $R_+(\zeta) = I - R + \zeta^{-n}R$. Так как оператор-функция $R_+^{-1}(\zeta) = I + \zeta^n R$ принадлежит $C_\omega^+(K, L(\mathfrak{B}))$, то оператор $\mathbf{PR}_+^{-1}$ является обратным справа к оператору $\mathbf{PR}_+$. С другой стороны, для любого вектора f из $C_\omega^+(K, \mathfrak{B})$ имеет место равенство

$$(\mathbf{PR}_+^{-1}\mathbf{PR}_+f)(\zeta) = f(\zeta) + \zeta^n \mathbf{P}(\zeta^{-n}Rf(\zeta)) - Rf(\zeta).$$

Легко видеть, что

$$\zeta^n \mathbf{P}(\zeta^{+n}Rf(\zeta) - Rf(\zeta)) = \sum_{j=0}^{n-1} \zeta^j Rf_j,$$

где f_j — коэффициенты Тейлора функции $f(\zeta)$ в точке $\zeta = 0$. Из последнего равенства вытекает, что $\dim \operatorname{Im}(\mathbf{PR}_+^{-1}\mathbf{PR}_+ - \mathbf{P}) = n$.

Таким образом, оператор $\mathbf{PR}_+$ является Φ-оператором.

Положим $B_+(\zeta) = N_+(\zeta)R_+(\zeta)$. Легко видеть, что оператор-функция B_+ голоморфна в $\bar{K}$ и вполне нормальна в F^+. Все значения $B_+(\zeta)$ при $\zeta \in (\bar{K} \cup F^+) \setminus \{0\}$ являются обратимыми операторами. Покажем, что оператор $\mathbf{PAB}_+$ является Φ-оператором. Без труда проверяются равенства $\mathbf{PAN}_+ = (\mathbf{PA})(\mathbf{PN}_+)$ и $(\mathbf{PN}_+^{-1})(\mathbf{PN}_+) = (\mathbf{PN}_+)(\mathbf{PN}_+^{-1}) = \mathbf{P}$. Следовательно, оператор $\mathbf{PAN}_+$ является Φ-оператором и $\operatorname{Im}\mathbf{PAN}_+ = \operatorname{Im}\mathbf{PA}$.

Оператор $(\mathbf{I} - \mathbf{P})\mathbf{R}_+\mathbf{P}$ конечномерен, так как

$$((\mathbf{I} - \mathbf{P})\mathbf{R}_+f)(\zeta) = (\mathbf{I} - \mathbf{P})(\zeta^{-n}Rf(\zeta)) = \sum_{j=0}^{n} \zeta^{-j}Rf_j,$$

где f_j — коэффициенты Тейлора функции f в точке $\zeta = 0$.

Из равенства

$$\mathbf{PAB}_+ = \mathbf{P}(\mathbf{AN}_+)\mathbf{PR}_+ + \mathbf{PAN}_+(\mathbf{I} - \mathbf{P})\mathbf{R}_+$$

в силу доказанного вытекает, что $\mathbf{PAB}_+$ является Φ — оператором. Перейдем к доказательству неравенства (2. 1).

Пусть g_+ — любая функция из $\operatorname{Im}\mathbf{PA}$ $(= \operatorname{Im}\mathbf{PAN}_+)$. Тогда для некоторого вектора $h_+ \in C_\omega^+(K, \mathfrak{B})$ будет иметь место равенство $g_+ = \mathbf{PAN}_+h_+$.

Полагая

$$\tilde{h}^+(\zeta) = (I - R + \zeta^n R)\, h_+(\zeta)$$

получим

$$\mathbf{PAB}_+\tilde{h}_+ = \mathbf{PAN}_+h_+ = g_+,$$

т. е. $g_+ \in \operatorname{Im} PAB_+$. Таким образом,

$$\operatorname{Im}\mathbf{PA} \subseteq \operatorname{Im}\mathbf{PAB}_+ \tag{2.10}$$

Рассмотрим функцию

$$v_+(\zeta) = \sum_{j=1}^{k} (\zeta - \zeta_j)y.$$

Из равенств (2. 9) и (2. 8) вытекает, что

$$B_+(\zeta)v_+(\zeta) = \zeta^{-n} \sum_{j=1}^{k} (\zeta-\zeta_j) N_+(\zeta) y = \zeta^{-n} f_+(\zeta).$$

Следовательно, в силу (2. 7) будем иметь

$$\mathbf{P}[A(\zeta) B_+(\zeta) v_+(\zeta)] = \mathbf{P}[\zeta^{-n} \varphi_+(\zeta) x + \zeta^{-n} f_-(\zeta)] = \alpha_n x,$$

где комплексное число $\alpha_n \neq 0$. Стало быть,

(2. 11) $$x \in \operatorname{Im} \mathbf{PAB}_+ .$$

Так как вектор x не принадлежит $\operatorname{Im} \mathbf{PA}$, то из соотношений (2. 10) и (2. 11) непосредственно вытекает неравенство (2. 1).

Лемма доказана.

3. Доказательство леммы 2. 2. Пусть f_+ $(\neq 0)$ — некоторая вектор-функция из $\operatorname{Ker} \mathbf{PA}$. Тогда

(2. 12) $$A(\xi) f_+(\zeta) = f_-(\zeta) \qquad (\zeta \in \overline{K}),$$

где $f_- \in C_\omega^-(K, \mathfrak{B})$. Так как оператор-функция $A: \overline{K} \to GL(\mathfrak{B})$ голоморфна на $\overline{K}$, то из (2. 12) следует, что вектор-функция f_- допускает голоморфное продолжение на $F^- \cup \overline{K}$. Пусть $\zeta_1, \ldots, \zeta_m$ — все числа из множества $(F^- \cup \overline{K}) \diagdown \{\infty\}$, являющиеся нулями функции f_-, и s — порядок нуля функции f_- на бесконечности. Положим

$$\tilde{f}_-(\zeta) = \zeta^s \sum_{j=1}^{m} (\zeta-\zeta_j)^{-1} f_-(\zeta).$$

Вектор-функция $\tilde{f}_-$ голоморфна на $F^- \cup \overline{K}$ и $\tilde{f}(\zeta) \neq 0$ $(\zeta \in F^- \cup \overline{K})$. Следовательно, в силу леммы 1. 3 существует голоморфная оператор-функция $N_-: F^- \cup \overline{K} \to GL(\mathfrak{B})$, такая, что

(2. 13) $$N_-(\zeta) \tilde{f}_-(\zeta) \equiv y \qquad (\zeta \in F^- \cup \overline{K}),$$

где $y \neq 0$ — некоторый постоянный вектор из $\mathfrak{B}$.

Положим

$$\tilde{f}_+(\zeta) = \sum_{j=1}^{m} (\zeta-\zeta_j)^{-1} f_+(\zeta).$$

Из равенства (2. 12) следует, что вектор-функция f_+ голоморфна на $\overline{K} \cup F^+$ и что f_+ и f_- на множестве $\overline{K}$ имеют один и те же нули. Следовательно, вектор-функция $\tilde{f}_+$ голоморфна на $F^+ \cup \overline{K}$. Из равенства (2. 12) также следует, что

(2.14) $$A(\zeta) \tilde{f}_+(\zeta) = \zeta^{-s} \tilde{f}_-(\zeta).$$

Пусть R — проектор, проектирующий пространство $\mathfrak{B}$ на одномерное пространство, порожденное вектором y. Положим

$$B_-(\zeta) = R_-(\zeta) N_-(\zeta),$$

где $R_-(\zeta) = I - R + \zeta^s R$, и покажем, что $\mathbf{PB_-A}$ является Φ-оператором в пространстве $C_\omega^+(K, \mathfrak{B})$. Это утверждение будем доказывать так же, как подобное утверждение в лемме 2. 1. В процессе доказательства леммы 2. 1, в частности, показано, что оператор $\mathbf{PR_-}$ является Φ-оператором.

Без труда проверяются равенства $\mathbf{PN_-A} = (\mathbf{PN_-})(\mathbf{PA})$ и $(\mathbf{PN_-^{-1}})(\mathbf{PN_-}) = (\mathbf{PN_-})(\mathbf{PN_-^{-1}}) = \mathbf{P}$. Следовательно, оператор $\mathbf{PN_-A}$ является Φ-оператором и $\operatorname{Ker} \mathbf{PN_-A} = \operatorname{Ker} \mathbf{PA}$.

Оператор $\mathbf{PR_-(I-P)}: C_\omega(K, \mathfrak{B}) \to C_\omega^+(K, \mathfrak{B})$ конечномерен, так как

$$(\mathbf{PR_-(I-P)}f)(\zeta) = \mathbf{P}(\zeta^S R(\mathbf{I}-\mathbf{P})f(\zeta)) = Rf_S + \zeta R f_{S-1} + \dots + \zeta^{S-1} R f_1$$

где $f_1, f_2, \dots$ — коэффициенты Тейлора функции $f(\zeta^{-1})$ в точке $\zeta = 0$.

Из равенства

$$\mathbf{PB_-A} = \mathbf{PR_-PN_-A} + \mathbf{PR_-(I-P)N_-A}$$

вытекает, что $\mathbf{PB_-A}$ является Φ-оператором.

Докажем теперь соотношение (2. 2). Отметим сначала, что все вектор-функции вида $\varphi_+(\zeta)y$ из $C_\omega^+(K, \mathfrak{B})$, где φ_+ — скалярная функция, принадлежат пространству $\operatorname{Im} \mathbf{PB_-A}$. А самом деле, в силу равенств (2. 13) и (2. 14) для функции $\varphi_+ \tilde{f}_+ \in C_\omega^+(K, \mathfrak{B})$ имеют место равенства

$$\mathbf{P}[B_-(\zeta) A(\zeta) \varphi_+(\zeta) \tilde{f}_+(\zeta)] = \mathbf{P}[B_-(\zeta) \varphi_+(\zeta) \zeta^{-S} \tilde{f}_-(\zeta)] =$$

$$= \mathbf{P}[(I - R + \zeta^S R) \varphi_+(\zeta) \zeta^{-S} y] = \mathbf{P}[\varphi^+(\zeta) y] = \varphi_+(\zeta) y.$$

Пусть тепрь g_+ — любая вектор-функция из $C_\omega^+(K, \mathfrak{B}) = \operatorname{Im} \mathbf{PN_-A}$. Тогда существуют функции $h_\pm \in C_\omega^\pm(K, \mathfrak{B})$ такие, что

$$\mathbf{N_-A} h_+ = g_+ + h_-. \tag{2. 15}$$

Имеет место равенство

$$\mathbf{P}[B_-(\zeta) A(\zeta) h_+(\zeta)] = \mathbf{P}[(I - R + \zeta^S R)(g_+(\zeta) + h_-(\zeta))] =$$

$$= g_+(\zeta) - R g_+(\zeta) + w_+(\zeta),$$

где $w_+(\zeta) = \mathbf{P}[\zeta^s R(g_+(\zeta) + h_-(\zeta))]$. Вектор-функции $Rg_+(\zeta)$ и $w_+(\zeta)$ имеют вид $\varphi_+(\zeta)y$, где φ_+ — скалярные функции. Следовательно, они принадлежат подпространству $\operatorname{Im} \mathbf{PB_-A}$. Отсюда вытекает, что и вектор $g_+ \in \operatorname{Im} \mathbf{PB_-A}$. Таким образом, $C_\omega^+(K' \mathfrak{B}) = \operatorname{Im} \mathbf{PB_-A}$.

Докажем теперь соотношения

$$\text{Ker}\,\mathbf{PB_-A} \subseteqq \text{Ker}\,\mathbf{PA} \tag{2.16}$$

и

$$\tilde{f}_+ \in \text{Ker}\,\mathbf{PA}/\text{Ker}\,\mathbf{PB_-A}, \tag{2.17}$$

из которых будет вытекать неравенство (2. 3).

Пусть g_+ — любая вектор-функция из подпространства $\text{Ker}\,\mathbf{PB_-A}$, т. е. $g_- \overset{\text{def}}{=} \mathbf{B_-A}g_+ \in C_\omega^-(K, \mathfrak{B})$. Тогда

$$A(\zeta)g_+(\zeta) = N_-^{-1}(\zeta)(I-R+\zeta^{-S}R)g_-(\zeta) \in C_\omega^-(K, \mathfrak{B}),$$

т. е. $g_+ \in \text{Ker}\,\mathbf{PA}$.

Соотношение $\tilde{f}_+ \in \text{Ker}\,\mathbf{PA}$ следует из равенства (2. 14). В силу равенств (2. 14) и (2. 13) получаем

$$\mathbf{P}[B_-(\zeta)A(\zeta)\tilde{f}_+(\zeta)] = \mathbf{P}[(I-R+\zeta^S R)N_-(\zeta)\zeta^{-S}\tilde{f}_-(\zeta)] =$$
$$= \mathbf{P}[(I-R+\zeta^S R)\zeta^S y] = y \neq 0.$$

Следовательно, $\tilde{f}_+ \notin \text{Ker}\,\mathbf{PB_-A}$.

Лемма доказана.

4. Доказательству теоремы 2. 1 предпошлем еще одну лемму, позволяющую допилнительно предполагать, что оператор-функция $A(\zeta)$ голоморфна на $\overline{K}$.

Лемма 2. 3. *Пусть оператор-функция* $A:\overline{K}\to GL(\mathfrak{B})$ *голоморфна в* K а *непрерывна на* $\overline{K}$. *Тогда существуют оператор-функция* $E_+:\overline{K}\cup F^+\to GL(\mathfrak{B})$ *голоморфная в* $K\cup F^+$ *и непрерывная на* $\overline{K}\cup F^+$, *и оператор-функция* $E_-:\overline{K}\cup \cup F^-\to GL(\mathfrak{B})$, *голоморфная в* $K\cup F^-$ *и непрерывная на* $\overline{K}\cup F^-$, *такие, что оператор-функция* E_-AE_+ — *голоморфна на* $\overline{K}$.

Доказательство. Обозначим через δ_K константу из леммы 1. 1. В силу леммы 1. 2 оператор-функцию можно представить в виде $A = (I+M)G$, где $M, G \in C_\omega(K, L(\mathfrak{B}))$, причем $\|M\|_{C_\omega(K, L(\mathfrak{B}))} < \delta_K$ и оператор-функция G голоморфна на $\overline{K}$. В силу леммы 1. 1 оператор-функция $I+M$ допускает каноническую факторизацию $I+M = X_-X_+$ относительно контура Γ. Оператор-функция X_+G, очевидно, голоморфна на $\overline{K}\cap(K\cup F^+)$. С помощью леммы 1. 2 оператор-функцию X_+G также можно представить в виде $X_+G = H(I+N)$, где H голоморфна на $\overline{K}$ и $\|N\|_{C_\omega(K, L(\mathfrak{B}))}<\delta_K$. Легко видеть, что оператор-функция N голоморфна на $\overline{K}\cap(K\cup F^+)$. В силу леммы 1. 1 оператор-функция $I+N$ допускает каноническую факторизацию $I+N = Y_-Y_+$ относительно Γ. Легко видеть, что оператор-функция Y_- голоморфна на $\overline{K}\cap(K\cup F^+)$ и, следовательно, она голоморфна на $\overline{K}\cup F^-$. Таким образом, $X_-^{-1}AY_+^{-1} = HY_-$, причем оператор-функция HY_- голоморфна $\overline{K}$.

Лемма доказана.

Доказательство теоремы 2. 1. Легко видеть, что условие теоремы является необходимым (см., например, доказательство теоремы 6. 1). Покажем его достаточность. Пусть оператор **PA** является Φ — оператором в $C^+_\omega(K, \mathfrak{B})$ и $E_\pm$ — оператор-функции из леммы 2. 3. Так как операторы $\mathbf{PE}_\pm$ обратимы в $C^+_\omega(K, \mathfrak{B})$ $((\mathbf{PE}_\pm)^{-1}=\mathbf{PE}^{-1}_\pm)$, то оператор $\mathbf{PE}_-\mathbf{AE}_+$ также является Φ — оператором в $C^+_\omega(K, \mathfrak{B})$. Оператор-функция E_-AE_+ голоморфна на $\bar{K}$. Применяя к ней последовательно несколько раз лемму 2. 1 и затем лемму 2. 2, получим голоморфные оператор-функции B_- и $B_+:\bar{K}\to GL(\mathfrak{B})$, допускающие вполне нормальные расширения соответственно в F^- и F^+, такие, что оператор $\mathbf{PB}_-\mathbf{E}_-\mathbf{AE}_+\mathbf{B}_+$ обратим в пространстве $C^+_\omega(K, \mathfrak{B})$.

Из теоремы 1. 1 вытекает, что оператор-функция $B_-E_-AE_+B_+$ допускает каноническую факторизацию $B_-E_-AE_+B_+=G_-G_+$ относительно контура Γ. Таким образом, для оператор-функции A получаем представление

$$A = (E_-^{-1}B_-^{-1}G_-)(G_+B_+^{-1}E_+^{-1}).$$

Согласно предложению 1. 1 оператор-функции $E_-^{-1}B_-^{-1}G_-$ и $G_+B_+^{-1}E_+^{-1}$ вполне нормальны соотвтственно в F^- и F^+. Отсюда в сулу теоремы 1. 2 следует, что оператор-функция A допускает факторизацию.

Теорема доказана.

5. Сделаем два замечания. Пусть $K-\Gamma$ — кольцо и оператор-функция $A:\bar{K}\to GL(\mathfrak{B})$ принадлежит $C_\omega(K, L(\mathfrak{B}))$. Тогда легко проверить, что имеет место равенство

$$\mathbf{Q}+\mathbf{PAP} = \mathbf{A}(\mathbf{P}+\mathbf{QA}^{-1}\mathbf{Q})(\mathbf{I}+\mathbf{PA}^{-1}\mathbf{Q})(\mathbf{I}-\mathbf{QAP}).$$

Так как операторы **A**, $\mathbf{I}+\mathbf{PA}^{-1}\mathbf{Q}$ и $\mathbf{I}-\mathbf{QAP}$ обратимы, то отсюда следует, что оператор **PA** является Φ-оператором в $C^+_\omega(K, \mathfrak{B})$ в том и только том случае, когда оператор $\mathbf{QA}^{-1}$ является Φ-оператором в $C^-_\omega(K, \mathfrak{B})$. При этом

$$\dim \operatorname{Ker} \mathbf{PA} = \dim \operatorname{Ker} \mathbf{QA}^{-1} \text{ и } \dim \operatorname{Coker} \mathbf{PA} = \dim \operatorname{Coker} \mathbf{QA}^{-1}.$$

Следовательно, в формулировке, теоремы 2. 1 оператор **PA** можно заменить оператором $\mathbf{QA}^{-1}$. Отметим еще, что если в предположениях этого предложения оператор-функция A допускает факторизацию $A=A_-DA_+$ относительно Γ, то, как легко видеть, имеют место равенства

$$\dim \operatorname{Ker} \mathbf{PA} = \dim \operatorname{Ker} \mathbf{QA}^{-1} = \dim \operatorname{Ker} \mathbf{D} = \sum_{\varkappa_j<0} \varkappa_j$$

и

$$\dim \operatorname{Coker} \mathbf{PA} = \dim \operatorname{Coker} \mathbf{QA}^{-1} = \dim \operatorname{Coker} \mathbf{D} = \sum_{\varkappa_j>0} \varkappa_j,$$

где $\varkappa_j$ — все частные индексы оператор-функции A.

§ 3. Обобщенная факторизация оператор-функций относительно контура

1. Пусть оператор-функция $A:\Gamma\to GL(\mathfrak{B})$ имеет вид

$$A(\zeta) = \sum_{j=-m}^{n} \zeta^j A_j, \tag{3.1}$$

где $A_j\in L(\mathfrak{B})$. Не всякая такая оператор-функция допускает факторизацию относительно контура Γ. В этом можно убедиться на примере $A_0(\zeta)=\zeta I$.

Рассмотрим менее жесткое понятие факторизации, определяемое следующим образом. *Обобщенной факторизацией* оператор-функции $A:\Gamma\to GL(\mathfrak{B})$ относительно контура Γ назовем ее представление в виде

$$A = A_- D A_+,$$

где оператор-функция $A_\pm:\bar{F}^\pm\to GL(\mathfrak{B})$ непрерывна $\bar{F}^\pm$ и голоморфна в $F^\pm$, а оператор-функция D имеет вид

$$D(\zeta) = \sum_{j=0}^{n} \zeta^{\varkappa_j} P_j,$$

где $P_1, \ldots, P_n$ — попарно дизъюнктные (т. е. $P_jP_k=0$ при $j\neq k$) проекторы из $L(\mathfrak{B})$, для которых $\sum P_j=I$.

В отличие от обычной факторизации в определении обобщенной факторизации не накладываются никакие ограничения на размерности проекторов P_j и на целые числа $\varkappa_j$ ($j=0, 1, \ldots, n$).

Как будет показано ниже, приведенное новое определение еще не является достаточно общим. Оказывается, что не всякая оператор-функция вида (3.1) допускает обобщенную факторизацию. Это будет доказано с помощью специального примера. Сперва установим одно необходимое условие для того, чтобы оператор-функция допускала обобщенную факторизацию.

Пусть $\mathfrak{H}$ — гильбертово пространство и $\mathbf{P}$ — проектор, проектирующий $\mathbf{L}_2(\Gamma,\mathfrak{H})$ на $\mathbf{L}_2^+(\Gamma,\mathfrak{H})$ параллельно $\mathbf{L}_2^-(\Gamma,\mathfrak{H})$*).

Предложение 3.1. *Для того чтобы непрерывная оператор-функция $A:\Gamma\to GL(\mathfrak{H})$ допускала обобщенную факторизацию относительно Γ, необходимо, чтобы оператор $\mathbf{PA}$ имел замкнутое множество значений в $\mathbf{L}_2^+(\Gamma,\mathfrak{H})$, т. е. он был нормально разрешим в $\mathbf{L}_2^+(\Gamma,\mathfrak{H})$.*

В самом деле, предположим, что оператор-функцию $A(\zeta)$ допускает обобщенную факторизацию $A=A_-DA_+$, тогда будем иметь $\mathbf{PA}=(\mathbf{PA}_-)(\mathbf{PD})(\mathbf{PA}_+)$.

---

*) Через $\mathbf{L}_2^+(\Gamma,\mathfrak{H})$ ($\mathbf{L}_2^-(\Gamma,\mathfrak{H})$) обозначается замыкание множества голоморфных функций $f:F^+\to\mathfrak{H}$ ($f:F^-\to\mathfrak{H}$, обращающиеся в нуль на бесконечности). Нетрудно показать (подробно это сделано в [8]), что $\mathbf{L}_2(\Gamma,\mathfrak{H}) = \mathbf{L}_2^+(\Gamma,\mathfrak{H})\dotplus\mathbf{L}_2^-(\Gamma,\mathfrak{H})$.

Операторы $\mathbf{PA}_{\pm}$ обратимы, причем $(\mathbf{PA}_{\pm})^{-1}=\mathbf{PA}_{\pm}^{-1}$. Оператор $\mathbf{PD}$, очевидно, распадается в прямую сумму односторонне обратимых операторов. Следовательно, оператор $\mathbf{PD}$, а вместе с ним и оператор $\mathbf{PA}$ имеют замкнутые множества значений.

2. Пусть Γ_0 — единичная окружность и $S=[0, 1]$ — единичный интервал. Рассмотрим гильбертово пространство $L_2^2(S)=L_2(S)\oplus L_2(S)$ двумерных вектор-функций с координатами из $L_2(S)$ и оператор-функцию $B:\Gamma_0\to L\big((L_2^2(S)\big)$, определенную равенством

$$B(\zeta) = \zeta^{-1}\beta_{-1}+B_0+\zeta B_1 \tag{3.2}$$

где $B_{-1}, B_0, B_1\in L\big(L_2^2(S)\big)$ — операторы умножения соответственно на матрицы-функции

$$\begin{pmatrix}0 & 0\\ t & 0\end{pmatrix}, \quad \begin{pmatrix}t-1 & 0\\ 0 & t+1\end{pmatrix} \quad \text{и} \quad \begin{pmatrix}0 & t\\ 0 & 0\end{pmatrix} \qquad (t\in S).$$

Так как

$$\det\begin{pmatrix}t-1 & \zeta t\\ \zeta^{-1}t & t+1\end{pmatrix}\equiv -1 \qquad (\zeta\in\Gamma_0,\ t\in S),$$

то $B(\zeta)\in GL\big(L_2^2(S)\big)$, для всех $\zeta\in\Gamma_0$.

Предложение 3. 2. *Оператор-функция B не допускает обобщенную факторизацию.*

Действительно, рассмотрим пространство $\mathbf{L}_2\big(\Gamma_0, L_2^2(S)\big)$ и оператор $\mathbf{PB}$ в $\mathbf{L}_2^+\big(\Gamma_0, L_2^2(S)\big)$. Согласно предпожению 3. 1 достаточно показать, что оператор $\mathbf{PB}$ не является нормально разрешимым.

Обозначим через $L_2^2(\Gamma_0)\,\big(=L_2(\Gamma_0)\oplus L_2(\Gamma_0)\big)$ гильбертово пространство двумерных вектор-функций с координатами из $L_2(\Gamma_0)$. Как известно, пространство $L_2(\Gamma_0)$ является распадающимся: $L_2(\Gamma_0)=L_2^+(\Gamma_0)\oplus L_2^-(\Gamma_0)$. Следовательно, $L_2^2(\Gamma_0)={}^+L_2^2(\Gamma_0)\oplus{}^-L_2^2(\Gamma_0)$, где ${}^{\pm}L_2^2(\Gamma_0)=L_2^{\pm}(\Gamma_0)\oplus L_2^{\pm}(\Gamma_0)$. Пусть P_2 — проектор, проектирующий $L_2^2(\Gamma_0)$ на ${}^+L_2^2(\Gamma_0)$ параллельно ${}^-L_2^2(\Gamma_0)$. Пространство $\mathbf{L}_2^+\big(\Gamma_0, L_2^2(S)\big)$ можно естественным образом отождествить с пространством $\mathbf{L}_2\big(S, {}^+L_2^2(\Gamma_0)\big)$. Обозначим через $G_t(t\in S)$ оператор умножения на матрицу-функцию

$$G_t(\zeta) = \begin{pmatrix}t-1 & \zeta t\\ \zeta^{-1}t & t+1\end{pmatrix} \qquad (\zeta\in\Gamma_0),$$

действующий в пространстве $L_2^2(\Gamma_0)$, а через $\mathbf{V}$ — оператор умножения на оператор-функцию $P_2G_t(t\in S)$, действующий в пространстве $\mathbf{L}_2\big(S, {}^+L_2^2(\Gamma_0)\big)$.

Легко видеть, что оператор $\mathbf{PB}$, рассматриваемый в пространстве $\mathbf{L}_2\big(S, {}^+L_2^2(\Gamma_0)\big)$, совпадает с оператором $\mathbf{V}$. Поэтому осталось показать, что оператор $\mathbf{V}$ не является нормально разрешимым.

Для всех $t \in S \setminus \{1\}$ матрица-функция $G_t(\zeta)$ $(\zeta \in \Gamma_0)$ допускает каноническую факторизацию относительно Γ:

$$G_t(\zeta) = \begin{pmatrix} 1 & 0 \\ \dfrac{-t}{\zeta(t-t)} & 1 \end{pmatrix} \begin{pmatrix} t-1 & \zeta t \\ 0 & \dfrac{1}{1-t} \end{pmatrix}.$$

Таким образом, операторы $P_2 G_t$ обратимы в ${}^{+}L_2^2(\Gamma_0)$ для всех $t \in S \setminus \{1\}$. Отсюда, в частности, вытекает, что $\operatorname{Ker} \mathbf{V} = \{0\}$.

С другой стороны,

$$G_1(\zeta) = \begin{pmatrix} 0 & \zeta \\ \zeta^{-1} & 2 \end{pmatrix}.$$

Следовательно, вектор $(1, 0) \in {}^{+}L_2^2(\Gamma_0)$ принадлежит ядру оператора $P_2 G_1$. Учитывая непрерывность оператор-функции $P_2 G_t (t \in S)$, отсюда легко получить последовательность вектор-функций $f_n \in \mathbf{L}_2(S, {}^{+}L_2^2(\Gamma_0))$, таких, что $\|f_n\|_{\mathbf{L}_2} = 1$, а $\|\mathbf{V} f_n\|_{\mathbf{L}_2} \leq 1/n$. Вместе с равенством $\operatorname{Ker} \mathbf{V} = \{0\}$ это противоречит нормальной разрешимости оператора $\mathbf{V}$.

Предложение доказано.

3. Как показано в предложении 3. 1, нормальная разрешимость оператора $\mathbf{PA}$ в пространстве $L_2^{+}(\Gamma, \mathfrak{H})$ является необходимым условием для существования обобщенной факторизации любой оператор-функции $A: \Gamma \to GL(\mathfrak{H})$ вида (3. 1). Оказывается, что это условие не является достаточным. В самом деле, рассмотрим оператор-функцию

$$\tilde{B}(\zeta) = \zeta B(\zeta) (\zeta \in \Gamma_0),$$

где B — оператор-функция (3. 2).

Так как оператор-функция B не допускает обобщенную факторизацию относительно Γ_0, то $\tilde{B}$ также ее не допускает. Но оператор $\mathbf{P}\tilde{\mathbf{B}}$ обратим слева, так как оператор-функция $\tilde{B}$ голоморфна внутри окружности Γ_0 и, следовательно, $(\mathbf{P}\tilde{\mathbf{B}}^{-1})(\mathbf{P}\tilde{\mathbf{B}}) = \mathbf{P}\tilde{\mathbf{B}}^{-1}\tilde{\mathbf{B}} = \mathbf{P}$.

4. Если в определении обобщенной факторизации средний множитель D отнести к множителю A_-, то мы приходим к следующему дальнейшему обобщению понятия факторизации.

Неполной факторизацией непрерывной оператор-функции $A: \Gamma \to GL(\mathfrak{B})$ назовем ее представление в виде

$$A = \tilde{A}_- A_+,$$

где оператор-функция $A_+ : \bar{F}^+ \to GL(\mathfrak{B})$ непрерывна на $\bar{F}^+$ и голоморфна в F^+, а оператор-функция $\tilde{A}_- : \bar{F}^- \setminus \{\infty\} \to GL(\mathfrak{B})$ непрерывна на $\bar{F}^- \setminus \{\infty\}$ и голоморфна в $F^- \setminus \{\infty\}$.

Оказывается, что любую оператор-функцию $A: \Gamma \to GL(\mathfrak{B})$ вида (3. 1) можно неполно факторизовать. Более того, имеет место следующая теорема.

Теорема 3.1. *Любая голоморфная оператор-функция $A: \Gamma \to GL(\mathfrak{B})$ допускает неполную факторизацию.*

Эта теорема вытекает из более общего результата Л. Бунгарта [15] о том, что любое голоморфное расслоенное пространство со структурной группой $GL(\mathfrak{B})$ и базисом $\mathbf{C}^1$ является тривиальным.

Элементарное доказательство этой теоремы содержится в статье авторов [16].

Цитированная литература

[1] И. Ц. Гохберг, Задача факторизации оператор-функций, *Известия АН СССР, серия матем.*, **28:5** (1964), 1055—1082.

[2] И. Ц. Гохберг и М. Г. Крейн, Основные положения о дефектных числах, корневых числах и индексах линейных операторов, *Усп. матем. наук*, **12:2** (1957), 44—118.

[3] Н. И. Мусхелишвили, *Сингулярные интегральные уравнения* (Москва, 1968).

[4] М. С. Будяну и И. Ц. Гохберг, Общие теоремы о факторизации матриц-функций. II. Некоторые признаки и их следствия, *Матем. исследования*, **3:3** (Кишинев, 1968), 3—18.

[5] И. Ц. Гохберг и М. Г. Крейн, Системы интегральных уравнений на полупрямой с ядрами, зависящими от разности аргументов, *Усп. матем. наук*, **13:2** (1958), 3—72.

[6] И. Ц. Гохберг и Ю. Лайтерер, Факторизация оператор-функций относительно контура. I. Конечномероморфные оператор-функции, *Math. Nachrichten*, **52:1—6** (1972), 259—282.

[7] И. Ц. Гохберг и Ю. Лайтерер, О голоморфных вектор-функциях одного переменного. I. Функции на компакте, *Матем. исследования*, **7:4** (Кишинев, 1972), 60—84.

[8] И. Ц. Гохберг и Ю. Лайтерер, Общие теоремы о канонической факторизации оператор-функций относительно контура, *Матем. исследования*, **7:3** (Кишинев, 1972), 87—134.

[9] П. М. Блехер, О линейных операторах, аналитически зависящих от параметра, *ДАН СССР*, **5** (1969), 30—36.

[10] И. Ц. Гохберг и Е. И. Сигал, Операторные обобщения теоремы о логарифмическом вычете и теоремы Руше, *Матем. сборник*, **84 (126):4** (1971), 607—630.

[11] И. Ц. Гохберг и И. А. Фельдман, *Уравнения в свертках и проекционные методы их решения* (Москва, 1971).

[12] H. Röhl, Das Riemann-Hilbertsche Problem der linearen Differentialgleichungen, *Math. Annalen*, **133** (1957), 1—25.

[13] H. Grauert, Analytische Faserungen über holomorph Räumen, *Math. Annalen*, **135** (1958) 263—273.

[14] H. Cartan, Espaces fibrés analytiques, *Symposium International de Topologia Algebraica* (1958), 97—121.

[15] L. Bungart, On analytic fiber bundles. 1. Holomorphic fiber bundles with infinite dimensional fibres, *Topology*, **7** (1968), 55—68.

[16] И. Ц. Гохберг и Ю. Лайтерер, О коциклах, оператор-фтнкциях и семействах подпространств, *Матем. исследования*, **8:2** (Кишинев, 1973).

(Поступило 3/VII/1972)

Bernstein-type inequalities for families of multiplier operators in Banach spaces with Cesàro decompositions. I. General theory

By E. GÖRLICH, R. J. NESSEL and W. TREBELS in Aachen (FRG)

Dedicated to Professor B. Sz.-Nagy on the occasion of his 60th birthday on July 29, 1973, in admiration

1. Introduction

Let $(\Pi_n)_{2\pi}$ be the set of all trigonometric polynomials $f(x)=\sum_{k=-n}^{n} c_k e^{ikx}$ of degree n. The classical Bernstein inequality states that

$$\|(d/dx)^r f(x)\|_{X_{2\pi}} \leqq n^r \|f(x)\|_{X_{2\pi}} \qquad (f\in(\Pi_n)_{2\pi}), \tag{1. 1}$$

where $X_{2\pi}$ is any of the spaces $L^p_{2\pi}$, $1\leqq p<\infty$, or $C_{2\pi}$ of periodic functions (cf. Section 3). As is well known, this inequality plays a central role in the proof of inverse theorems concerning best approximation by trigonometric polynomials. In a very general setting it was recently shown in some basic work of BUTZER—SCHERER [3, 4] (see also [6, 7]) that one may always obtain inverse approximation theorems, provided an inequality of type (1. 1) is available. In their spirit we may formulate the following problem:

Let X be an arbitrary (real or complex) Banach space, $[X]$ the Banach algebra of all bounded linear operators of X into itself, and let $\{T(\varrho)\}_{\varrho>0}\subset[X]$ be a family of operators depending on a parameter $\varrho>0$ (tending to infinity). Suppose B to be a closed linear operator with domain $D(B)\subset X$ and range in X. The family $\{T(\varrho)\}$ is said to satisfy a Bernstein-type inequality (with respect to B) if $T(\varrho)(X)\subset D(B)$ for each $\varrho>0$, and if there exists $\Omega(\varrho)>0$, defined on $(0,\infty)$, and a constant $A>0$ such that

$$\|BT(\varrho)f\| \leqq A\Omega(\varrho)\|f\| \qquad (f\in X,\ \varrho>0). \tag{1. 2}$$

In this paper we would like to study (1. 2) in the setting of [2], i.e., the operators in question are generated via multipliers in connection with Fourier expansions corresponding to general decompositions of Banach spaces. Then Bernstein inequalities of type (1. 2) in fact lead to a study of uniformly bounded multipliers (cf. (2. 4)). This is considered in Section 2 which gives convenient sufficient criteria in connection with Cesàro-(C, j)-decompositions. The most concrete version regarding uniform bounds is given in Corollary 2. 4 for multipliers of Fejér's type. This is in fact induced by a fundamental work of Sz.-NAGY [12] on the representation of functions as trigonometric integrals. Indeed, the case $j=1$ of Corollary 2. 4 may be considered as an elementary version of general results in [12] which are in turn used there as multiplier criteria to establish far reaching direct approximation theorems for trigonometric polynomials. Section 3 is concerned with particular choices of $\{T(\varrho)\}$ and B for arbitrary spaces X and decompositions. At the end of this section the trigonometric system is considered, mainly to discuss the question to which extent the classical inequalities may be covered by the present methods. The main bulk of concrete applications, however, will follow in Part II, thus illustrating the usefulness of this simple but nevertheless general and unifying approach to the subject. Finally, let us emphasize that we do not plan to reconstruct the (sometimes) long development of certain instances of Bernstein-type inequalities; for a brief historical account one may consult [10] (seemingly the latest paper on the subject of a survey nature).

The authors are very grateful to Professor P. L. BUTZER who inspired the work, read the manuscript, and made many valuable suggestions. The contribution of W. TREBELS was supported by a DFG fellowship.

2. Bernstein-type inequalities

Let $\mathbf{Z}$, $\mathbf{P}$, $\mathbf{N}$ be the sets of all, of all non-negative, of all positive integers, respectively. Let $\{P_k\}_{k\in\mathbf{P}}$ be a total sequence of mutually orthogonal continuous projections on X, i.e., (i) $P_k\in[X]$ for each $k\in\mathbf{P}$, (ii) $P_k f=0$ for all $k\in\mathbf{P}$ implies $f=0$, (iii) $P_j P_k=\delta_{jk}P_k$, δ_{jk} being Kronecker's symbol. Then with each $f\in X$ one may associate its unique Fourier series expansion

$$f\sim\sum_{k=0}^{\infty}P_k f \qquad (f\in X). \tag{2. 1}$$

With s the set of all sequences $\gamma=\{\gamma_k\}_{k\in\mathbf{P}}$ of scalars, $\gamma\in s$ is called a multiplier for X (corresponding to $\{P_k\}$) if for each $f\in X$ there exists an element $f^\gamma\in X$ such that $\gamma_k P_k f=P_k f^\gamma$ for all $k\in\mathbf{P}$, thus

$$f^\gamma\sim\sum_{k=0}^{\infty}\gamma_k P_k f \qquad (f\in X). \tag{2. 2}$$

Obviously, $Gf=f^\gamma$ defines a bounded linear operator G on X by the closed graph theorem. Conversely, operators T on X which permit an expansion of type (2. 2), i.e. $P_k(Tf)=\tau_k P_k f$, are called multiplier operators. Denoting the set of all multipliers for X by $M=M(X;\{P_k\})$, with the natural vector operations, coordinate-wise multiplication, and norm

(2. 3) $$\|\gamma\|_M = \sup\{\|f^\gamma\| \,|\, f\in X,\ \|f\|\leqq 1\},$$

M is a commutative Banach algebra, isometrically isomorphic to the subspace $[X]_M\subset[X]$ of multiplier operators on X.

Let $\alpha\in s$ be arbitrary and let X^α be the set of all $f\in X$ for which there exists $f^\alpha\in X$ such that $\alpha_k P_k f=P_k f^\alpha$ for all $k\in\mathbf{P}$. Obviously, if B^α is the operator with domain $X^\alpha\subset X$ and range in X defined by $B^\alpha f=f^\alpha$, then B^α is a closed linear operator for each $\alpha\in s$. Furthermore, if $\{P_k\}$ is fundamental, i.e., the linear span of $\bigcup_{k=0}^{\infty} P_k(X)$ is dense in X, then B^α is densely defined for each $\alpha\in s$.

On restricting oneself to operators with the above multiplier structure one may rephrase problem (1. 2) in terms of the corresponding sequences, namely

Theorem 2. 1. *Let $\alpha\in s$ and $\{T(\varrho)\}\subset[X]_M$ be a family of multiplier operators with associated multipliers $\tau(\varrho)$. If $\alpha\tau(\varrho)\in M$ for each $\varrho>0$, and if there exists $\Omega(\varrho)>0$ and a constant $A>0$ such that*

(2. 4) $$\|\alpha\tau(\varrho)/\Omega(\varrho)\|_M\leqq A$$

uniformly for $\varrho>0$, then $\{T(\varrho)\}$ satisfies the Bernstein-type inequality

(2. 5) $$\|B^\alpha T(\varrho)f\|\leqq A\Omega(\varrho)\|f\| \qquad (f\in X,\ \varrho>0).$$

Indeed, let $U^{\alpha\tau(\varrho)}\in[X]_M$ be associated with $\alpha\tau(\varrho)$. Then for any $f\in X$, $\varrho>0$, and $k\in\mathbf{P}$

$$P_k(U^{\alpha\tau(\varrho)}f)=\alpha_k\tau_k(\varrho)P_k f=\alpha_k P_k(T(\varrho)f),$$

so that $T(\varrho)(X)\subset X^\alpha$ and $B^\alpha T(\varrho)f=U^{\alpha\tau(\varrho)}f$. In view of (2. 3—4) this implies (2. 5).

Therefore, in the present setting, the problem is to verify the multiplier condition, particularly (2. 4), thus to establish convenient criteria concerning uniformly bounded multipliers. To this end we follow up the lines of [2] (see also the literature cited there), assuming (essentially) that $\{P_k\}$ is a Cesàro-(C,j)-decomposition of X. For basic facts concerning those decompositions (and bases) one may consult [8], [9], [11].

Let the (C,j)-means of (2. 1) be defined for $j\in\mathbf{P}$ by

(2. 6) $$(C,j)_n f=(A_n^j)^{-1}\sum_{k=0}^{n}A_{n-k}^j P_k f, \qquad A_n^j=\binom{n+j}{n}.$$

Obviously $(C,j)_n$ coincides for $j=0$ with the nth partial sum operator $S(n)=\sum_{k=0}^{n} P_k$. For some fixed $j\in\mathbf{P}$ assume that $(C,j)_n$ is uniformly bounded, i.e.

$$\|(C,j)_n f\| \leqq C_j\|f\| \qquad (f\in X), \tag{2.7}$$

the constant $C_j(\geqq 1)$ being independent of $n\in\mathbf{P}$ and $f\in X$.

Remark. In many cases of interest (cf. Part II) one deals with Fourier series in X associated with a total biorthogonal system $\{f_k,f_k^*\}$, $\{f_k\}\subset X$, $\{f_k^*\}\subset X^*$ (the dual of X). Then (2.1) and (2.2) read

$$f\sim\sum_{k=0}^{\infty} f_k^*(f)f_k, \qquad Tf\sim\sum_{k=0}^{\infty}\tau_k f_k^*(f)f_k, \tag{2.8}$$

respectively; $P_k(X)$ is the one-dimensional linear space spanned by f_k. If, furthermore, $\{f_k\}$ is fundamental, then it is clear by the Banach—Steinhaus theorem that (2.7) for $j=0$ is equivalent to the assumption that $\{f_k\}$ is a Schauder basis, i.e., for every $f\in X$

$$\lim_{n\to\infty}\left\|\sum_{k=0}^{n} f_k^*(f)f_k-f\right\|=0,$$

whereas for $j=1$ condition (2.7) is equivalent to the statement that $\{f_k\}$ is a Cesàro basis, i.e., for every $f\in X$

$$\lim_{n\to\infty}\left\|\sum_{k=0}^{n}\left(1-\frac{k}{n+1}\right)f_k^*(f)f_k-f\right\|=0.$$

To study multipliers in connection with systems $\{P_k\}$ satisfying (2.7), let us introduce the following spaces of (scalar-) sequences:

$$bv_{j+1}=\left\{\gamma\in l^\infty \,\Big|\, \|\gamma\|_{bv_{j+1}}=\sum_{k=0}^{\infty}\binom{k+j}{j}|\Delta^{j+1}\gamma_k|+\lim_{m\to\infty}|\gamma_m|<\infty\right\}, \tag{2.9}$$

$$l^\infty=\{\gamma\in s \mid \sup_{k\in\mathbf{P}}|\gamma_k|<\infty\},\quad \Delta\gamma_k=\gamma_k-\gamma_{k+1},\quad \Delta^{j+1}=\Delta^j\Delta.$$

Note that $\gamma\in l^\infty$ and the convergence of the series in (2.9) imply the existence of the limit $\lim_{m\to\infty}\gamma_m=\gamma_\infty$. Furthermore, $bv_{j+1}\subset bv_j$ in the sense of continuous embedding (cf. [5]). Obviously, bv_{j+1} is the space of all sequences of bounded variation if $j=0$, and the space of all bounded, quasi-convex sequences if $j=1$, respectively.

Theorem 2.2. *Let $\{P_k\}\subset[X]$ be a total sequence of mutually orthogonal projections satisfying* (2.7) *for some $j\in\mathbf{P}$. Then every $\gamma\in bv_{j+1}$ is a multiplier and*

$$\|\gamma\|_M\leqq C_j\|\gamma\|_{bv_{j+1}}. \tag{2.10}$$

Indeed, to each $f \in X$ one may associate (cf. [2II])

$$f^{\gamma} = \sum_{k=0}^{\infty} \binom{k+j}{j} \Delta^{j+1} \gamma_k \cdot (C, j)_k f + \gamma_{\infty} f.$$

Therefore, to verify (2. 4) one has to check whether (for suitably chosen $\Omega(\varrho)$) the bv_{j+1}-norms of the sequences $\{\alpha_k \tau_k(\varrho)/\Omega(\varrho)\}_{k\in\mathbf{P}}$ are uniformly bounded for $\varrho > 0$. For this purpose, let BV_{j+1} be the class of all bounded continuous functions f on $[0, \infty)$ for which $f, \dots, f^{(j-1)}$ are locally (i.e. on every compact subinterval) absolutely continuous on $(0, \infty)$ and $f^{(j)}$ is locally of bounded variation on $(0, \infty)$ such that $\int_0^\infty x^j |df^{(j)}(x)| < \infty$.

Then one may use the following result (cf. [2II])

Theorem 2. 3. *Let $\gamma \in s$ be such that there exists a function $g \in BV_{j+1}$ with $\gamma_k = g(k)$. Then $\gamma \in bv_{j+1}$ and*

$$\sum_{k=0}^{\infty} \binom{k+j}{j} |\Delta^{j+1} \gamma_k| \leqq D \int_0^\infty x^j |dg^{(j)}(x)|, \tag{2. 11}$$

the constant D being independent of γ and j.

As an immediate consequence one has the following criterion concerning uniformly bounded multipliers.

Corollary 2. 4. *Let the system $\{P_k\}$ satisfy (2. 7) for some $j \in \mathbf{P}$. Let $\{\gamma(\varrho)\} \subset s$ be such that there exists $\{g_\varrho\} \subset BV_{j+1}$ with $\lim_{x\to\infty} g_\varrho(x) = 0$ and $\gamma_k(\varrho) = g_\varrho(k)$ for each $k \in \mathbf{P}$, $\varrho > 0$. Then*

$$\|\gamma(\varrho)\|_M \leqq C_j D \int_0^\infty x^j |dg_\varrho^{(j)}(x)|. \tag{2. 12}$$

In particular, if g_ϱ is of Fejér's type, i.e., there exists $G \in BV_{j+1}$ such that $g_\varrho(x) = G(x/\varrho)$, then $\{\gamma(\varrho)\}$ is a family of uniformly bounded multipliers.

3. Particular operators in arbitrary spaces

Let X be an arbitrary Banach space and $\{P_k\} \subset [X]$ be any total system of orthogonal projections satisfying (2. 7) for some $j \in \mathbf{P}$. In this section we would like to discuss certain particular choices of families $\{T(\varrho)\}$ and sequences α. Throughout this section A stands for constants which may generally be distinct.

First, let us consider Bernstein inequalities of the classical type (1. 1). Here it is essential that the elements f only belong to the direct sum

$$\Pi_n = \bigoplus_{k=0}^{n} P_k(X) = \left\{ f \in X \,\middle|\, f = \sum_{k=0}^{n} P_k f \equiv S(n) f \right\}$$

rather than to the whole space X. In reducing this situation to that of Theorem 2. 1, we will have to restrict ourselves to the cases $j=0$ or $j=1$.

In case $j=0$ one has $\|\sigma(n)\|_M \leqq C_0$ by hypothesis, $\sigma(n) \in M$ being associated with the partial sum operator $S(n)$. For given non-negative $\alpha \in s$ consider $\alpha\sigma(n)$, the continuous parameter ϱ being replaced by the discrete one n. Since $\alpha\sigma(n) = \beta(n)\sigma(n)$ with $\beta_k(n) = \alpha_k$ for $0 \leqq k \leqq n$, $= \alpha_n$ for $k > n$, Theorems 2. 1—2 imply

$$\|B^\alpha f\| \leqq A\alpha_n \|f\| \qquad (f \in \Pi_n), \tag{3. 1}$$

provided $\|\beta(n)\|_{bv_1} \leqq A\alpha_n$ for all $n \in \mathbf{P}$. In particular, if α is monotonely increasing on $\mathbf{P}$, then $\|\beta(n)\|_{bv_1} = \alpha_n - \alpha_0$.

In case $j=1$ consider the family $\{I(n)\} \subset [X]_M$ with associated $\imath(n) \in M$, defined by $\imath_k(n) = 1$ for $0 \leqq k \leqq n$, $= 2-(k/n)$ for $n < k \leqq 2n$, $=0$ for $k > 2n$. Then $\imath(n) \in bv_2$ uniformly for $n \in \mathbf{P}$, and the restriction of $I(n)$ to Π_n is the identity mapping. For given non-negative $\alpha \in s$ consider $\alpha\imath(n)$. Since $\alpha\imath(n) = \eta(n)\imath(n)$ with

$$\eta_k(n) = \alpha_k \text{ for } 0 \leqq k \leqq 2n, = \alpha_{2n} \text{ for } k > 2n, \tag{3. 2}$$

it is sufficient to examine $\|\eta(n)\|_{bv_2}$ in order to apply Theorems 2. 1—2. 2. Thus for the restriction of $B^\alpha I(n)$ to Π_n we have

Proposition 3. 1. *Let the system $\{P_k\}$ satisfy* (2. 7) *for $j=1$. Let $\alpha \in s$ be non-negative and assume that $\eta(n)$ is defined by* (3. 2) *and satisfies $\|\eta(n)\|_{bv_2} \leqq A\alpha_{2n}$ for all $n \in \mathbf{P}$. Then*

$$\|B^\alpha f\| \leqq A\alpha_{2n} \|f\| \qquad (f \in \Pi_n). \tag{3. 3}$$

In particular, Proposition 3. 1 immediately applies to concave sequences α. For, then α is monotonely increasing so that also $\eta(n)$ of (3. 2) is concave, and thus $\|\eta(n)\|_{bv_2} = \alpha_{2n} - \alpha_0$. Concerning convex sequences α compare the remarks at the end of this section.

In this paper we restrict ourselves to three illustrative examples of sequences α, the significance of this choice in approximation theory being exhibited in [6, 7]. Let $\omega > 0$ be arbitrary, fixed. Then

$$\text{(i) } \alpha = \{k^\omega\}_{k \in \mathbf{P}}, \quad \text{(ii) } \alpha = \{\log(1+k^\omega)\}_{k \in \mathbf{P}}, \quad \text{(iii) } \alpha = \{e^{a(k)}\}_{k \in \mathbf{P}}, \tag{3. 4}$$

where $a(x)$ is a non-negative function, defined and monotonely increasing on $[0, \infty)$.

Obviously, (3. 1) applies to (3. 4) (iii) in case $j=0$. Concerning examples (3. 4) (i), (ii) it follows for the corresponding $\eta(n)$ (cf. (3. 2)) that $\|\eta(n)/\alpha_{2n}\|_{bv_2}\leqq A$ uniformly for $n>0$ by Corollary 2. 4 (for (ii) cf. (2. 12)). Thus

Corollary 3. 2. (a) *Let the system* $\{P_k\}$ *satisfy* (2. 7) *for* $j=0$. *Given* $a(x)$ *as specified in* (3. 4) (iii), *then for any* $f\in X$, $n\in\mathbf{P}$

$$\left\|\sum_{k=0}^{n} e^{a(k)}P_k f\right\| \leqq Ae^{a(n)}\left\|\sum_{k=0}^{n} P_k f\right\|. \tag{3. 5}$$

(b) *Let the system* $\{P_k\}$ *satisfy* (2. 7) *for* $j=1$. *Then for any* $\omega>0$ *and* $f\in X$, $n\in\mathbf{P}$

$$\left\|\sum_{k=0}^{n} k^{\omega}P_k f\right\| \leqq An^{\omega}\left\|\sum_{k=0}^{n} P_k f\right\|, \tag{3. 6}$$

$$\left\|\sum_{k=0}^{n} \log(1+k^{\omega})P_k f\right\| \leqq A\log(1+n^{\omega})\left\|\sum_{k=0}^{n} P_k f\right\|. \tag{3. 7}$$

In each case the constant A is independent of $f\in X$, $n\in\mathbf{P}$.

Now, let us apply Theorem 2. 1 directly to several particular families $\{T(\varrho)\}$. We consider the Abel—Cartwright means of order $\varkappa>0$ of the Fourier series (2. 1) of f

$$\text{(i) } W_\varkappa(\varrho)f\sim\sum_{k=0}^{\infty} e^{-(k/\varrho)^\varkappa}P_k f \qquad (f\in X,\ \varrho>0), \tag{3. 8}$$

the Bessel potentials of order $\varkappa>0$

$$\text{(ii) } L_\varkappa(\varrho)f\sim\sum_{k=0}^{\infty}(1+(k/\varrho)^2)^{-\varkappa/2}P_k f \qquad (f\in X,\ \varrho>0), \tag{3. 8}$$

and the Riesz means of order $\varkappa, \lambda>0$ ($\varrho = n+1\in\mathbf{N}$ being discrete)

$$\text{(iii) } R_{\varkappa,\lambda}(n)f\sim\sum_{k=0}^{n}\left(1-\left(\frac{k}{n+1}\right)^{\varkappa}\right)^{\lambda}P_k f \qquad (f\in X,\ n\in\mathbf{P}). \tag{3. 8}$$

Since (cf. [2II]) $\|P_k\|_{[X]}\leqq Ak^j$ in case (2. 7) holds for $j\in\mathbf{P}$, one has equality for all $\varrho>0$ in (i) for $\varkappa>0$, in (ii) for $\varkappa>j+1$, and trivially in (iii) for $\varkappa$, $\lambda>0$. Furthermore, $L_\varkappa(\varrho)\in[X]_M$ for all $\varkappa>0$ since $(1+x^2)^{-\varkappa/2}\in BV_{j+1}$.

For these families $\{T(\varrho)\}$ let us consider $\alpha=\{k^\omega\}$, $\omega>0$, with $\Omega(\varrho)=\varrho^\omega$. For the corresponding $\alpha\tau(\varrho)$ one has

$$\frac{\alpha_k\tau_k(\varrho)}{\Omega(\varrho)}=g_\varrho(k)=G\left(\frac{k}{\varrho}\right),\quad G(x)=\begin{cases} x^\omega\exp(-x^\varkappa), \\ x^\omega(1+x^2)^{-\varkappa/2}, \\ \begin{cases} x^\omega(1-x^\varkappa)^\lambda, & 0\leqq x\leqq 1, \\ 0\quad, & x>1, \end{cases}\end{cases}$$

respectively. Since $G\in BV_{j+1}$ for each $\omega, \varkappa>0$ in case (i), for each $0<\omega<\varkappa$ in case (ii), and for each $\lambda\geqq j$ and $\omega, \varkappa>0$ in case (iii), it follows by Corollary 2. 4 that

Corollary 3. 3. *Let the system* $\{P_k\}$ *satisfy* (2. 7) *for some* $j\in\mathbf{P}$. *Then for every* $f\in X$, $\varrho>0$ ($n\in\mathbf{P}$):

$$\left\|\sum_{k=0}^{\infty} k^{\omega} e^{-(k/\varrho)^{\varkappa}} P_k f\right\| \leqq A\varrho^{\omega}\|f\| \qquad (\varkappa, \omega > 0), \tag{3.9}$$

$$\|B^{\{k^{\omega}\}} L_{\varkappa}(\varrho) f\| \leqq A\varrho^{\omega}\|f\| \qquad (0<\omega<\varkappa), \tag{3.10}$$

where for $0<\omega<\varkappa-j-1$ *the corresponding sum exists and therefore*

$$\left\|\sum_{k=0}^{\infty} k^{\omega}(1+(k/\varrho)^2)^{-\varkappa/2} P_k f\right\| \leqq A\varrho^{\omega}\|f\| \qquad (0<\omega<\varkappa-j-1),$$

$$\left\|\sum_{k=0}^{n} k^{\omega}\left(1-\left(\frac{k}{n+1}\right)^{\varkappa}\right)^{\lambda} P_k f\right\| \leqq An^{\omega}\|f\| \qquad (\lambda\geqq j; \varkappa, \omega>0). \tag{3.11}$$

Analogously, Bernstein-type inequalities may be derived for further sequences α.

Remark. The methods employed here may also be used to treat the following counterpart to the general problem (1. 2):

Let $\{T^{(1)}(\varrho)\}$, $\{T^{(2)}(\varrho)\}\subset[X]$ be two families of operators and B a closed linear operator with domain $D(B)\subset X$ and range in X. The family $\{T^{(1)}(\varrho)\}$ is said to satisfy a Bernstein-type inequality (with respect to B and $\{T^{(2)}(\varrho)\}$) if $T^{(1)}(\varrho)(X)\subset$ $\subset D(B)$ for each $\varrho>0$, and if there exists $\Omega(\varrho)>0$ such that

$$\|BT^{(1)}(\varrho)f\| \leqq \Omega(\varrho)\|T^{(2)}(\varrho)f\| \qquad (f\in X, \varrho>0). \tag{3.12}$$

From the point of view of applications following in Part II, however, formulations (1. 2) and (3. 12) are parallel.

Furthermore, note that (2. 5) may be interpreted as a weak and (3. 1), (3. 3) as strong Bernstein-type inequalities, respectively, as introduced in Butzer—Scherer [3, 4]. However, for commutative operators (as considered here), (2. 5) may be sharper than (3. 1), (3. 3), as the particular de la Vallée Poussin process shows (cf. [3]). In the noncommutative case, strong Bernstein-type inequalities seem to be essential.

So far, we have discussed the results of Section 2 in connection with certain particular choices of families $\{T(\varrho)\}$ and sequences α for arbitrary Banach spaces X and systems $\{P_k\}$. Thus it remains to specify X and $\{P_k\}$. However, this will be examined in detail in Part II, devoted to explicit applications to classical orthogonal expansions. Here we only consider the trigonometric system in order to provide a feeling to which extent the classical results are covered by the present approach.

Let $X_{2\pi}=L^p_{2\pi}$, $1\leqq p\leqq\infty$, or $C_{2\pi}$ be the Banach space of 2π-periodic functions with standard norm $\|\cdot\|_{X_{2\pi}}$:

$$\Big(\int_{-\pi}^{\pi}|f(x)|^p dx\Big)^{1/p}(1\leqq p<\infty),\quad \text{ess. sup}\;|f(x)|,\quad \max|f(x)|,$$

respectively. Defining the system $\{P_k\}_{k\in\mathbf{P}}$ by

(3.13) $$(P_0f)(x)=f^\wedge(0),\quad (P_kf)(x)=f^\wedge(k)e^{ikx}+f^\wedge(-k)e^{-ikx}\qquad(k\in\mathbf{N}),$$

$f^\wedge(k)$ being the usual Fourier coefficient

$$f^\wedge(k)=(1/2\pi)\int_{-\pi}^{\pi}f(x)e^{-ikx}\,dx\qquad(k\in\mathbf{Z}),$$

$\{P_k\}$ is a total sequence of mutually orthogonal continuous projections on $X_{2\pi}$ and

(3.14) $$f\sim\sum_{k=0}^{\infty}P_kf\left(\equiv\sum_{k=-\infty}^{\infty}f^\wedge(k)e^{ikx}\right)\qquad(f\in X_{2\pi}).$$

It is well known that $\{P_k\}$ satisfies (2.7) with $j=0$ in case $X_{2\pi}=L^p_{2\pi}$, $1<p<\infty$, and with $j=1$ in all $X_{2\pi}$-spaces. Thus an application of (3.6) yields for any $\omega>0$

(3.15) $$\Big\|\sum_{k=-n}^{n}|k|^\omega c_ke^{ikx}\Big\|_{X_{2\pi}}\leqq An^\omega\Big\|\sum_{k=-n}^{n}c_ke^{ikx}\Big\|_{X_{2\pi}}.$$

Note that $\sum_{k=-n}^{n}|k|^\omega c_ke^{ikx}$ corresponds to the ωth Riesz derivative $t_n^{\{\omega\}}(x)$ of the trigonometric polynomial $t_n(x)=\sum_{k=-n}^{n}c_ke^{ikx}$ (for the definition and basic properties of this fractional derivative see [1, Sec. 11.5]).

Obviously, apart from the constants, (3.15) coincides with the classical inequality (1.1), thus with

(3.16) $$\Big\|\sum_{k=-n}^{n}(ik)^rc_ke^{ikx}\Big\|_{X_{2\pi}}\leqq n^r\Big\|\sum_{k=-n}^{n}c_ke^{ikx}\Big\|_{X_{2\pi}}$$

only in case of even values of r. The case of odd values, particularly $r=1$, is not covered for arbitrary $X_{2\pi}$-spaces.

Of course, there are several proofs of (3.16) for $r=1$ and all spaces $X_{2\pi}$, using particular features of the trigonometric system. Here we may mention the classical proof of F. RIESZ. In its extended form (cf. [6, 7]) it deals with (even or odd) sequences $\{\alpha_k\}_{k=-\infty}^{\infty}$, non-negative and convex on $\mathbf{P}$ with $\alpha_0=0$. Taking into account addition formulae, specific for the trigonometric system, the proof of the inequality

$$\Big\|\sum_{k=-n}^{n}\alpha_kc_ke^{ikx}\Big\|_{X_{2\pi}}\leqq 2\alpha_n\Big\|\sum_{k=-n}^{n}c_ke^{ikx}\Big\|_{X_{2\pi}}$$

reduces to a verification of the convexity on **P** of the sequence α_{n-k}/α_n for $0 \leqq k \leqq n$, 0 for $k > n$. Whether this method of proof may be extended to more general systems $\{P_k\}$ remains open.

Finally, let us observe that the classical Bernstein inequality (3. 16) for $r=1$, $X_{2\pi}=C_{2\pi}$, for example, may of course be derived by using different methods as a (direct) consequence of theorems in arbitrary Banach spaces. Thus, for example, one may take (3. 16) for $r=2$ and interpolation techniques in order to establish (3. 16) for any $0<r<2$ (see [13]).

References

[1] P. L. Butzer—R. J. Nessel, *Fourier Analysis and Approximation*. Vol. I., Birkhäuser, and Academic Press (Basel and New York, 1971).

[2] P. L. Butzer—R. J. Nessel—W. Trebels, On summation processes of Fourier expansions in Banach spaces. I. Comparison theorems; II. Saturation theorems, *Tôhoku Math. J.*, **24** (1972), 127—140; 551—569.

[3] P. L. Butzer—K. Scherer, Approximation theorems for sequences of commutative operators in Banach spaces, *Proc. the Conf. Constructive Function Theory*, Varna (Bulgaria) 1970. Publishing House Bulgarian Acad. Sci. (Sofia, 1972), 137—146.

[4] P. L. Butzer—K. Scherer, Jackson and Bernstein-type inequalities for families of commutative operators in Banach spaces, *J. Approximation Theory*, **5** (1972), 308—342.

[5] D. F. Dawson, Variation properties of sequences, *Mat. Vesnik*, **6** (21) (1969), 437—441.

[6] E. Görlich, *Logarithmische und exponentielle Ungleichungen vom Bernstein-Typ und verallgemeinerte Ableitungen*. Habilitationsschrift, TH Aachen 1971, pp. 1—68.

[7] E. Görlich, Logarithmic and exponential variants of Bernstein's inequality and generalized derivatives, *Linear Operators and Approximation* (P. L. Butzer—J. P. Kahane—B. Sz.-Nagy, Eds.), ISNM 20, Birkhäuser (Basel, 1972), 325—337.

[8] J. T. Marti, *Introduction to the Theory of Bases*, Springer (Berlin, 1969).

[9] V. D. Milman, Geometric theory of Banach spaces. I. The theory of bases and minimal systems, *Russian Math. Surveys*, **25** (1970), 111—170.

[10] A. C. Schaeffer, Inequalities of A. Markoff and S. Bernstein for polynomials and related functions, *Bull. Amer. Math. Soc.*, **47** (1941), 565—579.

[11] I. Singer, *Bases in Banach Spaces*. I, Springer (Berlin, 1970).

[12] B. Sz.-Nagy, Sur une classe générale de procédés de sommation pour les séries de Fourier, *Hungarica Acta Math.*, **1**, no **3** (1948), 14—52.

[13] W. Trebels—U. I. Westphal, A note on the Landau—Kallman—Rota—Hille inequality, *Linear Operators and Approximation* (P. L. Butzer—J. P. Kahane—B. Sz.-Nagy, Eds.), ISNM 20, Birkhäuser (Basel, 1972), 115—119.

(Received October 6, 1972)

Limits of shifts

By P. R. HALMOS in Bloomington (Indiana, U.S.A.)

Dedicated to B. Sz.-Nagy on his sixtieth birthday, July 29, 1973

What is the closure of the unilateral shifts?

The question looks odd; by long-standing tradition the unilateral shift is regarded as one operator, not a set of operators. On the occasions when the plural is used it usually indicates multiplicities, or weights, but neither of those is what is meant here. A moment's thought reveals that the question makes unambiguous sense. An operator S on a Hilbert space H is a unilateral shift (of multiplicity 1) in case there exists an orthonormal basis $\{e_0, e_1, e_2, \dots\}$ for H such that $Se_n = e_{n+1}$, $n = 0, 1, 2, \dots$. From this point of view there are as many unilateral shifts of multiplicity 1 as there are orthonormal bases enumerated by the natural numbers. The problem is to determine the closure of the set of all such shifts with respect to the norm topology of operators.

The same question can be asked and the same comments can be made about bilateral shifts, which shift an orthonormal basis enumerated by all integers.

Unilateral shifts are isometric, and, therefore, so are their limits. (Reason: if $S_n \to T$, then $S_n^* S_n \to T^* T$.) If, moreover, all the terms of a convergent sequence of unilateral shifts have the same multiplicity, then the co-rank of the limit is equal to that common multiplicity. (Reason: for n large, the projections $1 - S_n S_n^*$ and $1 - TT^*$ are near, and, therefore, they have the same rank; the rank of $1 - S_n S_n^*$ is the multiplicity of S_n.)

Bilateral shifts are unitary, and, therefore, so are their limits. Since, moreover, the spectrum of every bilateral shift is the entire unit circle, it follows that the spectrum of a limit of bilateral shifts is also the entire unit circle. (Reason: the spectrum is upper semicontinuous, [6, Problem 86].)

The preceding two paragraphs describe some necessary conditions that limits of shifts must satisfy; it is natural to ask how near those conditions come to being sufficient. Can a limit of unilateral shifts of multiplicity 1, say, have a unitary direct

*) Research supported in part by a grant from the National Science Foundation.

summand? Can a limit of bilateral shifts of multiplicity 1 have anything other than an absolutely continuous spectral measure of uniform multiplicity? On first consideration both questions seem to call for a negative answer. It is remarkable, however, that the already stated necessary conditions turn out to be sufficient also. The facts are described in the following statement; the main purpose of the sequel is to prove it.

Theorem. *On a separable Hilbert space the norm closure of the set of unilateral shifts of multiplicity n* ($1 \leqq n \leqq \infty$) *is the set of all isometries of co-rank n, and the norm closure of the set of bilateral shifts of multiplicity n* ($1 \leqq n \leqq \infty$) *is the set of all unitary operators whose spectrum is the entire unit circle.*

The proof uses a slight sharpening of the proof of a result of R. G. Douglas (which will be described later). That result became part of the oral tradition sometime in 1971. I learned the statement from P. A. Fillmore and the proof of the central lemma (which appears as Lemma 2 below) from I. D. Berg. A treatment of the Douglas result in an extended context is to appear later [3]. The present sharpening is applied, along the way, to the proof of a theorem of von Neumann's (the so-called von Neumann converse of Weyl's theorem [11]). The result (Lemma 4) is a quantitative improvement of von Neumann's theorem for a large class of normal operators (the ones for which the spectrum coincides with the essential spectrum).

Lemma 1. *If A is a normal operator on a separable Hilbert space, then $A = D+C$, where D is diagonal, with its spectrum included in that of A, and C is compact, with its norm arbitrarily small.*

Except for the statement about the spectrum of D, this is the Berg extension [2] to normal operators of the Weyl—von Neumann theorem [11] for Hermitian ones. In my subsequent proof [9] no restriction was placed on the spectrum of D or on the size of C. There is perhaps some merit in knowing that the restrictions can be captured in the framework of that proof; the next two paragraphs show how that can be done.

As far as the size of C is concerned, the result in the Hermitian case goes back to von Neumann [11], who proved that the compact summand of a Hermitian operator could in fact be made a Hilbert—Schmidt operator with arbitrarily small Hilbert—Schmidt norm. (Cf. also [8, p. 904].) To extend the result to the normal case, use the fact that if A is normal, then $A=\varphi(A')$, where A' is Hermitian and φ is continuous [9]. Recall now that the mapping $X \mapsto \varphi(X)$, defined for each Hermitian operator X whose spectrum is in the domain of φ, is continuous in the norm topology. (This is an easy exercise whose proof uses nothing more than the Weierstrass polynomial approximation theorem and the norm continuity of the algebraic operations on operators. The statement is true for continuous functions of normal operators, as well as Hermitian ones; the only additional technique needed is the planar version of the Weierstrass theorem.) Consequence: if $A' = D' + C'$, with D' diagonal and C' compact, the norm of the (compact) operator $C = A - D (= \varphi(A') - \varphi(D'))$ can be made as small as desired by making $\|C'\|$ small enough. (Observe that because of the passage to a limit implied by the formation of a continuous function, the Hilbert—Schmidt character of the compact summand cannot automatically be asserted in the normal case. It is not known whether the reason is in the proof or in the facts.)

The problem of putting the spectrum of D into the spectrum of A can be handled as follows. For each positive number δ, there can be only finitely many eigenvalues of D farther than δ from the spectrum of A. (Reason: otherwise the eigenvalues of D would have a cluster point not in the spectrum of A, in contradiction to the fact that A and D have the same essential spectrum.) Suppose now that $A = D + C$, with D diagonal, C compact, and $\|C\|$ small enough for two purposes: (1) if the ultimate C is to have norm below ε, make the present one have norm below $\varepsilon/2$, and (2) use the upper semicontinuity of the spectrum [6, Problem 86] to guarantee that if $\|A - X\| < \|C\|$, then the spectrum of X is in the $\varepsilon/2$ neighborhood of the spectrum of A. Consider, successively, the values of δ equal to $\|C\|, \|C\|/2, \|C\|/3, \ldots$, and, each time, replace the eigenvalues of D outside the δ neighborhood of the spectrum of A by numbers in the spectrum as near as possible. The total alteration is compact and has norm not more than $\varepsilon/2$. Absorb it in C, increasing $\|C\|$ thereby to ε at worst. In case C happened to have not only small norm but small Hilbert—Schmidt norm as well, the altered C will have the same property.

Lemma 2. *If S is a shift of multiplicity 1 (unilateral or bilateral), if $\{\lambda_1, \lambda_2, \lambda_3, \ldots\}$ is a sequence of complex numbers of modulus 1, and if $\varepsilon > 0$, then there exist operators D and E such that D is diagonal, with eigenvalues $\lambda_1, \lambda_2, \lambda_3, \ldots$, and such that $S = (D \oplus E) + C$, where C is a Hilbert—Schmidt operator with Hilbert—Schmidt norm not greater than ε.*

To prove the lemma, consider an orthonormal set $\{e_0, e_1, e_2, \ldots\}$ that S shifts. If $|\lambda| = 1$, $m = 0, 1, 2, \ldots$, and $n = 1, 2, 3, \ldots$, write

$$f = (1/\sqrt{n})(e_m + e_{m+1}/\lambda + \cdots + e_{m+n-1}/\lambda^{n-1}).$$

Clearly $\|f\| = 1$. Since

$$Sf = (1/\sqrt{n})(e_{m+1} + e_{m+2}/\lambda + \cdots + e_{m+n}/\lambda^{n-1})$$

$$= \lambda(1/\sqrt{n})(e_{m+1}/\lambda + e_{m+2}/\lambda^2 + \cdots + e_{m+n}/\lambda^n),$$

it follows that

$$Sf - \lambda f = (1/\sqrt{n})(e_{m+n}/\lambda^{n-1} - \lambda e_m),$$

and hence that

$$\|Sf-\lambda f\|^2 = 2/n.$$

In other words, the vector $f=f(\lambda, m, n)$ is an approximate eigenvector for S, with approximate eigenvalue λ and degree of approximation $\sqrt{2/n}$. Since

$$S^*f-\bar{\lambda}f = -\bar{\lambda}S^*(Sf-\lambda f),$$

it follows that f is, at the same time, an approximate eigenvector for S^*, with approximate eigenvalue $\bar{\lambda}$ and degree of approximation $\sqrt{2/n}$.

The preceding construction can be applied to each of the given numbers λ_k. Choose n_k so that

$$\Sigma_k(2/n_k) \leqq (\varepsilon/2)^2,$$

and choose m_k so that the index intervals $[m_k, m_k+n_k-1]$ are pairwise disjoint. If $f_k=f(\lambda_k; m_k, n_k)$, then $\{f_1, f_2, f_3, \dots\}$ is an orthonormal sequence such that $\Sigma_k\|Sf_k-\lambda_k f_k\|^2 \leqq (\varepsilon/2)^2$; $\Sigma_k\|S^*f_k-\bar{\lambda}_k f_k\|^2 \leqq (\varepsilon/2)^2$. Let M be the span of $\{f_1, f_2, f_3, \dots\}$, let P be the projection with range M, and let D be the diagonal operator defined on M by

$$Df_k = \lambda_k f_k, \quad k = 1, 2, 3, \dots .$$

Write the Hilbert space as $M\oplus M^\perp$, and, correspondingly, consider the matrices

$$S=\begin{pmatrix} X & Y \\ Z & E\end{pmatrix}, \quad P=\begin{pmatrix} 1 & 0 \\ 0 & 0\end{pmatrix}, \quad Q=\begin{pmatrix} D & 0 \\ 0 & 0\end{pmatrix}; \quad C=\begin{pmatrix} X-D & Y \\ Z & 0\end{pmatrix}.$$

Assertion: $X-D$, Y, and Z are Hilbert—Schmidt operators, and the sum of the squares of their Hilbert—Schmidt norms is not more than ε^2. Indeed:

$$\|(X-D)f_k\|^2+\|Zf_k\|^2 = \|(S-Q)f_k\|^2 = \|Sf_k-\lambda_k f_k\|^2 = 2/n_k,$$

and, similarly,

$$\|(X^*-D^*)f_k\|^2+\|Y^*f_k\|^2 = 2/n_k;$$

the proof of Lemma 2 is complete.

Remark. The control that the proof gives over the differences $Sf_k-\lambda_k f_k$ is strong enough to make it possible to put C into the trace class. Indeed: choose the n_k's so large that $\Sigma_k\sqrt{2/n_k}$ is small, and apply the lemma of Dunford and Schwartz [4, p. 1116] according to which $\Sigma_k\|Tf_k\|<\infty$, for an orthonormal basis $\{f_1, f_2, f_3, \dots\}$, implies that T is in the trace class. (The statement in [4] does not seem to be formulated for the right set of exponents. In any event, it is true and the proof is valid for the exponent $p=1$.)

Lemma 3. *If $\{a_n\}$ and $\{b_n\}$ are sequences with the same cluster set C in a compact metric space, if $a_n \in C$ and $b_n \in C$ for all n, and if $\varepsilon > 0$, then there exists a permutation π of the natural numbers such that $\Sigma_n d(a_n, b_{\pi n}) < \varepsilon$.*

It is convenient to have a word to describe the sequences that occur in this statement: call a sequence recurrent if each of its terms is a cluster point of it. (The "cluster set" of a sequence is, of course, the set of all cluster points. In this language a sequence is recurrent if it is included in its own cluster set.) Lemma 3 is a sharpened version of the von Neumann permutation theorem [7], which is used in the proof of the von Neumann converse of Weyl's theorem. The original version does not assume that the given sequences are recurrent, and cannot conclude that, after the permutation, the sum of distances is small. If, for instance, $a_1 = 1$, $a_n = 0$ for $n > 1$, and $b_n = 0$ for all n, then, clearly, there is no permutation π such that $d(a_n, b_{\pi n}) \leqq 1/2$ for all n. The trouble is not that the ranges of the sequences are different; if $b_1 = b_2 = 1$ and $b_n = 0$ for $n > 2$, the inequalities $d(a_n, b_{\pi n}) \leqq 1/2$ for all n can still not be achieved. The trouble is that the cluster sets (which, to be sure, are the same) do not contain all the terms; in the first example one of the sequences fails to be recurrent, and in the second example they both do.

Now for the proof of Lemma 3.

Write $\sigma(1) = 1$. Since $a_{\sigma(1)} \in C$, there exists an index $\tau(1)$ such that $d(a_{\sigma(1)}, b_{\tau(1)}) \leqq \varepsilon/2$. Let $\tau(2)$ be the smallest index distinct from $\tau(1)$ (so that, typically, $\tau(2)$ will be 1). Since $b_{\tau(2)} \in C$, there exists an index $\sigma(2)$ distinct from $\sigma(1)$ such that $d(a_{\sigma(2)}, b_{\tau(2)}) \leqq \varepsilon/4$. The preceding four sentences describe a two-step process that is now to be applied infinitely often. The second application will indicate how the general one is to be made. Let $\sigma(3)$ be the smallest index not contained in $\{\sigma(1), \sigma(2)\}$. Find $\tau(3)$ not contained in $\{\tau(1), \tau(2)\}$ so that $d(a_{\sigma(3)}, b_{\tau(3)}) \leqq \varepsilon/8$. Let $\tau(4)$ be the smallest index not contained in $\{\tau(1), \tau(2), \tau(3)\}$. Find $\sigma(4)$ not contained in $\{\sigma(1), \sigma(2), \sigma(3)\}$ so that $d(a_{\sigma(4)}, b_{\tau(4)}) \leqq \varepsilon/16$.

When, ultimately, $\sigma(n)$ and $\tau(n)$ are defined for all n, each of σ and τ is a permutation of the set of all natural numbers. Indeed: since the definition of $\sigma(n)$ guarantees that $\sigma(n)$ is not contained in $\{\sigma(1), \ldots, \sigma(n-1)\}$, $n > 1$, the mapping σ is one-to-one; the definition for odd values of n guarantees that every natural number is in the range of σ. The argument for τ is, of course, the same, except that "odd" has to be replaced by "even".

The result is a pair of permutations σ and τ such that $d(a_{\sigma(n)}, b_{\tau(n)}) \leqq \varepsilon/2^n$ for all n. If π is defined so that $\tau(n) = \pi(\sigma(n))$ for all n, i.e., if $\pi = \tau\sigma^{-1}$, then $\Sigma_n d(a_n, b_{\pi(n)}) = \Sigma_n d(a_{\sigma(n)}, b_{\tau(n)}) \leqq \varepsilon$.

My original statement of Lemma 3 had "$d(a_n, b_{\pi(n)}) \to 0$ and $d(a_n, b_{\pi(n)}) \leqq \varepsilon$ for all n" instead of "$\Sigma_n d(a_n, b_{\pi(n)}) \leqq \varepsilon$", and my proof of it was longer; the simplification is due to J. G. Stampfli.

To apply Lemma 3, I introduce a new concept: a normal operator is *essential* if its spectrum is the same as its essential spectrum.

Lemma 4. *If A and B are essential normal operators with the same spectrum, on a separable Hilbert space, and if $\varepsilon > 0$, then there exists a unitary operator U and a compact operator K such that $A = U^* BU + K$ and $\|K\| \leqq \varepsilon$.*

Lemma 4 is a sharpened version of the von Neumann converse of Weyl's theorem. The original version does not assume that the given operators are essential, and cannot conclude that, after the unitary equivalence, they are within ε of one another. VON NEUMANN [11] remarked that, in fact, if a single compact operator is excluded from the competition, the conclusion becomes false. The point of Lemma 4 is that for essential normal operators the compact operators that appear can be made to satisfy severe and useful size restrictions.

To prove Lemma 4, use Lemma 1 to write $A = D_A + C_A$ and $B = D_B + C_B$, where D_A and D_B are diagonal, with each diagonal entry in the common spectrum, and C_A and C_B are compact, with $\|C_A\| \leqq \varepsilon/3$, $\|C_B\| \leqq \varepsilon/3$. Since D_A and A have the same essential spectrum, and since the essential spectrum of D_A is the cluster set of the diagonal, it follows that that cluster set is the common spectrum of A and B. (This step uses the assumption that A is essential.) Similarly the cluster set of D_B is that common spectrum. By Lemma 3 there exists a unitary operator U (induced by a permutation) and a compact operator C such that $D_A = U^* D_B U + C$ and $\|C\| \leqq \varepsilon/3$. Consequence:

$$A = D_A + C_A = U^* D_B U + C + C_A = U^*(B - C_B) U + C + C_A =$$
$$= U^* BU - U^* C_B U + C + C_A;$$

since $-U^* C_B U + C + C_A$ is compact and has norm not greater than ε, the proof of Lemma 4 is complete.

Remark. In case A and B are such that C_A and C_B can be made to have small Hilbert—Schmidt norm (e.g., in case A and B are Hermitian or unitary), then K can be made to have small Hilbert—Schmidt norm; the perturbation C that Lemma 3 introduces belongs, in fact, to the trace class.

The statement of Lemma 4 does not include the von Neumann converse (for not necessarily essential operators) as a special case, but the proof of Lemma 4 is, in spirit, the same as that of the unmodified version; cf. [1], [10], [11]. The main difference is that the present proof uses the quantitative version (Lemma 3) of the von Neumann permutation theorem.

For some of the statements that follow it is convenient to introduce a shorthand notation: if A and B are operators and $\varepsilon > 0$, write

$$A \sim B \quad (\varepsilon)$$

in case there exists an operator B', unitarily equivalent to B, such that $A-B'$ has Hilbert—Schmidt norm not greater than ε. (The operators A and B need not even be defined on the same Hilbert space. The generality gained thereby is shallow but useful.)

Lemma 5. *If S is a shift of multiplicity n $(1 \leqq n \leqq \infty)$ (unilateral or bilateral), if U is a unitary operator on a separable Hilbert space, and if $\varepsilon > 0$, then*

$$S \sim U \oplus S \quad (\varepsilon).$$

For the proof, let $\{\lambda_1, \lambda_2, \lambda_3, \ldots\}$ be a sequence (of complex numbers of modulus 1) whose closure is dense in the spectrum of U, in which each term occurs infinitely often. Apply Lemma 2 to write

$$S \sim D \oplus E \quad (\varepsilon/4), \tag{1}$$

where D is a diagonal operator with eigenvalues $\lambda_1, \lambda_2, \lambda_3, \ldots$. (The unitary equivalence is, in fact, effected by the identity operator in this case, but nothing is lost by forgetting that.) In Lemma 2, to be sure, the shift was assumed to have multiplicity 1. The lemma is, however, applicable to shifts of all non-zero multiplicities; all that needs to be done is to break off a shift of multiplicity 1 as a direct summand, apply Lemma 2 as is, and then glue the fracture together again.

Let U^∞ be the direct sum of countably infinitely many copies of U; observe that U^∞ is unitary (and hence, in particular, normal) and that U^∞ is essential. Since D and $U^\infty \oplus D$ are essential unitary operators with the same spectrum, the remark following the proof of Lemma 4 shows that

$$D \sim U^\infty \oplus D \quad (\varepsilon/4). \tag{2}$$

Substitute (2) into (1) to get

$$S \sim U^\infty \oplus D \oplus E \quad (\varepsilon/2).$$

It follows that

$$U \oplus S \sim U \oplus U^\infty \oplus D \oplus E \quad (\varepsilon/2). \tag{3}$$

Since $U \oplus U^\infty$ is unitarily equivalent to U^∞, (3) implies

$$U \oplus S \sim U^\infty \oplus D \oplus E \quad (\varepsilon/2)$$

and hence, by (2)

$$U \oplus S \sim D \oplus E \quad (3\varepsilon/4). \tag{4}$$

Use (1) to replace the right side of (4) by S, and conclude that

$$U \oplus S \sim S \quad (\varepsilon).$$

The proof of Lemma 5 is complete.

Proof of the theorem. Suppose that V is an isometry of co-rank n $(1 \leqq n \leqq \infty)$ on a separable Hilbert space, and suppose that $\varepsilon > 0$. Write V as $U \oplus S$, where U

is unitary and S is a unilateral shift (with, of course, multiplicity n) [6, Problem 118]. By Lemma 5

$$V \sim S \quad (\varepsilon).$$

Since an operator unitarily equivalent to a unilateral shift is a unilateral shift, this proves that in every ε neighborhood of V there is a unilateral shift (necessarily of the same co-rank as V), and the first half of the theorem follows.

The second half is proved similarly. Suppose that U is a unitary operator whose spectrum is the entire unit circle, and suppose that $\varepsilon > 0$. By Lemma 5

$$U \oplus S \sim S \quad (\varepsilon/2),$$

where S is a bilateral shift of multiplicity n. Since U and $U \oplus S$ are essential unitary operators with the same spectrum (here is where the hypothesis about the spectrum of U is used), it follows from the remark following the proof of Lemma 4 that

$$U \sim U \oplus S \quad (\varepsilon/2).$$

Consequence:

$$U \sim S \quad (\varepsilon),$$

and the proof is completed as in the unilateral case.

Scholium. *On a separable Hilbert space every isometry of non-zero co-rank is the sum of a pure isometry and an operator of arbitrarily small Hilbert—Schmidt norm.*

Except for the description of the size of the perturbation, this is the original version of the Douglas result mentioned after the statement of the theorem.

Experience shows that norm approximation theorems are likely to be difficult but worth the trouble; they give useful analytic insights into the behavior of operators. Strong and weak approximation theorems are usually easier to prove, but harder to find applications for. A comparison of the theorem proved above and the proposition below indicates that for approximation by shifts the customary situation prevails.

In what follows it is convenient to use the word "shift" ambiguously. A true statement and a valid proof result if it is interpreted consistently as either "unilateral shift" or "bilateral shift".

Proposition. *On a separable infinite-dimensional Hilbert space the strong closure of the set of shifts of multiplicity* 1 *is the set of all isometries; the weak closure of the set of shifts of multiplicity* 1 *is the set of all contractions.*

For the proof, consider first an arbitrary operator A on the given Hilbert space H, and a direct sum of the form $A \oplus B$ on $H \oplus H$. Assertion 1: if $f_1, \ldots, f_n$ are in H, then there exists an operator on H unitarily equivalent to $A \oplus B$ that agrees with

A on each f_j. To prove that, let V be an isometry from H onto $H\oplus H$ such that if f is in the (finite-dimensional) subspace spanned by $f_1, \ldots, f_n$, $Af_1, \ldots, Af_n$, then $Vf=[f, 0]$. (Here is where the infinite-dimensionality of H is used.) It follows that $V^*(A\oplus B)Vf_j = V^*(A\oplus B)[f_j, 0] = V^*[Af_j, 0] = Af_j$ for $j=1, \ldots, n$.

Suppose now that U is an arbitrary unitary operator on H. Assertion 2: every strong neighborhood of U contains a shift of multiplicity 1. To prove this, consider a basic strong neighborhood of U, consisting of all operators T such that $\|Uf_j-Tf_j\| < \varepsilon$, $j=1, \ldots, n$, where $f_1, \ldots, f_n$ are unit vectors in H and $\varepsilon>0$. If S is a shift of multiplicity 1, then, by Assertion 1, there exists an operator unitarily equivalent to $U\oplus S$ that agrees with U on each f_j. Since, by Lemma 5, $U\oplus S\sim S$ (ε), it follows that some operator unitarily equivalent to S differs from U by less than ε on each f_j; this implies Assertion 2.

The preceding two paragraphs imply that the strong closure of the set of shifts of multiplicity 1 contains all unitary operators, and from this all else follows. Indeed, the *strong* closure of the set of unitary operators is known to be the set of all isometries [8, p. 892], and the *weak* closure of the set of unitary operators is known to be the set of all contractions [5, p. 128].

Problem. What are the answers to the corresponding questions for weighted shifts?

References

[1] S. K. Berberian, The Weyl spectrum of an operator, *Indiana U. Math. J.*, **20** (1970), 529—544.

[2] I. D. Berg, An extension of the Weyl—von Neumann theorem to normal operators, *Trans. Amer. Math. Soc.*, **160** (1971), 365—371.

[3] L. G. Brown, R. G. Douglas, and P. A. Fillmore. Unitary equivalence modulo compact operators and extensions of C^*-algebras. (Preprint, October 1972).

[4] N. Dunford and J. T. Schwartz, *Linear operators*. Part II, Interscience (New York, 1963).

[5] P. R. Halmos, Normal dilations and extensions of operators, *Summa Brasil. Math.*, **2** (1950), 125—134.

[6] P. R. Halmos, *A Hilbert space problem book*, Van Nostrand (Princeton, 1967).

[7] P. R. Halmos, Permutations of sequences and the Schröder—Bernstein theorem, *Proc. Amer. Math. Soc.*, **19** (1968), 509—510.

[8] P. R. Halmos, Ten problems in Hilbert space, *Bull. Amer. Math. Soc.*, **76** (1970), 887—933.

[9] P. R. Halmos, Continuous functions of Hermitian operators, *Proc. Amer. Math. Soc.*, **31** (1972), 130—132.

[10] W. Sikonia, The von Neumann converse of Weyl's theorem, *Indiana U. Math. J.*, **21** (1971), 121—123.

[11] J. von Neumann, *Charakterisierung des Spektrums eines Integraloperators*, Hermann (Paris, 1935).

INDIANA UNIVERSITY
BLOOMINGTON, INDIANA 47401

(Received May 27, 1972)

On a class of adjoint functional equations

By EINAR HILLE in La Jolla (California, U.S.A.)*)

To Béla Szőkefalvi-Nagy on his sixtieth birthday

1. Problem. Let A and B denote *vector mean values* in R^m, this term to be defined below. Consider the two *adjoint functional equations*

$$f[A(s,t)] = B[f(s), f(t)], \tag{1.1}$$

$$g[B(a,b)] = A[g(a), g(b)]. \tag{1.2}$$

A study of these equations is proposed involving intrinsic properties of the solutions, classification of the latter with continuous solutions as one type and boundary affine solutions as the other, spread of continuity, and construction of solutions by a process of successive interpolation. The equations are *inverses* of each other in the sense that there exist solutions f_0 and g_0 such that

$$f_0[g_0(a)] = a, \quad g_0[f_0(s)] = s \tag{1.3}$$

in the domains of definition of g_0 and f_0.

In the theory of vector meanvalues the case

$$B(a,b) = \tfrac{1}{2}(a+b) \tag{1.4}$$

is basic and it plays an important role also in this paper where we are trying to extend the results obtained in this special case (see HILLE [2]) to the more general one. Some additional material for the special case is to be found in Section 11 below.

2. On vector meanvalues. In [2] the author based the discussion of vector meanvalues on a system of postulates analogous to those of A. N. KOLMOGOROV and M. NAGUMOV in the linear case.

The discussion involves various consequences of the natural partial ordering of R^m. We write

$$x = (x_1, x_2, \ldots, x_m), \quad y = (y_1, y_2, \ldots, x_m)$$

*) Research supported in part by NSF Grant GU-2582.

and define

$$x \leqq y \tag{2.1}$$

iff

$$x_j \leqq y_j, \qquad j = 1, 2, \ldots, m. \tag{2.2}$$

Should "$<$" hold here for all j, we write $x<y$. For a finite set of vectors x_p in R^m, distinct or not, we write

$$u = (u_1, u_2, \ldots, u_m) = \inf V \tag{2.3}$$

for the vector whose jth coordinate is the infimum of the jth coordinates of the vectors x_p of V. Similarly we define

$$v = (v_1, v_2, \ldots, v_m) = \sup V. \tag{2.4}$$

The set of all vectors $x=(x_1, x_2, \ldots, x_m)$ such that

$$u_j \leqq x_j \leqq v_j, \qquad j = 1, 2, \ldots, m, \tag{2.5}$$

is called the *closed cellular hull* $C[V]$ of V. If in all the inequalities where $u_j<v_j$ we replace "$\leqq$" by "$<$" we obtain a subset $C^0[V]$ of $C[V]$ known as the *open cellular hull* of $C[V]$. This is not necessarily an open set in the topology of R^m, but, unless all the x_p's are equal, there is a subspace of lower dimension in which $C^0[V]$ is open. These concepts are due to J. B. MILLER [3] and have been further explored by C. T. Ng [4, 5, 6].

The vectors to be admitted in forming meanvalues will be restricted to an open convex set G in R^m. G is supposed to contain the closed pyramid Π defined by

$$x_j \geqq 0, \quad j = 1, 2, \ldots, m, \quad x_1 + x_2 + \cdots + x_m \leqq 1. \tag{2.6}$$

We can now formulate the postulates:

M_1. For each finite set V of vectors $x_1, x_2, \ldots, x_n$ in G, not necessarily distinct, there exists a meanvalue $M(V)=M(x_1, x_2, \ldots, x_n)$, a vector in G.

M_2. M is a symmetric continuous function, strictly increasing in each of its arguments.

M_{21}. $M(V)\subset C^0[V]$.

M_{22}. For t fixed and s variable, both in G, the mapping $t\to M(s, t)$ is open, injective and continuous, uniformly with respect to t on compact subsets of G.

M_3. $M(x, x, \ldots, x)=x$.

M_4. Let $1<k<n$ and set $M(x_1, x_2, \ldots, x_k)=y$. Then

$$M(x_1, \ldots, x_k, x_{k+1}, \ldots, x_n)=M(y, \ldots, y, x_{k+1}, \ldots, x_n)$$

where y is repeated k times.

The meanvalues A and B in equations (1. 1) and (1. 2) are supposed to satisfy these conditions. We write D for G in the first case, E in the second; they may be distinct or identical.

3. Some properties of meanvalues. As consequences of M_3 and M_4 we note

Lemma 1. *The meanvalue of k copies of the set V is the same as of one copy*

$$M(kV) = M(V). \tag{3. 1}$$

Lemma 2. *We have*

$$M[M(V_1), M(V_2)] = M(V_1 \cup V_2). \tag{3. 2}$$

Here it should be noted that each vector figures with its proper multiplicity: if x occurs k_1 times in V_1 and k_2 times in V_2, then it figures k_1+k_2 times in $V_1 \cup V_2$.

Lemma 3. *From a set V of n vectors, select k vectors, $k<n$, and form the meanvalue. Repeat this for the $N = \binom{n}{k}$ choices of k objects from a set of n. Let V_k be the set of the corresponding N meanvalues. Then*

$$M(V_k) = M(V). \tag{3. 3}$$

Proof. We consider an enlarged set V^* made up of

$$\binom{n-1}{k-1}$$

copies of V. By Lemma 1, $M(V^*)=M(V)$. Here the set V^* has kN vectors which may be arranged into N sets of k vectors each. This is to be done in such a manner that the elements in the pth set are precisely those k vectors chosen in the pth selection. Let y_p be the meanvalue of these vectors. In forming $M(V^*)$ we can replace each of the vectors in the pth set by y_p using M_4. We repeat this for each p so that

$$M(V^*) = M(y_1, \ldots, y_1, \ldots, y_N, \ldots, y_N)$$

where each y_j figures k times. Using Lemma 1 again we can contract the last expression to $M(y_1, \ldots, y_N)=M(V_k)$ as asserted.

Corollary. *For each $k<n$*

$$M(V) \subset C^0[V_k]. \tag{3. 4}$$

A simple argument shows that

$$C^0[V] \supset C^0[V_2] \supset \cdots \supset C^0[V_k] \tag{3. 5}$$

which expresses that meanvalues are *variation-reducing*.

Lemma 4. *We have*

$$\|M(s,t)-\tfrac{1}{2}(s+t)\| \leqq \tfrac{1}{2}\sqrt{3}\,\|s-t\|. \tag{3.6}$$

Proof. The coordinates of the vectors M, s and t satisfy the inequalities

$$\inf(s_j, t_j)-s_j \leqq M_j-s_j \leqq \sup(s_j, t_j)-s_j$$

for $j=1, 2, \ldots, m$. This implies that

$$|M_j-s_j| \leqq |t_j-s_j| \tag{3.7}$$

for all j so that

$$\|M(s,t)-s\| \leqq \|s-t\| \tag{3.8}$$

and by symmetry

$$\|M(s,t)-t\| \leqq \|s-t\|. \tag{3.9}$$

The geometric meaning of these two inequalities is that $M(s, t)$ lies in the domain common to two spheres, one with center at s, the other at t having the same radius $\|s-t\|$. The inequality (3. 6) follows. Here equality can hold iff $s=t$ in which case, of course, equality holds for all j in (3. 7).

The main result of [2] can be expressed as follows.

Theorem M. *The equation*

$$h[M(s,t)] = \tfrac{1}{2}[h(s)+h(t)] \tag{3.10}$$

with initial conditions

$$h(0)=0, \quad h(u_1)=u_1, \ldots, h(u_m)=u_m \tag{3.11}$$

where the u's are the unit vectors, $u_j=(\delta_{jk})$, has a unique solution h which is one-to-one. In terms of this solution

$$h[M(s_1, s, \ldots, s_n)] = \frac{1}{n}\sum_{1}^{n} h(s_j) \tag{3.12}$$

for all n and all s_j in G.

This representation leads to further properties of the meanvalues. It is clear that

$$M(s,u) \neq M(s,v) \text{ if } u \neq v. \tag{3.13}$$

A more general form of this inequality is

Lemma 5. *If $u=M(u_1, \ldots, u_p)$, $v=M(v_1, \ldots, v_p)$ and $u\neq v$, then*

$$M(s_1, \ldots, s_k, u_1, \ldots, u_p) \neq M(s_1, \ldots, s_k, v_1, \ldots, v_p). \tag{3.14}$$

Proof. We use the representation of M given by (3. 12). Thus

$$(k+p)h[M(s_1, \ldots, u_p)] = \sum_1^k h(s_j) + \sum_1^p h(u_n) = \sum_1^k h(s_j) + ph(u) \neq$$

$$\neq \sum_1^k h(s_j) + ph(v) = (k+p)h[M(s_1, \ldots, v_p)].$$

Since the mapping $s \to h(s)$ is one-to-one, (3. 14) follows.

A subspace of R^m is said to be *principal* if it is obtained by equating k of the coordinates to zero, say $x_j=0$ for $j=n_1, n_2, \ldots, n_k$ where the n's are fixed.

Lemma 6. *If a set V of vectors belongs to a principal linear subspace of R^m, then $C^0[V]$ belongs to the same subspace.*

Proof by inspection.

Repeated averaging with one entry fixed leads to this entry:

Lemma 7. *Let x and y be vectors in R^m and form the sequence*

$$x_n = M(x_{n-1}, y), \qquad n = 2, 3, \ldots, x_1 = x. \tag{3.15}$$

Then $\lim x_n = x_0$ *exists and* $x_0 = y$.

Proof. Let z_{nj} denote the j^{th} coordinate of $x_n - y$. Then for fixed j the sequence $\{z_{nj}\}$ is monotone, all the members of the sequence have the same sign namely that of the first term. The sequence is strictly increasing, zero, or strictly decreasing according as z_{1j} is negative, zero, or positive. All this follows from the definition of $M(s, t)$. Hence $\lim (x_n - y) = x_0 - y$ exists and by the continuity of M, we have $x_0 = M(x_0, y)$ so that $x_0 = y$ as asserted.

Let X and Y be subsets of G and set

$$M(X, Y) = \{u; u = M(s, t); s \in X, t \in Y\}. \tag{3.16}$$

If $Y=X$, then $X \subset M(X, X)$ and the inclusion is proper unless X reduces to a single vector.

4. Boundedness and singularities. For the case of the general system (1. 1) and (1. 2) we need some notions which become self-evident in the special case where B is the arithmetic mean.

The convex set G is bounded by a convex hypersurface $K = \vartheta G$. If G is not bounded, then part or all of ϑG may be infinitary. It is understood and admitted that some boundary elements may have coordinates equal to $+\infty$ or $-\infty$.

If $\{s_n\} \subset G$ and converges to an element of ϑG, we call $\{s_n\}$ a *boundary sequence.*

M is *boundary preserving* if for all $t \in G$ the sequence $\{M(s_n, t)\}$ is a boundary sequence whenever $\{s_n\}$ has this property.

As an illustration take $m=1$ and

$M(s,t)$	G	∂G
$\frac{1}{2}(s+t)$	$(-\infty, +\infty)$	$s=-\infty, +\infty;$
$\arctan[\frac{1}{2}(\tan s+\tan t)]$	$(-\frac{1}{2}\pi, \frac{1}{2}\pi)$	$s=-\frac{1}{2}\pi, +\frac{1}{2}\pi;$
$\arcsin[\frac{1}{2}(\sin s+\sin t)]$	$(-\frac{1}{2}\pi, \frac{1}{2}\pi)$	$s=-\frac{1}{2}\pi, +\frac{1}{2}\pi.$

Here the first two means are boundary preserving but the third is not.

For the special case (1. 4) local boundedness implies global boundedness. In the general case affinity for the boundary plays a similar role. This leads to the notion of a singular set.

A point $s=s_0 \in D$ is *singular* with respect to a solution $s \to f(s)$ of (1. 1) iff there exists a sequence $\{s_n\} \in D$ converging to s_0 such that $\{f(s_n)\}$ is a boundary sequence with respect to ϑE. *The set of all singular points is the singular set of f.*

5. Intern transformations. The equations (1. 1) and (1. 2) are special cases of functional equations of the form

$$f[F(x,y)] = H[f(x), (y), x, y]. \tag{5. 1}$$

Here F and H are given and f is to be found. The study of this class was initiated by J. Aczél [1] in 1964 for x and y real variables. F is assumed to be an *intern transformation* in the sense that

$$x<y \text{ implies } x<F(x,y)<y. \tag{5. 2}$$

Further F is continuous and H is assumed to be injective with respect to either the first or the second argument. Under these assumptions Aczél could prove that the equation has at most one continuous solution satisfying a given two-point condition

$$f(x_0)=y_0, \quad f(x_1)=y_1. \tag{5. 3}$$

Note that this is a uniqueness theorem, nothing is said about the existence of any solution.

Extensions to R^m have been proved by J. B. Miller [3] and C. T. Ng [4, 5]. Here there are two concepts to be generalized: (i) the internity and (ii) the initial conditions. *Cellular internity* in the sense that

$$F(x,y) \subset C^0(x;y) \tag{5. 4}$$

turns out to be a suitable generalization of the first notion. As to the initial conditions it is fairly clear that $m+1$ conditions

$$f(x_j)=y_j, \qquad j=0,1,2,\ldots,m, \tag{5. 5}$$

are required, but it is not obvious that the x's are subject to restrictions which are satisfied if the vectors $x_k - x_0$, $k=1, 2, \dots, m$ are linearly independent. We note that the vertices of Π are admissible choices for the x's. Assuming further F is continuous and that H is injective with respect to the first or the second argument, Ng [5] could prove that equation (5. 1) has at most one continuous solution satisfying an admissible $(m+1)$-point condition (5. 5).

This result applies in particular to equation (1. 1) and (1. 2). For by M_{21} both A and B are continuous and cellularly intern and by M_{22} both right hand sides are injective with respect to both arguments.

The case where $A=B$ is of some interest.

Theorem 1. *The equation*

$$h[A(s,t)] = A[h(s), h(t)] \tag{5.6}$$

has a unique solution $s \to h(s)$ which is continuous and leaves the vertices of Π invariant, namely $h(s) \equiv s$.

Proof by inspection. It should be noted that the solution is independent of A. This leads to the following composition theorem.

Theorem 2. *Suppose that* (1. 1) *has the continuous solution f_0 of domain D and range E which satisfies the initial conditions* (3. 11). *Suppose that g_0 is the continuous solution of* (1. 2) *with domain E and range D, satisfying the same conditions. Then $g_0[f_0(s)]$ is defined in D and equals s identically, $f_0[g_0(a)]$ is defined in E and equals a identically.*

Proof. The existence of the composite functions is obvious. In equation (1. 2) we substitute $a=f_0(s)$, $b=f_0(t)$ and obtain

$$A\{g_0[f_0(s)],\quad g_0[f_0(t)]\} = g_0\{B[f_0(s), f_0(t)]\} = g_0\{f_0[A(s,t)]\}.$$

This implies that $s \to h(s) = g_0[f_0(s)]$ is a continuous solution of (5. 6) which leaves the vertices of Π invariant. Hence $h(s) \equiv s$ as asserted. The case of $f_0[g_0(a)]$ is handled in the same manner.

6. The dichotomy. We shall prove

Theorem 3. *If f is a solution of* (1. 1) *with domain D and range E, if B is boundary preserving with respect to E and if the singular set U of f in D is not all of D, then $U=\emptyset$.*

Proof. Since D is open and connected, it is enough to prove that U is both open and closed in D for then D is either D itself or void. The first alternative being excluded by assumption, the second must hold. Suppose $s_0 \in U$. Then there is a sequence $\{s_n\}$ converging to s_0 such that $\{f(s_n)\}$ is a boundary sequence. Let $t \in D$ be arbitrary and consider the sequence $\{A(s_n, t)\}$ which converges to $A(s_0, t)$ when $n \to \infty$. Now

(6. 1) $$f[A(s_n, t)] = B[f(s_n), f(t)], \qquad n = 1, 2, 3, \ldots$$

and here the right members form a boundary sequence since B is boundary preserving. It follows that $\{f[A(s_n, t)]\}$ is a boundary sequence so that $A(s_0, t) \in U$. In particular, there exists a full neighborhood of $t = s_0$ which belongs to U so that U is open in D.

On the other hand, if $\{x_n\} \subset U$ and $x_n \to x_0 \in D$, there exist sequences $\{x_{nk}\}$, $n = 1, 2, 3, \ldots$, such that each $\{f(x_{nk}): k = 1, 2, 3, \ldots\}$ is a boundary sequence. It is then seen that the diagonal sequence $\{f(x_{nn})\}$ either is a boundary sequence or contains one and here $x_{nn} \to x_0$ so that $x_0 \in U$. Thus U is also closed and, as observed above, this implies that $U = \emptyset$.

With the appropriate changes of assumptions a similar result holds for equation (1. 2).

We note that for these equations the singular set is either void or coincides with the domain of definition of the solution. In the special case (1. 4) the alternatives are a solution everywhere unbounded or one bounded on compact sets which implies continuity. In the general case where A and B are arbitrary (boundary preserving) meanvalues, we note that if the singular set is void, then the solution is necessarily bounded on compact subsets. The proof given above uses the assumption of boundary preserving but it is not obvious that this is a necessary condition for the validity of the result. For the special case the idea of the proof is due to C. T. Ng.

7. Boundedness and continuity. We shall prove

Theorem 4. *If a solution of* (1. 1) *is continuous at a single point of D then it is continuous everywhere in D.*

Proof. Suppose that $s \to f(s)$ is continuous at $s = s_0$. Then for $\|h\|$ small and t arbitrary in D,

$$f[A(s_0 + h, t)] = B[f(s_0 + h), f(t)] \to B[f(s_0), f(t)]$$

while $A(s_0 + h, t) \to A(s_0, t)$. This shows that f is continuous at the point $A(s_0, t)$ for all t in D. These points form an open set $D_1 = A(s_0, D)$, a neighborhood of $s = s_0$. If $D_1 = D$ we are through. If not, we define an expanding sequence of sets $\{D_n\}$ where in the notation of (3. 16)

(7. 1) $$D_n = A(D_{n-1}, D), \qquad n = 2, 3, \ldots.$$

Here f is continuous in each of the sets D_n and hence also in their union D_0 and

(7. 2) $$D_0 = A(D_0, D).$$

by (7. 1). If t were in D but not in D_0, then we would have $A(s, t) \in D_0$ for all s in D, in particular for $s=t$. Since $A(t, t)=t$ this contradiction shows that $D_0=D$ so that f is continuous everywhere in D.

Corollary 1. *A solution of* (1. 1) *is continuous either everywhere in D or nowhere.*

The same argument applies to equation (1. 2).

In the construction of solutions given below the interpolation process leads to a solution which is not defined in all of D but merely in a dense subset T of Π. It is important for us to note that such a solution already has continuity properties provided T satisfies certain conditions. The proof given for Theorem 4 leads to

Corollary 2. *Suppose that a solution f of* (1. 1) *is defined in a subset T of D with the following properties* (i) *T is dense in Π and* (ii) *$A(T, T) \subset T$. Suppose that f is continuous at a single point s_0 of T for approach to s_0 in T, then f is continuous everywhere in T. In the case of equation* (1. 2) *replace* (ii) *by* (ii') *$B(T, T) \subset T$.*

Here $M(T, T) \subset T$ means that

$$\{x: x = M(s, t), s \in T, t \in T\} \subset T. \tag{7.3}$$

Proof. The argument used above carries over if s_0, s_0+h and t are confined to T which replaces D throughout so that

$$T_1 = A(s_0, T), \quad T_n = A(T_{n-1}, T,) \quad T_0 = \bigcup_n T_n. \tag{7.4}$$

Theorem 5. *A solution of* (1. 1) *or of* (1. 2) *which is zero at the origin is continuous there and hence continuous everywhere in D. The same conclusion with D replaced by T holds if the domain of definition is a set T of the type defined above.*

Proof. Take, for instance, (1. 1) with $s \to f(s)$ as the solution with D as domain of definition and $f(0)=0$. Then for s in D we have $f[A(0, s)]=B[f(0), f(s)]=B[0, f(s)]$. We introduce two sequences of vectors $\{s_n\}$ and $\{t_n\}$ where

$$s_{n+1} = A(0, s_n), \quad s_1 = s, \quad t_{n+1} = f(s_n). \tag{7.5}$$

Here $\lim s_n=0$ by Lemma 7. Since $t_{n+1}=B(0, t_n)$, a second appeal to Lemma 7 gives $\lim t_n=0$, so that $\lim f(s_n)=0=f(0)$. This holds for every s in D so that $f(s)$ can approach no other limit than 0 as $s \to 0$. Hence f is continuous at $s=0$ and by Theorem 4 this means continuity everywhere in D. The modifications of the proof that become necessary if D is replaced by T are obvious.

Equation (1. 2) is handled in the same manner.

The arithmetic mean $\frac{1}{2}(s+t)$ is not merely a special case of the means considered here but in a certain sense it is asymptotic to the general mean. For in important cases the right member of (3. 6) can be lowered to the second power so that

$$M(s,t) = \tfrac{1}{2}(s+t) + O(\|s-t\|^2) \tag{7. 6}$$

as $t \to s$. See further below, Theorem 12.

Under these circumstances it makes sense to examine the consequences of imposing temporarily a further restriction on M, namely

M_5. There exists a constant k, $\frac{1}{2} \leqq k < 1$, such that

$$\|M(s,t) - s\| \leqq k\|s-t\|, \qquad \forall s, t \in G. \tag{7. 7}$$

It should be noted that a value of $k < \frac{1}{2}$ is not admissible for this would give $\|s-t\| \leqq 2k\|s-t\|$ which can hold only for $t=s$ if $2k<1$. On the other hand, $k=1$ is no restriction by (3. 8).

We have now

Theorem 6. *If the meanvalue B satisfies* (7. 7), *then a locally bounded solution of* (1. 1) *is locally continuous.*

Proof. Suppose that in some neighborhood of $s=s_0$ we have $\|f(s)\| \leqq F$ and set

$$\limsup_{\|h\| \to 0} \|f(s_0+h) - f(s_0)\| = \delta(s_0). \tag{7. 8}$$

Then $f[A(s_0+h, s_0)] - f(s_0) = B[f(s_0+h), f(s_0)] - f(s_0)$. By virtue of M_5 this gives

$$\|f[A(s_0+h, s_0)] - f(s_0)\| \leqq k\|f(s_0+h) - f(s_0)\|. \tag{7. 9}$$

Here we let $\|h\| \to 0$. The superior limit of the right hand side is $k\delta(s_0)$ which is non-negative and at most equal to $2kF$. In the left member $h \to A(s_0+h, s_0)$ maps a sphere with center at $h=0$ onto a full neighborhood of $s=s_0$. It follows that the superior limit of the left member is $\delta(s_0)$ when $\|h\| \to 0$ so we have $\delta(s_0) \leqq k\delta(s_0)$. This implies $\delta(s_0)=0$ and makes f continuous at $s=s_0$ and hence everywhere.

The same argument applies to (1. 2) if A satisfies M_5. Moreover, we can replace D by a set T with but little change in the argument.

8. Primary mappings. So far we have operated under the assumption that solutions of our equations do exist. This will now be proved. More precisely, we shall prove that solutions satisfying the initial conditions (3. 11) can be constructed in the basic pyramid Π. These are the so called *fundamental solutions* and since they are continuous at the origin they are continuous wherever they are defined. We concentrate on equation (1. 1) but the same method applies to (1. 2). Actually we construct solutions of the auxiliary equations

(8. 1) $$h[\tfrac{1}{2}(a+b)] = A[h(a), h(b)],$$

(8. 2) $$k[\tfrac{1}{2}(a+b)] = B[k(a), k(b)],$$

with the aid of which we obtain the solutions of (1. 1) and (1. 2). These equations involve the arithmetic mean. We set

(8. 3) $$C(a, b) = \tfrac{1}{2}(a+b).$$

Let R be the set of points in Π with dyadic rational coordinates. It is clear that R is dense in Π and we have also

(8. 4) $$R \supset C(R, R)$$

so that R is a set T in the sense defined above. Next we construct two sets S and Y by applying the operations A and B repeatedly to R. It simplifies matters to construct these sets step by step. Let $R_0 = S_0 = Y_0$ be the set of $(m+1)$ vertices of Π and define the sets $\{R_n\}$, $\{S_n\}$ and $\{Y_n\}$ recursively as follows

(8. 5) $$R_n = C(R_{n-1}, R_{n-1}), \quad S_n = A(S_{n-1}, S_{n-1}), \quad Y_n = B(Y_{n-1}, Y_{n-1}).$$

These are sequences of expanding bounded sets and clearly tend to limits which are the sets R, S, T.

None of these sets is closed but we shall prove

Lemma 8. *The closures of the sets R, S, Y are connected, even simply connected.*

Proof. For R this is trivial. We shall give the argument for S, the same proof applies to Y. If a and b are two points of S, we set $T_0 = a \cup b$ and $T_n = A(T_{n-1}, T_{n-1})$ and $T = \bigcup_0^\infty T_n$. Then $\bar{T}$ is a continuous arc joining a with b and lies in $\bar{S}$. Since S is closed under the operation A, it is clear that $T \subset S$. By Lemma 7 every point of T is a limit point of T and since A is uniformly continuous, it follows that $\bar{T}$ is a continuous arc.

To prove that $\bar{S}$ is simply-connected we have to show that the complement of $\bar{S}$ is connected. If this is not so, then there is a point z_0 in the complement of $\bar{S}$ that cannot be joined to a point in the infinitary component of the complement of $\bar{S}$ by a Jordan arc not having points in common with $\bar{S}$. We can then "box in" z_0, that is we can find a parallelopiped π such that: (i) The faces of π are planes $x_j = \alpha_j$, $x_j = \beta_j$, $\alpha_j < \beta_j$, $j = 1, 2, \ldots, m$. (ii) There is no point of $\bar{S}$ in the interior of π. (iii) For each j there is a point $P_{j1} \in \bar{S}$ on the face $x_j = \alpha_j$ and a point $P_{j2} \in \bar{S}$ on the face

$x_j=\beta_j$. Since these points may fall on edges or vertices of π the total number of distinct points does not have to be $2m$ in number, but there are at least two. To see that this extreme case can happen, note that each of the endpoints of a diagonal of π lies on m distinct faces and that these faces have nothing in common with the faces associated with the other endpoint. If this case should be present, we obtain a contradiction right away for the A-mean of the endpoints is an interior point of π and belongs to $\bar{S}$ if the endpoints do. In the general case we may have to perform as many as $(m+1)$ A-operations before a contradiction results. To see this we form $P_j=A(P_{j1}, P_{j2})$. This is a point of $\bar{S}$, it may conceivably fall on an edge or a face of π but in any case its j^{th} coordinate belongs to the open interval (α_j, β_j) by virtue of the properties of $C^0(x, y)$. If one of the points P_j belongs to the interior of π, we are through. If not, one more averaging will lead to the desired contradiction. We form $P=A(P_1, P_2, \ldots, P_m)$. Again this is a point of $\bar{S}$ but now for each j the jth coordinate belongs to the open interval (α_j, β_j) and this forces P to be an interior point of π thus violating assumption (ii). This contradiction shows that $\bar{S}$ is simply-connected.

A closer examination of the generation of the sets R, S, Y is now in order. An element of R is obtained by applying the arithmetic means, the operation C, to some of the elements of the set R_0; i.e., the vertices of Π, say

(8.6) $\quad k_0$ copies of 0, k_1 copies of $u_1, \ldots, k_m$ copies of u_m,

where $\sum_{j=0}^{m} k_j=2^n$ for some positive integer n. Denote this aggregate of 2^n vectors by V. Then

$$a = C(V). \tag{8.7}$$

To this element of R correspond vectors of S and Y

$$s = A(V) = p(a), \quad y = B(V) = q(a). \tag{8.8}$$

Our first object is to study these mappings of R into S and Y.

Lemma 9. *The mappings $R \xrightarrow{p} S$ and $R \xrightarrow{q} Y$ are one-to-one.*

Proof. Take two distinct points a and b of R and suppose that

$$a = C(V_1), \quad b = C(V_2) \tag{8.10}$$

in the notation of (8. 6) and (8. 7). Since we can always enlarge our vector sets using Lemma 1 we may assume that V_1 and V_2 both contain 2^n vectors. The sets V_1 and V_2 are distinct, but may have elements in common. Let V_0 be the set of common vectors; if u_j occurs μ_j times in V_1 and ν_j times in V_2 then it occurs $\min(\mu_j, \nu_j)$ in V_0. Set

$$V_1 = V_0 \cup U_1, \quad V_2 = V_0 \cup U_2. \tag{8.11}$$

Then U_1 and U_2 have the same number of elements but have no vectors in common. Moreover, the linear manifolds spanned by U_1 and U_2 have only the zero element in common. In any case Lemma 6 shows that $C^0[U_1] \cap C^0[U_2] = \emptyset$ and this implies that $A(U_1) \neq A(U_2)$. By Lemma 2

$$p(a) = A(V_1) = A(V_0 \cup U_1) = A[A(V_0), A(U_1)],$$

$$p(b) = A(V_2) = A(V_0 \cup U_2) = A[A(V_0), A(U_2)],$$

and by (3. 13) $p(a) \neq p(b)$. In the same manner it is shown that $q(a) \neq q(b)$.

Theorem 7. *The fundamental solution of* (8. 1) *is* $h(a) = p(a)$ *and the fundamental solution of* (8. 2) *is* $k(a) = q(a)$.

Proof. Consider $A[p(a), p(b)]$ with $p(a) = A(V_1)$, $p(b) = A(V_2)$. Then $A[p(a), p(b)] = A[A(V_1), A(V_2)] = A(V_1 \cup V_2)$. Without loss of generality we may suppose that V_1 and V_2 contain the same number of elements, namely 2^n. Then $V_1 \cup V_2$ contains 2^{n+1} vectors and $C(V_1 \cup V_2) = \frac{1}{2}(a+b)$. Hence $A(V_1 \cup V_2) = p[\frac{1}{2}(a+b)]$. This proves that p satisfies (8. 1) for $a, b \in R$. Equation (8. 2) is handled in the same manner.

Lemma 10. *The mappings* $a \to p(a)$ *and* $a \to q(a)$ *are continuous on* R, *uniformly in* $(1-2^{-n})R$ *for any* n.

Proof. Since R is a set of type T and since $p(0) = q(0) = 0$, Theorem 5, gives continuity in R. Only the uniformity remains to be proved. Suppose that b and h are in R. Then with M and g as generic notations

$$M[g(b), g(h)] = g[\tfrac{1}{2}(b+h)].$$

Letting $h \to 0$ in R, we get continuity at $\frac{1}{2}b$ uniformly with respect to b in R. This gives uniform continuity in $(1-2^{-n})R$ for $n=1$ and the general case follows by induction.

9. First extension. Our functions p and q are defined in R which is dense in Π. The next step is to extend to all of Π.

Theorem 8. *The mappings* $a \to p(a)$, $a \to q(a)$ *can be extented as continuous mappings to all of* Π.

This follows from the fact that R is dense in Π and the primary mappings are uniformly continuous on any set $(1-2^{-n})R$.

Lemma 11. *The extented mappings are one-to-one.*

Proof. The argument is essentially that used for Lemma 8. Consider two distinct points a and b in Π. Here

$$a = \sum_{j=1}^{m} \alpha_j u_j, \quad \sum \alpha_j \leqq 1, \quad b = \sum_{j=1}^{m} \beta_j u_j, \quad \sum \beta_j \leqq 1, \tag{9.1}$$

and the α's and β's are real, non-negative. Set

$$c = \inf(a, b), \tag{9.2}$$

$$a = c + a_1, \quad b = c + b_1. \tag{9.3}$$

Here the unit vectors u_j which enter in a_1 are distinct from those of b_1. Either a_1 or b_1 may be zero but not both. If this should happen we have $g(a_1) \neq g(b_1)$. If both are positive vectors, they belong to principal linear subspaces having only the zero vector in common. Again $g(a_1) \neq g(b_1)$. Now

$$M[g(c), g(a_1)] = g[\tfrac{1}{2}(c+a_1)] = g(\tfrac{1}{2}a), \quad M[g(c), g(b_1)] = g[\tfrac{1}{2}(c+b_1)] = g(\tfrac{1}{2}b).$$

Here the first members are distinct by (3. 13), hence also the last. But if $g(a) = g(b)$, then $g(\frac{1}{2}a) = g(\frac{1}{2}b)$ since we would have $g(\frac{1}{2}a) = M[g(a), 0] = M[g(b), 0] = g(\frac{1}{2}b)$ which is a contradiction.

We can now construct fundamental solutions of the original equations (1. 1) and (1. 2). Since $a \to p(a)$ is one-to-one on Π there exists a unique inverse

$$a = P(s), \quad s = p(a), \quad a \in \Pi, \tag{9.4}$$

which is continuous and one-to-one.

Theorem 9. *The fundamental solution of* (1. 1) *is given by the mapping*

$$s \to f(s) = q[P(s)]. \tag{9.5}$$

Proof. Let $s = p(a)$, $t = p(b)$ where a and b are in Π. Then

$$A(s, t) = A[p(a), p(b)] = p[\tfrac{1}{2}(a+b)]$$

so that

$$f[A(s, t)] = f\{p[\tfrac{1}{2}(a+b)]\} = q[\tfrac{1}{2}(a+b)]$$

by the definition of f. On the other hand,

$$B[f(s), f(t)] = B\{q[P(s)], q[P(t)]\} = B[q(a), q(b)] = q[\tfrac{1}{2}(a+b)].$$

This shows that f is a solution of (1. 1) defined on the image of Π under the mapping $a \xrightarrow{p} s$. Since both p and q leave the vertices of Π invariant, the same is true for f so that f is the fundamental solution of (1. 1). It is obviously continuous since P and q are continuous.

In the same manner it is shown, that

$$p[Q(u)] = q(u), \quad Q(u) = a, \quad u = q(a)$$

is the fundamental solution of (1. 2).

10. Second extensions. We shall indicate a method of extending our functions to the outside of Π. We base the argument on the functional equations and are guided by a Principle of Permanence of Functional Equations. The letters G and M are used in the generic sense as above. We start by giving an elaborate interpretation of Postulate M_{22}.

Theorem 10. *For a given $z \in G$ there exists a positive number σ and a mapping $x \to H(x, z)$ with the following properties.*

(1) *H is defined and continuous for $\|x-z\| < \sigma$ and x its values lie in G.*

(2) *H is injective with respect to x.*

(3) *$H(z, z) = z$.*

(4) *Por $\|x-z\| < \sigma$*

(10. 1) $$M[x, H(x, z)] \equiv z.$$

(5) *H is unique.*

Proof. By M_{22} the mapping $t \to M(s, t)$ is open, injective and continuous, uniformly on compact subsets of G. Suppose that G_0 is such a set and $z \in G_0$. Let r be a fixed number $0 < r < \frac{1}{2} d[G_0, \partial G]$. Then there exists a positive number $\varrho = \varrho(r)$ such that for each $s \in G_0$ the set

(10. 2) $$E_s = \{u; u = M(s, t), \|t-s\| < r\}$$

contains an open sphere with center at $u = s$ and radius at least $\varrho(r)$. Take now the point z and choose an x at a distance $\leqq \frac{1}{2}\varrho(r)$ from z. Then the set E_x contains the open sphere $\|u-x\| < \varrho(r)$ and, in particular, the point z. Hence there is a vector $t = y$ with $\|y-x\| < r$ such that

(10. 3) $$M(x, y) = z.$$

Moreover, since the mapping is injective, there is one and only one such vector

(10. 4) $$y \doteq H(x, z).$$

This is H and for σ we can take $\frac{1}{2}\varrho(r)$. The continuity of M implies the continuity of H and H is injective in x because M has this property. This completes the proof.

We come now to the applications of this theorem to the extension problem. The results will be stated as lemmas. We consider first radial extension.

Lemma 12. *Let $a \to g(s)$ be a solution of*

(10. 5) $$M[g(a), g(b)] = g[\tfrac{1}{2}(a+b)]$$

which is defined, continuous and with values in G for a on a ray, $a = \alpha a_0$, $0 \leqq \alpha \leqq 1$, $a_0 \neq 0$. *Then there exists a* τ, $1 < \tau < \infty$ *such that g can be defined for* $1 < \alpha \leqq \tau$ *with the same properties.*

Proof. In Theorem 10 we take

$$g[(1-\eta)a_0] = x, \quad g(a_0) = z, \quad g[(1+\eta)a_0)] = y. \tag{10.6}$$

Here η is a small positive number, so small that $\|x-z\| < \frac{1}{2}\varrho(r)$ with r referring to a compact set containing such parts of the ray which are needed for the argument. Theorem 10 then asserts the existence of a unique $y = H(x, z)$ and we can satisfy (10. 5) by defining

$$g[(1+\eta)a_0] = H(x, z) = H\{g[(1-\eta)a_0], g(a_0)\}. \tag{10.6}$$

We can take $\tau = 1+\eta$ for any admissible η.

This lemma enables us to cross the "base" of the pyramid Π, i.e. the plane

$$x_1 + x_2 + \cdots + x_m = 1.$$

In particular, g will be defined along the positive axes some distance beyond the vertices of Π.

So far g has been defined only for non-negative vectors. The extension to negative vectors is again based on Theorem 10.

Lemma 13. *If the origin is an interior point of G and if* $g(0) = 0$ *and* a_0 *is a vector such that* $g(a)$ *is defined for* $a = \alpha a_0$, $0 \leqq \alpha \leqq 1$, $a_0 \neq 0$, *then there exists a* $\tau > 0$ *such that*

$$g(-\alpha a_0) = H[g(\alpha a_0), 0] \tag{10.7}$$

for $0 \leqq \alpha \leqq \tau$.

Proof by inspection. In particular, g is definable on the negative axes for sufficiently small values of $\|a\|$.

Lemma 14. *The domain of definition of g is a convex subset of* R^m.

Proof. Equation (10. 5) shows that g is defined at the point $\frac{1}{2}(a+b)$, if it is known at a and at b and if $g(a)$ and $g(b)$ belong to G where the meanvalue M is defined.

Lemma 15. *If g satisfies* (10. 4) *for all a and b in a convex domain* D^*, *then for any n and for any choice of* $a_1, a_2, \ldots, a_n$ *in* D^* *we have*

$$M[g(a_1), g(a_2), \ldots, g(a_n)] = g\left[\frac{1}{n}(a_1 + a_2 + \cdots + a_n)\right]. \tag{10.8}$$

Proof. The lemma is true for $n=2^k$ by induction on k and by retrogressive induction (following the precepts of Cauchy) from n to $n-1$.

Lemma 16. *The mapping $a \to g(a)$ is one-to-one.*

Proof. The argument given for Lemma 11 applies to the present case.

Applying these results to our functional equations (1. 1), (1. 2), (8. 1) and (8. 2) we are led to the following existence theorem.

Theorem 11. *Let the meanvalues A and B be defined in the convex domains D and E respectively, and satisfy the M-postulates. Let $\Pi \in D \cap E$.*

The fundamental solution $p(a)$ of (8. 1) is defined and continuous in a convex domain C_1, $\Pi \subseteq C_1$. The range R_1 of p is contained in D. The mapping $a \to p(a)$ is one-to-one.

For the fundamental solution $q(a)$ of (8. 2) replace C_1, R_1, D by C_2, R_2, E.

The fundamental solution of (1. 1) is defined by

$$f(s) = q[P(s)], \quad s = p(a), \quad a = P(s) \tag{10.9}$$

for $s \in S = p(C_1 \cap C_2) \subset R_1$ where f is continuous and one-to-one.

The fundamental solution g of (1. 2) is defined by

$$g(u) = p[Q(u)], \quad Q(u) = q, \quad u = q(a) \tag{10.10}$$

for $u \in U = q(C_1 \cap C_2) \subset R_2$ where g is continuous and one-to-one.

If $C_1 = C_2$, then f is defined in R_1 and g in R_2. It is reasonable to expect that R_1 and R_2 are also convex but this question must be left open.

11. Further comments. We have restricted ourselves to fundamental solutions. Initial conditions of the type

$$g(0) = y_0, \quad g(u_j) = y_j, \qquad j = 1, 2, \ldots, m, \tag{11.1}$$

or more generally

$$g(x_k) = y_k, \qquad k = 0, 1, 2, \ldots, m, \tag{11.2}$$

present a new and more difficult problem. The methods used above may or may not be effective in this case. Moreover, normally there is no simple relation between different solutions so information on the fundamental solution has little bearing on the behavior of other solutions.

This is totally different from the case

$$f[A(s, y)] = \tfrac{1}{2}[f(x) + f(y)] \tag{11.3}$$

where the general solution f and the fundamental solution f_0 satisfy

$$f(x) = \mathscr{C} f_0(x) + v_0. \tag{11.4}$$

Here $\mathscr{C}$ is an arbitrary m by m matrix and v_0 is an arbitrary vector. It is of course the linear character of (11. 3) with respect to f that accounts for the difference. Let us note in passing

Lemma 17. *The mapping $x \to f(x)$ is one-to-one iff $\mathscr{C}$ is non-singular.*

Proof. If $\mathscr{C}$ is non-singular then $f(x_1)-f(x_2) = \mathscr{C}[f_0(x_1)-f_0(x_2)]$. The quantity in square brackets is different from 0 if $x_1 \neq x_2$ so the same must hold for the left member. On the other hand, if $\mathscr{C}$ is singular we can find a vector w_0 such that $w_0 \neq 0$ and $\mathscr{C}w_0=0$. Here $\|w_0\|$ is at our disposal and may be chosen so small that there is an $x_0 \neq 0$ but small so that $f_0(x_0)=w_0$. Then $f(x_0) = \mathscr{C}f(x_0)+v_0 = \mathscr{C}w_0+v_0 = v_0 = f(0)$, so the mapping is not one-to-one.

It was observed above, formulas (3. 6) and (7. 6), that $M(s,t)-\frac{1}{2}(s+t)$ is small and at most $O(\|s-t\|)$. This can be made more precise.

Lemma 18. *Let $m=1$ and suppose that the generating function h of Theorem M has continuous first and second order derivatives. If $s<t$ there exists an s_1, $s<s_1<t$, such that*

$$M(s,t) = \tfrac{1}{2}(s+t)+\tfrac{1}{8}\frac{h''(s_1)}{h'(s_1)}(t-s)^2. \tag{11.5}$$

Proof. Since $M(s,s)=s$ and $\frac{1}{2}(s+t) = s+\frac{1}{2}(t-s)$ it is evidently required to prove that

$$M_s(s,t)|_{t=s} = \tfrac{1}{2}, \quad M_{ss}(s,t)|_{t=s} = \tfrac{1}{4}\frac{h''(s)}{h'(s)}. \tag{11.6}$$

These relations we obtain from the functional equation (3. 10)

$$h[M(s,t)] = \tfrac{1}{2}[h(s)+h(t)].$$

Differentiation with respect to s gives

$$h'[M(s,t)]\,M_s(s,t) = \tfrac{1}{2}h'(s),$$

$$h''[M(s,t)]\,[M_s(s,t)]^2+h'[M(s,t)]\,M_{ss}(s,t) = \tfrac{1}{2}h''(s).$$

Here we put $t=s$ and solve for $M_s(s,s)$ and $M_{ss}(s,s)$, noting that $h'(s)\neq 0$ since h is strictly monotone. The result is (11. 6), and (11. 5) is Taylor's theorem with remainder.

This is elementary and so is the extension to $m>1$ but the latter requires a number of devices and a much more thorough use of the functional equation.

Theorem 12. *If the generating function h of Theorem M has continuous first and second order partial derivatives, then*

$$M(s,t) = \tfrac{1}{2}(s+t)+O(\|s-t\|^2). \tag{11.7}$$

Proof. Let $z=(z_1, z_2, \ldots, z_m)\in R^m$, $s=(s_1, s_2, \ldots, s_m)\in G$ and let α be real and so small that $s+\alpha z\in G$. Consider $M(s, s+\alpha z)$. We have

$$(11.8)\qquad h[M(s, s+\alpha z)] = \tfrac{1}{2}[h(x)+h(x+\alpha z)].$$

The assumption on h implies that the right member may be differentiated with respect α and hence also the left member so that $M(s, s+\alpha z)$ is differentiable. Set

$$h = (h_1, h_2, \ldots, h_m), \quad M = (M_1, M_2, \ldots, M_m).$$

Each h_j is a real-valued function of the m real arguments $M_1, M_2, \ldots, M_m$ and is differentiable with respect to each of them. Let $h_{j,k}$ denote the partial of h_j with respect to the k^{th} argument with similar notation for other vector functions. For a fixed j we equate the j^{th} components of the two sides in (11. 7) and differentiate once with respect to α to obtain

$$(11.9)\qquad \sum_{k=1}^{m} h_{j,k}(*) \sum_{p=1}^{m} M_{k,p}(*)z_p = \tfrac{1}{2}\sum_{k=1}^{m} h_{j,k}(\cdot)z_k$$

where

$$(\cdot) = (s_1+\alpha z_1, s_2+\alpha z_2, \ldots, s_m+\alpha z_m), \quad (*) = (M_1(\cdot), M_2(\cdot), \ldots, M_m(\cdot)).$$

We now set $\alpha=0$ in (11. 9). Both $(\cdot)$ and $(*)$ collapse and become $(s_1, s_2, \ldots, s_m)=s$ so that (11. 9) becomes

$$(11.10)\qquad \sum_{k=1}^{m} h_{j,k}(s) \sum_{p=1}^{m} M_{k,p}(s)z_p = \tfrac{1}{2}\sum_{k=1}^{m} h_{j,k}(s)z_k.$$

Since z is arbitrary, (11. 10) must hold identically in the components $z_1, z_2, \ldots, z_m$. This gives a system of m equations

$$(11.11)\qquad \sum_{k=1}^{m} h_{j,k}(s)[M_{k,p}(s)-\tfrac{1}{2}\delta_{kp}] = 0, \qquad p = 1, 2, \ldots, m,$$

which, regarded as a system of equations for the derivatives of the components of h, certainly has non-trivial solutions. Hence the matrix

$$(11.12)\qquad \mathcal{J}(s)-\tfrac{1}{2}\mathcal{E}_m$$

is singular. Here $\mathcal{E}_m$ is the unit matrix in R^m and $\mathcal{J}(s)$ is a Jacobian

$$(11.13)\qquad \mathcal{J}(s) = (M_{k,p}(s)).$$

We note that $\mathcal{J}(s)$ has the characteristic value $\tfrac{1}{2}$ for all values of s. Actually this is the only characteristic value and we shall prove the stronger result

$$(11.14)\qquad \mathcal{J}(s) \equiv \tfrac{1}{2}\mathcal{E}_m.$$

To this end we note that (11. 8), regarded as a functional equation for h when M is given, has infinitely many solutions which, as observed above, are of the form (11. 4)

$$h(s) = \mathscr{C} h_0(s) + v_0.$$

Here h_0 is the fundamental solution, $\mathscr{C}$ is an arbitrary m by m constant matrix and v_0 is a vector with constant components. We can use this freedom of choice to normalize equations (11. 11) for a particular but arbitrary value of s, $s = s_0 \in G$. The set of m systems of m equations will obviously simplify very much, in fact become trivial, if we can determine the matrix $\mathscr{C}$ so that

$$h_{j,k}(s_0) = \delta_{j,k}, \qquad j, k = 1, 2, \dots, m \tag{11. 15}$$

To attain this, note that the matrix

$$\mathscr{H}^0(s) = (h^0_{j,k}(s)) \tag{11. 16}$$

of the first order derivatives of the components of the fundamental solution h_0 is necessarily non-singular since the mapping $s \to h_0(s)$ is one-to-one. It follows that we can solve equations of the form

$$\sum_{p=1}^{m} c_{jp} h^0_{p,k}(s_0) = h_{j,k}(s_0), \qquad j, k = 1, 2, \dots, m \tag{11. 17}$$

for the c's. We have m systems corresponding to a fixed value of k, all systems with the same non-vanishing determinant, namely that of the matrix $\mathscr{H}^0(s_0)$. In particular, this is possible if the right hand sides are given by (11. 15). This determines the matrix $\mathscr{C}$ and the normalizing transformations at $s = s_0$. The equations (11. 11) now give

$$M_{k,p}(s_0) = \tfrac{1}{2}\delta_{kp}, \qquad k, p = 1, 2, \dots, m \tag{11. 16}$$

and (11. 14) holds for $s = s_0$. Since s_0 is arbitrary in G, (11. 14) must hold identically in s.

This gives

$$M(s, s+\alpha z)|_{\alpha=0} = \left(\sum_{p=1}^{m} M_{1,p}(s) z_p, \dots, \sum_{p=1}^{m} M_{m,p}(s) z_p \right) = \tfrac{1}{2}(z_1, z_2, \dots, z_m) = \tfrac{1}{2} z.$$

We take $z = t - s$ and note again that $s + \frac{1}{2}(t-s) = \frac{1}{2}(s+t)$ which is the first term in the right member of (11. 7).

To get the remainder we can compute

$$\frac{\partial^2}{\partial \alpha^2} M(s, s+\alpha z)|_{\alpha=0} = \delta^2 M(s; z).$$

We do not insist on the exact expression for the second variation, it is clear that it is $O(\|z\|^2)$ uniformly in s on compact subsets of G.

Corollary. *Under the assumptions of Theorem* 12 *we have, for any n and for any choice of n vectors in G*

$$M(s_1, s_2, \ldots, s_m) = \frac{1}{n}(s_1 + s_2 + \cdots + s_m) + R \tag{11.19}$$

where $\|R\|$ *is dominated by a constant multiple of*

$$\sum_{1 \leq j < k \leq n} (\|s_j - s_k\|^2)$$

uniformly on compact subsets of G.

Proof. For $n = 2^k$ use induction on k. Complete by retrogressive induction from n to $n-1$. This gives the leading term. Note that the remainder must be a symmetric function of the s-vectors and should vanish when they are all equal.

References

[1] J. Aczél, Ein Eindeutigkeitssatz in der Theorie der Funktionalgleichungen und einige ihrer Anwendungen, *Acta Math. Acad. Sci. Hung.*, **15** (1964), 355—362.

[2] E. Hille, Vector Meanvalues and Related Functional Equations, *Rendiconti di Matematica, Università di Roma* (to appear).

[3] J. B. Miller, Aczél's Uniqueness Theorem and Cellular Internity, *Aequationes Math.*, **5** (1970), 319—325.

[4] C. T. Ng, Uniqueness Theorems for a General Class of Functional Equations, *J. Australian Math. Soc.*, **11** (1970), 362—366.

[5] ——— On Uniqueness Theorems of Aczél and Cellular Internity of Miller, *Aequationes Math.*, **7** (1971), 132—139.

[6] ——— A Characterization of the Quasilinear Weighted Mean on K^m. (Personal communication)

UNIVERSIDAD DE LA LAGUNA
LA LAGUNA, TENERIFE, SPAIN
UNIVERSITY OF CALIFORNIA SAN DIEGO
LA JOLLA, CALIFORNIA, 92037, U.S.A.

Operators similar to contractions

By JOHN A. R. HOLBROOK in Guelph (Ontario, Canada)*)

In honor of Professor Béla Szőkefalvi-Nagy on his 60th birthday

1. Introduction. The purpose of this note is to establish the following characterization of those operators (continuous linear transformation) on a Hilbert space $\mathfrak{H}$ which are similar to contraction operators, and to point out some consequences.

Theorem. *An operator $T:\mathfrak{H}\to\mathfrak{H}$ is similar to a contraction if, and only if, there exist a Hilbert space $\mathfrak{K}$ and operators $A:\mathfrak{H}\to\mathfrak{K}$, $C:\mathfrak{K}\to\mathfrak{K}$, and $B:\mathfrak{K}\to\mathfrak{H}$ such that C is a contraction on $\mathfrak{K}$ (that is, $\|C\|\leqq 1$), and*

$$\sum_{n=0}^{\infty}\|BC^nA-T^n\|^2<\infty. \tag{1}$$

In one direction, this statement is trivial: if T is similar to a contraction $C:\mathfrak{H}\to\mathfrak{H}$, i.e. $T=SCS^{-1}$ for some operator $S:\mathfrak{H}\to\mathfrak{H}$ with operator inverse S^{-1}, then, taking $\mathfrak{K}=\mathfrak{H}$, $A=S^{-1}$, and $B=S$, we see that $BC^nA=T^n$ $(n=0, 1, 2, \ldots)$ so in this case (1) certainly holds. The reverse implication (proved in section 3) has some content, however, and provides a general principle from which we can immediately derive several of the known criteria for similarity to a contraction.

2. Applications. Perhaps the most direct application of this type is to the result of B. SZ.-NAGY and C. FOIAŞ that every operator of class C_ϱ is similar to a contraction; indeed, the theorem above, and the construction used in its proof, were inspired in part by a study of the Sz.-Nagy—Foiaş result. Recall that an operator $T:\mathfrak{H}\to\mathfrak{H}$ is said to be of class C_ϱ (where ϱ is some positive real number) when there exists a Hilbert space $\mathfrak{K}$ containing $\mathfrak{H}$ and a unitary operator $U:\mathfrak{K}\to\mathfrak{K}$ such that

$$T^n=\varrho P_{\mathfrak{H}}U^n|\mathfrak{H}\qquad (n=1, 2, 3, \ldots). \tag{2}$$

If we let $A:\mathfrak{H}\to\mathfrak{K}$ be the identity (inclusion) map, let C be the (contraction) U, and let $B=\varrho P_{\mathfrak{H}}$, then the Theorem certainly applies, since every term in (1) vanishes

*) Work supported in part by AF Grant AFOSR 1322-67.

except the one for which $n=0$; it follows that such an operator T is similar to a contraction. The original proof of Sz.-Nagy and Foiaş may be found in [9] or in the book [10, pp. 92—95]; it depends upon clever use of the special properties of the ϱ-dilation U.

A generalization of the C_ϱ classes has been proposed by H. Langer (see [10, p. 55]): given a non-negative operator A on $\mathfrak{H}$, T is said to be of class C_A whenever there is a Hilbert space $\mathfrak{K}$ containing $\mathfrak{H}$ and a unitary operator U on $\mathfrak{K}$ such that

$$T^n = A^{1/2} P_{\mathfrak{H}} U^n A^{1/2} \qquad (n = 1, 2, \ldots). \tag{3}$$

Since V. Istrăţescu has shown (see [5]) that $C_A \in C_{\|A\|}$, it follows by the theorem of Sz.-Nagy and Foiaş that such an operator T is similar to a contraction. Using our theorem it is just as easy to prove this directly by observing that (1) holds when we let the operator A of the theorem be $A^{1/2}$ (forgive the notation), let B be $A^{1/2} P_{\mathfrak{H}}$, and let C be the contraction U on $\mathfrak{K}$.

Our theorem also yields immediately the result of G.-C. Rota (see [7, Theorem 2]) that every operator T with spectral radius $\nu(T)<1$ is similar to a contraction. To see this, choose $\mathfrak{K}$, A and B as you please, and let $C=0$; then (1) reduces to

$$\|BA-I\|^2 + \sum_{n=1}^{\infty} \|T^n\|^2 < \infty, \tag{4}$$

and this follows directly from the spectral radius formula: $\lim_{n\to\infty} \|T^n\|^{1/n} = \nu(T)(<1)$. Rota's result may also be derived from the theorem of Sz.-Nagy and Foiaş discussed above since, as we pointed out in [4; see Theorem 5. 1], the inequality $\nu(T)<1$ implies that $T \in C_\varrho$ for all large enough values of ϱ.

Our remarks in Section 4 point out a role that maybe be played by the Theorem in establishing yet another sufficient condition for similarly to a contraction: T is similar to a contraction providing its characteristic function $\Theta_T(\lambda)$ is uniformly bounded on the unit disc $\{\lambda : |\lambda| < 1\}$. This condition is due to C. Davis and C. Foiaş (see [1]).

Finally, the present theorem is an extension of the lemma which occurs in our paper [3], where further applications of these ideas may be found. It is possible, by making appropriate constructions, to prove our theorem by reducing it to the special case handled by the lemma of [3]. However, it is more efficient to proceed directly, and we give below a self-contained proof.

3. Proof of the theorem. It remains to show that, given A, B, and contraction $\mathscr{C}$ such that (1) holds, T is similar to a contraction. Equivalently, we must construct an equivalent inner product norm $|\cdot|$ on $\mathfrak{H}$ such that $|T| \leqq 1$, that is, such that T is a contraction with respect to $|\cdot|$.

To this end, define $|h|$ for each $h\in\mathfrak{H}$ by the relation

$$|h|^2 = \inf\{\|\sum_{n=0}^{\infty} C^n A h_n\|^2 + \sum_{n=0}^{\infty} \|h_n\|^2 : \sum_{n=0}^{\infty} T^n h_n = h\}; \tag{5}$$

here it is understood that only a finite number of the h_n are different from 0. It is a simple matter, using the triangle inequality in l^2, to show that $|\cdot|$ is a seminorm on $\mathfrak{H}$. Since $h = T^0 + T\cdot 0 + T^2\cdot 0 + \cdots$,

$$|h| \leqq (\|Ah\|^2 + \|h\|^2)^{1/2} \leqq (\|A\|^2+1)^{1/2}\,\|h\|. \tag{6}$$

On the other hand, if $\sum_{n=0}^{\infty} T^n h_n = h$, then

$$\begin{aligned} \|h\| &= \|B\,(\sum_{n=0}^{\infty} C^n A h_n) + \sum_{n=0}^{\infty} (T^n - BC^n A)\,h_n\| \leqq \\ &\leqq \|B\|\,\|\sum_{n=0}^{\infty} C^n A h_n\| + \sum_{n=0}^{\infty} \|T^n - BC^n A\|\,\|h_n\| \leqq \\ &\leqq \left(\|B\|^2 + \sum_{n=0}^{\infty} \|T^n - BC^n A\|^2\right)^{1/2} \left(\|\sum_{n=0}^{\infty} C^n A h_n\|^2 + \sum_{n=0}^{\infty} \|h_n\|^2\right)^{1/2}, \end{aligned} \tag{7}$$

using the Schwarz inequality in l^2. It follows from (5) and (7) that

$$\|h\| \leqq \left(\|B\|^2 + \sum_{n=0}^{\infty} \|T^n - BC^n A\|^2\right)^{1/2} |h|. \tag{8}$$

and the constant in this inequality is finite by (1). We now know (via (6) and (8)) that $|\cdot|$ is a norm on $\mathfrak{H}$ equivalent to the given norm $\|\cdot\|$. Moreover, if $\sum_{n=0}^{\infty} T^n h_n = h$, then $\sum_{n=1}^{\infty} T^n h_{n-1} = Th$, and

$$\begin{aligned} \|\sum_{n=1}^{\infty} C^n A h_{n-1}\|^2 + \sum_{n=1}^{\infty} \|h_{n-1}\|^2 &= \|C\,(\sum_{n=0}^{\infty} C^n A h_n)\|^2 + \sum_{n=0}^{\infty} \|h_n\|^2 \leqq \\ &\leqq \|\sum_{n=0}^{\infty} C^n A h_n\|^2 + \sum_{n=0}^{\infty} \|h_n\|^2, \end{aligned} \tag{9}$$

since C is a contraction. From (9) it follows that $|Th| \leqq |h|$ $(h\in\mathfrak{H})$ so that $|T| \leqq 1$.

It remains to show that $|\cdot|$ is an inner product norm. Recall the characterization of inner product norms due to P. Jordan and J. von Neumann (see [6]): a norm $|\cdot|$ on $\mathfrak{H}$ is an inner product norm if (and only if) the "parallelogram law" holds, that is,

$$|h+g|^2 + |h-g|^2 = 2(|h|^2 + |g|^2) \qquad (h, g\in\mathfrak{H}). \tag{10}$$

Now if $\alpha > 2(|h|^2+|g|^2)$, then there exist h_n, g_n such that $\sum_{n=0}^{\infty} T^n h_n = h$, $\sum_{n=0}^{\infty} T^n g_n = g$, and

$$(11) \qquad \alpha > 2\,\Big(\|\sum_0^\infty C^n A h_n\|^2 + \|\sum_0^\infty C^n A g_n\|^2 + \sum_0^\infty (\|h_n\|^2 + \|g_n\|^2)\Big).$$

By the parallelogram law for the given norms in $\mathfrak{K}$ and $\mathfrak{H}$, the right-hand side of (11) can be replaced by

$$(12) \qquad \|\sum_0^\infty C^n A(h_n+g_n)\|^2 + \|\sum_0^\infty C^n A(h_n-g_n)\|^2 + \sum_0^\infty (\|h_n+g_n\|^2 + \|h_n-g_n\|^2).$$

Since $\sum_{n=0}^{\infty} T^n(h_n \pm g_n) = h \pm g$, it follows that $\alpha > |h+g|^2+|h-g|^2$, and hence

$$(13) \qquad |h+g|^2+|h-g|^2 \leqq 2\,(|h|^2+|g|^2) \qquad (h, g \in \mathfrak{H}).$$

Finally, (13) is equivalent to (10), since the reverse inequality follows from (13) upon replacing h by $h+g$ and g by $h-g$.

4. Remarks. Condition (1) of our Theorem involves the operators BC^nA-T^n for *all* n. To demonstrate that some such condition is necessary, it may be well to point out that for an *arbitrary* operator T and any finite N we can satisfy the equalities $BC^nA=T^n$ $(n=0, 1, 2, \dots, N)$, with A, B, and C as in the Theorem. In fact, as L. J. WALLEN and J. S. JOHNSON have observed (see P. R. HALMOS [2, p. 910]), we may assume that $\mathfrak{H}$ is a subspace of $\mathfrak{K}$, that A is simply the identity on $\mathfrak{H}$, that C is an isometry, and that B is a skew projection of $\mathfrak{K}$ onto $\mathfrak{H}$. We then have, for $n=0, 1, 2, \dots, N$,

$$(14) \qquad BC^n|\mathfrak{H} = T^n.$$

If, however, (14) holds for *all* $n=0, 1, 2, \dots$, then it follows as a corollary of our Theorem that T must be similar to a contraction. This corollary may be used as an alternative to some of the arguments of Davis and Foiaş in establishing their subtle result concerning operators with bounded characteristic function (see [1]). We shall indicate briefly how this may be done.

Given an operator T on Hilbert space $\mathfrak{H}$, let $Q_T=|I-T^*T|^{1/2}$, $J_T=\operatorname{sgn}(I-T^*T)$, and let $\mathfrak{D}_T$ denote the closure in $\mathfrak{H}$ of the subspace $Q_T\mathfrak{H}$. The characteristic function of T is the following operator-valued function of the complex variable λ:

$$(15) \qquad \Theta_T(\lambda) = \big(-TJ_T + \lambda Q_{T^*}(I-\lambda T^*)^{-1}Q_T\big)|\,\mathfrak{D}_T.$$

The assumption of Davis and Foiaş in [1] is that $\Theta_T(\lambda)$ is defined throughout the open unit disc (that is, the spectral radius $\nu(T)\leqq 1$) and that

$$(16) \qquad \sup_{|\lambda|<1} \|\Theta_T(\lambda)\| < \infty.$$

Davis and Foiaş show that the condition (16) implies that T is similar to a contraction, and in the course of their proof they conclude that the J-isometric dilation U of T is power-bounded. That is, they show that $\sup_{n\geq 0} \|U^n\| < \infty$, where U is the operator defined on the Hilbert space

$$K_+ \quad H \oplus \mathfrak{D}_T \oplus \mathfrak{D}_T \oplus \dots \tag{17}$$

by the relation

$$U(h_0 \oplus h_1 \oplus h_2 \oplus \cdots) = Th_0 \oplus Q_T h_0 \oplus h_1 \oplus h_2 \oplus \cdots. \tag{18}$$

Note that $P_{\mathfrak{H}} U^n | \mathfrak{H} = T^n$ ($n=0, 1, 2, \dots$), where $P_{\mathfrak{H}}$ denotes orthogonal projection of $\mathfrak{K}_+$ onto $\mathfrak{H}$ (imbedded in $\mathfrak{K}_+$ as in (17)). It is easy to see from (18) that U is expansive; that is, $\|Uk\| \geq \|k\|$ ($k \in \mathfrak{K}_+$). Once it is established, then, that U is power-bounded, the well-known technique of B. Sz.-Nagy (see [8]) allows us to define a new, equivalent, inner product on $\mathfrak{K}_+$, with respect to which U is isometric. In this new geometry on $\mathfrak{K}_+$, $P_{\mathfrak{H}}$ is generally a skew projection (no longer orthogonal), but, with $B = P_{\mathfrak{H}}$ and $C = U$, (14) holds for all n. By the corollary, then, T is similar to a contraction. Thus, with a further change in the geometry of $\mathfrak{H}$, T becomes a contraction.

The approach to the Davis—Foiaş theorem outlined above emphasizes the fact that assuming the J-isometric dilation U of T to be power-bounded is in itself sufficient to ensure that T is similar to a contraction. This assumption may be expressed directly in terms of T: it is easy to verify that $\sup_{n\geq 0} \|U^n\| < \infty$ if, and only if,

$$\sup_{\|h\| \leq 1} \sum_{n=0}^{\infty} \|Q_T T^n h\|^2 < \infty. \tag{19}$$

The arguments of Davis and Foiaş show that (16)$\Rightarrow$(19), but it appears that (19) may be a weaker condition (which still allows the conclusion that T is similar to a contraction). In any case, (19) (and hence (16)) is not a *necessary* condition for similarity to a contraction; if, for example, T is the operator on $\mathbf{C}^2$ corresponding to the matrix $\begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}$, then $T^2 = I$ and $Q_T \neq 0$, so that (19) is certainly violated, but it is well known that any finite-dimensional, power-bounded operator is similar to a contraction.

References

[1] C. Davis and C. Foiaş, Operators with bounded characteristic function and their J-unitary dilation, *Acta Sci. Math.*, **32** (1971), 127—140.

[2] P. R. Halmos, Ten problems in Hilbert space, *Bull. Amer. Math. Soc.*, **76** (1970), 887—933.

[3] J. Holbrook, Spectral dilations and polynomially bounded operators, *Indiana Univ. Math. J.*, **20** (1971), 1027—1034.

[4] J. HOLBROOK, On the power-bounded operators of Sz.-Nagy and Foiaş, *Acta Sci. Math.*, **29** (1968), 299—310.
[5] V. ISTRĂŢESCU, A remark on a class of power-bounded operators in Hilbert space, *Acta Sci. Math.*, **29** (1968), 311—312.
[6] P. JORDAN and J. VON NEUMANN, On inner products in linear metric spaces, *Ann. Math.*, **36** (1935), 719—723.
[7] G.-C. ROTA, On models for linear operators, *Comm. Pure Appl. Math.*, **13** (1960), 468—472.
[8] B. SZ.-NAGY, On uniformly bounded linear transformations in Hilbert space, *Acta Sci. Math.*, **11** (1947), 152—157.
[9] B. SZ.-NAGY and C. FOIAŞ, Similitude des opérateurs de classe C_ϱ à des contractions, *C. R. Acad. Sci. Paris*, série A, **264** (1967), 1063—1065.
[10] B. SZ.-NAGY and C. FOIAŞ, *Harmonic Analysis of Operators on Hilbert Space*, North-Holland Publishing Company and Akadémiai Kiadó (Amsterdam and Budapest, 1970).

UNIVERSITY OF GUELPH
GUELPH, ONTARIO, CANADA

(Received September. 3, 1972)

Eine neue Charakterisierung der halbeinfachen Ringe

Von A. KERTÉSZ in Debrecen

Herrn Prof. B. Sz.-Nagy anläßlich seines 60. Geburtstages in Verehrung und Zuneigung gewidmet

Ein (assoziativer) Ring R heißt *halbeinfach*, falls er (rechts-)artinsch und (im Sinne von Jacobson) radikalfrei ist. Die Klasse der halbeinfachen Ringe ist in der Literatur — wegen ihrer großen Bedeutung für die Ringtheorie — sehr oft und intensiv behandelt worden. Das Ziel dieser kleinen Note ist, eine neue Charakterisierung der halbeinfachen Ringe anzugeben.

Nach der Terminologie von [4] nennen wir ein Rechtsideal A des Ringes R *quasimodular*, wenn es für jedes $x \notin A$ $(x \in R)$ ein $y (\in R)$ gibt, so daß $yx \notin A$ gilt. Ein Rechstsideal B von R heißt *modular*, wenn R ein Linkseinselement e mod B besitzt, d.h., wenn für jedes $r (\in R)$

$$r - er \in B$$

gilt. Offensichtlich ist jedes modulare Rechtsideal in R auch quasimodular. Es gibt dagegen Ringe, die gewisse quasimodulare Rechtsideale besitzen, welche aber nicht modular sind (vgl. [3]).

Es gilt der folgende

Satz. *Ein Ring R ist genau dann halbeinfach, falls es in ihm endlich viele quasimodulare maximale Rechtsideale gibt, deren Durchschnitt das Nullideal ist.*[1])

Beweis. Es sei zunächst R halbeinfach. Dann besitzt R nach der wohlbekannten Noetherschen Charakterisierung der halbeinfachen Ringe ein Linkseinselement und ist eine (modultheoretische) direkte Summe von endlich vielen minimalen Rechtsidealen:

$R = A_1 + \cdots + A_n$ (A_i minimales Rechtsideal in R; $i = 1, \ldots, n$) (diesbezüglich vgl. etwa Satz 8. 9, S. 189 in [2]). Die Rechtsideale

$$A^{(i)} = A_1 + \cdots + A_{i-1} + A_{i+1} + \cdots + A_n \qquad (i = 1, \ldots, n)$$

---

[1]) Dieser Satz ist eine Verschärfung der Behauptung der Äquivalenz der Aussagen (V) und (X) in Satz 8.9 des Buches [2] (S. 189).

sind offensichtlich quasimodulare maximale Rechtsideale in R und es gilt

$$A^{(i)} \cap \cdots \cap A^{(n)} = (0).$$

Umgekehrt sei (0) der Durchschnitt von endlich vielen quasimodularen maximalen Rechtsidealen. Dann ist R gemäß Satz 5. 24 in [2] (S. 132) radikalfrei. Ferner ist R nach Lemma von [1] eine (modultheoretische) direkte Summe von endlich vielen seiner minimalen Rechtsideale und erst recht artinsch. Folglich ist R ein halbeinfacher Ring.

Damit ist der Beweis des Satzes erbracht.

Literatur

[1] A. Kertész, Modules and semi-simple rings. II, *Publ. Math. Debrecen*, **4** (1955—56), 229—236.

[2] A. Kertész, *Vorlesungen über artinsche Ringe* (Budapest—Leipzig, 1968).

[3] F. Szász, Lösung eines Problems bezüglich einer Charakterisierung des Jacobsonschen Radikals, *Acta Math. Acad. Sci. Hung.*, **18** (1967), 261—272.

[4] F. Szász, Eine Charakterisierung des Jacobsonschen Radikals eines Ringes, *Bull. Acad. Polon. Ser. math. astr. phys.*, **15** (1967), 53—56.

(Eingegangen am 28. August 1972)

On commutative universal algebras

By L. KLUKOVITS in Szeged

To Professor Béla Szőkefalvi-Nagy on his 60th birthday

Universal algebras are called commutative (or Abelian) in [2], [3] and [4] under different conditions. It has to be remarked that in the case of groups these conditions are satisfied only for the Abelian groups and this is the reason why the term "commutative" is used. The conditions mentioned above are the following:

M. (B. I. PLOTKIN [2], p. 32.) In case of a universal algebra (shortly algebra) (A, Ω) for any operations μ and ν (m- and n-ary ones, respectively) the equality $(a_{11} \dots a_{1m}\mu)\dots(a_{n1} \dots a_{nm}\mu)\nu = (a_{11} \dots a_{n1}\nu)\dots(a_{1m} \dots a_{nm}\nu)\mu$ holds under any matrix $(a_{ij})_{n\times m}$ over A (see also P. M. COHN [1], p. 127).

H. (A. G. KUROŠ [3], p. 92.) The set of all homomorphisms from any algebra (G, Ω) into the algebra (A, Ω) admits the operations in Ω, i.e., if $\varphi_1, \dots, \varphi_m$ are homomorphisms from (G, Ω) into (A, Ω) and $\mu(\in\Omega)$ is any m-ary operation, then the mapping

$$(\varphi_1 \dots \varphi_m \mu): G \to A$$

defined for any element $g \in G$ as

$$g(\varphi_1 \dots \varphi_m \mu) = (g\varphi_1)\dots(g\varphi_m)\mu$$

is also a homomorphism. It has to be mentioned that in [3], the case of 0-ary operation was studied separately. However, this fact is irrelevant here, since an operation of this type can be considered a special unary operation ω with $a\omega = b\omega$ for any elements a, b. Thus we can suppose that in the set Ω there are no 0-ary operations.

E. (B. CSÁKÁNY [4].) The set of all endomorphisms of the algebra (A, Ω) admits the operations in Ω, i.e., for any endomorphisms $\varepsilon_1, \dots, \varepsilon_m$ and any m-ary operation μ (in Ω) the mapping

$$(\varepsilon_1 \dots \varepsilon_m \mu): A \to A$$

defined for any element $a \in A$ as

$$a(\varepsilon_1 \dots \varepsilon_m \mu) = (a\varepsilon_1)\dots(a\varepsilon_m)\mu$$

is an endomorphism, too.

On the basis of a notion due to M. SERVI [5] we introduce a fourth condition. Consider any category $\mathscr{C}$, admitting finite direct composition. For any object $A(\in \text{Ob}\, \mathscr{C})$, each morphism $\varphi \in \text{Hom}\,(A^m, A)$ (m is a natural number) is called m-ary operation on A. If $A \in \text{Ob}\, \mathscr{C}$ and Φ is a set of operations defined above, let the ordered couple (A, Φ) be called a Servi-algebra in the given category $\mathscr{C}$ (e.g., the topological algebras are exactly the Servi-algebras in the category of all topological spaces).

So the fourth condition is as follows:

C. The algebra (A, Ω) is a Servi-algebra in the category of algebras of the same type $\mathscr{A}$.

We prove that these conditions are equivalent for varieties of algebras.

Lemma. *For any universal algebra* (A, Ω) *the conditions* **M**, **H** *and* **C** *are equivalent and* **E** *follows from each of them.*

Proof. The equivalence of conditions **M** and **H** is proved in [3], therefore we have only to prove the equivalence of **M** and **C**.

Consider any category $\mathscr{A}$ of universal algebras similar to a given algebra (A, Ω). We want to prove that the mapping

$$\bar{\mu}: A^m \to A$$

defined by

$$(x_1, \ldots, x_m)\,\bar{\mu} = x_1 \ldots x_m\, \mu$$

is a homomorphism from (A^m, Ω) into (A, Ω). Let $(a_{11}, \ldots, a_{1m}), \ldots, (a_{n1}, \ldots, a_{nm}) \in A^m$, and let ν be any n-ary operation in Ω. Then we have

$$\big((a_{11}, \ldots, a_{1m}) \ldots (a_{n1}, \ldots, a_{nm})\,\nu\big)\,\bar{\mu} = \big((a_{11} \ldots a_{n1}\,\nu), \ldots, (a_{1m} \ldots a_{nm}\,\nu)\big)\,\bar{\mu} =$$

$$= (a_{11} \ldots a_{n1}\,\nu) \ldots (a_{1m} \ldots a_{nm}\,\nu)\,\mu,$$

and, on the other hand, we have

$$\big((a_{11}, \ldots, a_{1m})\,\bar{\mu}\big) \ldots \big((a_{n1}, \ldots, a_{nm})\,\bar{\mu}\big)\,\nu = (a_{11} \ldots a_{1m}\,\mu) \ldots (a_{n1} \ldots a_{nm}\,\mu)\,\nu.$$

Now, if condition **M** holds, then we can see that $\bar{\mu}$ is a homomorphism and conversely, if $\bar{\mu}$ is a homomorphism, then condition **M** holds true. The second statement of the lemma is obvious.

A simple counter-example shows that, in general from the condition **E** the condition **M** does not follow. In fact let us consider the groupoid G, defined by the following multiplication table:

γ	a	b	c
a	a	b	c
b	b	a	b
c	c	c	a

G has three endomorphisms given by the following table

	ω	ι	α
a	a	a	a
b	a	b	c
c	a	c	b

Hence it is obvious that condition **E** holds for G. On the other hand,

$$(ab\gamma)(ca\gamma)\gamma = b, \qquad (ac\gamma)(ba\gamma)\gamma = c,$$

i.e., condition **M** is not valid.

We shall say that for a variety of universal algebras the conditions mentioned above hold if they hold for each algebra in the given variety.

Theorem. *For any variety of universal algebras all the conditions* **M**, **H**, **E** *and* **C** *are equivalent.*

Proof. In view of the lemma, it is enough to prove if condition **E** holds for the variety $\mathfrak{A}$, then condition **M** holds, too. The way of proving this is similar to that of T. Evans [6] concerning groupoids.

Let F denote the free algebra generated by the set $X=\langle x_1, x_2, \ldots\rangle$ in the variety $\mathfrak{A}$. It is sufficient to show that the equality in condition **M** holds for any $n\times m$ matrix over X.

Let μ and ν be arbitrary operations (m- and n-ary respectively) in the variety $\mathfrak{A}$ and $(x_{ij})_{n\times m}$ any matrix over X. We define the following mappings ε_k ($k=1, 2, \ldots, m$) on the set X,

$$x_{j1}\varepsilon_k = x_{jk}$$

for all $1\leqq j\leqq n$. These mappings can be extended to endomorphisms of the free algebra F. Thus we have

$$x_{j1}(\varepsilon_1 \ldots \varepsilon_m \mu) = (x_{j1}\varepsilon_1)(x_{j1}\varepsilon_2)\ldots(x_{j1}\varepsilon_m)\mu = x_{j1}x_{j2}\ldots x_{jm}\mu$$

and therefore

$$(x_{11}\ldots x_{1m}\mu)\ldots(x_{n1}\ldots x_{nm}\mu)\nu = (x_{11}\ldots x_{n1}\nu)(\varepsilon_1\ldots\varepsilon_m\mu) =$$

$$= (x_{11}\ldots x_{n1}\nu\varepsilon_1)\ldots(x_{11}\ldots x_{n1}\nu\varepsilon_m)\mu =$$

$$= ((x_{11}\varepsilon_1)\ldots(x_{n1}\varepsilon_1)\nu)\ldots((x_{11}\varepsilon_m)\ldots(x_{n1}\varepsilon_m)\nu)\mu =$$

$$= (x_{11}\ldots x_{n1}\nu)\ldots(x_{1m}\ldots x_{nm}\nu)\mu.$$

This completes the proof.

References

[1] P. M. Cohn, *Universal algebra*, Harper & Row (New York, Evanston and London, 1965).
[2] Б. И. Плоткин, *Группы автоморфизмов алгебраических систем*, Наука (1966).
[3] А. Г. Курош, *Общая алгебра* (лекции 1969—70 учебного года), МГУ (Москва, 1970).
[4] Б. Чакань, Об абелебых свойствах примитивных классов универсальных алгебр, *Acta Sci. Math.*, **25** (1964), 202—208.
[5] M. Servi, Questioni di algebra universale in una categoria astratta. II, *Annali Univ. Ferrara*, **15** (1970), 57—92.
[6] T. Evans, Properties of algebras almost equivalent to identities, *J. London Math. Soc.*, **35** (1962), 53—59.
[7] L. Klukovits, Commutative universal algebras, *Proceedings of the Mini-Conference on Universal Algebra*, Szeged (1971), 30—32.

(Received September 5, 1972)

H^p-spaces on compact nilmanifolds

By ADAM KORÁNYI in New York (N.Y., U.S.A.)

To Professor Béla Szőkefalvi-Nagy on his 60th birthday

§ 1. Introduction

The purpose of this paper is to generalize to several complex variables the following classical facts. Let D be the upper halfplane, $\operatorname{Im} z > 0$. In analogy to the usual Hardy spaces $H^p(D)$ one can define the spaces $H^p(\mathbf{Z} \setminus D)$ of periodic holomorphic functions (with period 1) on D. Adjoining a point at infinity and introducing the new variable $z' = e^{2\pi i z}$, these spaces are transformed to the H^p-spaces of the unit disc; their theory can however be studied without using this transformation. It is immediate that $\{e^{2\pi i m z}\}_{m=0}^{\infty}$ is a complete orthonormal system in $H^2(\mathbf{Z} \setminus D)$, hence the Cauchy—Szegő kernel (the reproducing kernel of $H^2(\mathbf{Z} \setminus D)$) is

$$\sum_{m=0}^{\infty} e^{2\pi i m(z-\bar{w})} = \tfrac{1}{2} + \frac{i}{2} \operatorname{cotg} \pi (z-\bar{w}). \tag{1.1}$$

This is also equal to

$$\tfrac{1}{2} + \frac{i}{2\pi}\left[\frac{1}{z-\bar{w}} + \sum_{n \neq 0}\left(\frac{1}{z+n-\bar{w}} - \frac{1}{n}\right)\right] \tag{1.2}$$

which can be regarded as a periodicization of $i(2\pi)^{-1}(z-\bar{w})^{-1}$, the Cauchy—Szegő kernel of D. As z and w approach the real line, (1.2) becomes a periodicized Hilbert transform kernel, to which the theory of periodic singular integrals [3] applies. In this way, one gets a proof of the theorem of M. Riesz on the conjugate function.

In the present paper, D will be a generalized halfplane analytically equivalent with the unit ball in $\mathbf{C}^n$ ($n > 1$). For simplicity of notation, we will assume $n = 2$, but everything extends to the general case. The boundary of D can be identified with a certain nilpotent Lie group $\mathfrak{N}$, often referred to as a Heisenberg group, and the role of $\mathbf{Z}$ is taken by a discrete co-compact subgroup Γ of $\mathfrak{N}$. In the case $n > 1$, $\Gamma \setminus D$ is not contained in any Stein manifold, as one sees easily by the boundedness of holomorphic functions at infinity (§ 4) and by topological reasons. We prove the analogue of the identity of (1. 1) and (1. 2) by Fourier transform methods ((5. 3),

(5. 6), Theorem 5. 2). In § 2, we develop in a slightly more general setting the necessary theorems about periodic singular integrals on nilpotent groups. In § 6, we prove the generalized M. Riesz theorem.

Along the way, in § 4, we also discuss the position of $H^2(\Gamma\backslash D)$ in relation to the harmonic analysis of the regular representation on $L^2(\Gamma\backslash\mathfrak{N})$ and we point out that the transformation formula for θ-functions is, up to a constant factor of modulus one, a consequence of our results. In § 7, we make some observations about the Bergman kernel of $\Gamma\backslash D$.

Independently and in a different context, H. Rossi has also considered the manifold $\Gamma\backslash D$ and also constructed the orthogonal system (4. 3). I wish to express to him my thanks for several useful conversations on this subject.

§ 2. Singular integrals on compact nilmanifolds

Let G be a connected, simply connected nilpotent Lie group, e its identity element. Let $\{a(t)\}$ be a multiplicatively written one-parameter group of automorphisms of G such that

$$\lim_{t\to 0} a(t)\, g = e \tag{2.1}$$

for all $g\in G$, and such that the induced automorphisms $a_*(t)$ of the Lie algebra are diagonalizable.

Let dg be a fixed Haar measure on G. Then there exists $q>0$ such that

$$d\big(a(t)\,g\big) = t^q\, dg \tag{2.2}$$

for all $t>0$.

Given any real s, a function f defined on $G-\{e\}$ is said to be *homogeneous of degree* s if, for all $t>0$, $g\in G$,

$$f\big(a(t)\,g\big) = t^s f(g).$$

Let $g\mapsto |g|$ be a strictly positive, continuously differentiable function on $G-\{e\}$, homogeneous of degree 1, and such that $|g^{-1}|=|g|$. This function is then necessarily a *gauge* in the sense of [7], i.e.

$$|gh| \leqq \varkappa(|g|+|h|) \tag{2.3}$$

for all $g, h\in G$ with some universal $\varkappa$, and the sets

$$B(r) = \{g\in G \,|\, |g| < r\}$$

which form a base of neighborhoods of e, have measure $m(B(r))$ proportional to r^q.

In [7], the largest eigenvalue of the infinitesimal generator of $\{a_*(t)\}$ was denoted by α. As explained on p. 604 of [7], it is no restriction of generality to assume $\alpha=1$.

We will assume this in this section; our results (Theorems 2. 1 and 2. 2) are, of course, valid without change for any value of α.

Before stating our first theorem, we recall the special cases of Lemmas 5. 1. and 5. 2. in [7] which we will need repeatedly:

Let f be homogeneous of degree s and continuously differentiable on $G-\{e\}$. Then, for all $0<a<b$,

$$\int_{a<|g|<b} f(g)\,dg = \begin{cases} C(b^{q+s}-a^{q+s}) & \text{if } s\neq -q \\ C(\log b-\log a) & \text{if } s=-q \end{cases} \tag{2.4}$$

with some constant C. Furthermore, there exist numbers $M, N\geqq 1$ such that,

$$\left.\begin{matrix} |f(gh)-f(g)| \\ |f(hg)-f(g)| \end{matrix}\right\} < M|h|\,|g|^{s-1} \tag{2.5}$$

whenever $N|h|<|g|$.

Theorem 2. 1. *Let $k:G-\{e\}\to\mathbf{R}$ be continuously differentiable, homogeneous of degree $-q$, and such that*

$$\int_{a<|g|<b} k(g)\,dg = 0$$

for some (hence all) $0<a<b$. Let $k(e)=0$. Let Γ be a discrete subgroup of G. Then, for g, h in any compact subset of G,

$$k^*(g,h) = \sum_{\gamma\in\Gamma} [k(h^{-1}\gamma g)-k(\gamma)] \tag{2.6}$$

converges normally after the omission of finitely many terms. For any $\gamma,\delta\in\Gamma$.

$$k^*(\gamma g,\delta h) = k^*(g,h). \tag{2.7}$$

Proof. Let g, h be in a compact subset; then $|g|$, $|h|\leqq C$ with the some constant C. For $|g|\leqq C$ and $|\gamma|$ large we have

$$|\gamma g| \geqq \tfrac{1}{2}|\gamma|. \tag{2.8}$$

In fact, this follows since the homogeneity of the gauge and (2. 5) imply $\big||\gamma g|-|\gamma|\big| < M|g|$ whenever $N|g|<|\gamma|$.

By (2. 8) and (2. 5), it follows that

$$|k(h^{-1}\gamma g)-k(\gamma)| \leqq C'|\gamma|^{-q-1}$$

for $|h|$, $|g|\leqq C$ and $|\gamma|$ large. The first assertion of the theorem follows if we show

$$\sum_{|\gamma|>R} |\gamma|^{-q-1} < \infty$$

for some R.

Let $r>0$ be such that the sets $\gamma B(r)$ $(\gamma\in\Gamma)$ are all disjoint. Now, for $|g|<r$ and $|\gamma|>R$, with an appropriate R, we have (2. 8), and hence

$$\sum_{|\gamma|>R}|\gamma|^{-q-1}\leqq\frac{2^{q+1}}{m(B(r))}\sum_{|\gamma|>R}\int_{B(r)}|\gamma g|^{-q-1}\,dg\leqq\frac{2^{q+1}}{m(B(r))}\int_{|g|>R/2}|g|^{-q-1}\,dg$$

which is finite by (2. 4).

To prove the second assertion, we show that $k^*(\delta g, h)=k^*(g, h)$ for any fixed $\delta\in\Gamma$. Since the same method also proves $k^*(g,\delta h)=k^*(g,h)$, this will finish the proof.

By absolute convergence of (2. 6), it suffices to show that

$$\sum_{\gamma\in\Gamma}[k(h^{-1}\gamma\delta g)-k(h^{-1}\gamma g)]=0. \tag{2.9}$$

We note that, for large R, $|\gamma|<R-M|\delta|$ implies

$$|\gamma\delta|<R. \tag{2.10}$$

This follows from (2. 5) when $|\gamma|>N|\delta|$ and from (2. 3) when $|\gamma|\leqq N|\delta|$.

Consequently,

$$\sum_{|\gamma|<R}[k(h^{-1}\gamma\delta g)-k(h^{-1}\gamma g)] \tag{2.11}$$

is majorized by

$$\sum_{R-M|\delta|<|\gamma|<R+M|\delta|}|k(h^{-1}\gamma g)|.$$

By homogeneity, $|k(g)|\leqq C|g|^{-q}$. We use (2. 8) as in the proof of the first assertion, and then use (2. 10) which is true for any fixed δ (not necessarily in Γ) to obtain the majorization of (2. 11) by

$$\frac{2^{q+1}C}{m(B(r))}\int_{R-c<|g|<R+c}|g|^{-q}dg$$

where $c=M|\delta|+Mr$. By (2. 4) this integral tends to 0 as $R\to\infty$, finishing the proof.

From now on let Γ be a discrete subgroup such that $\Gamma\backslash G$ is compact. As usual, we will identify functions on $\Gamma\backslash G$ with functions f on G such that $f(\gamma g)=f(g)$ for all $\gamma\in\Gamma$, $g\in G$. We choose a fundamental domain Ω for the action of Γ on G such that $e\in\Omega$; as it is shown in [1] (p. 54), Ω can be chosen as the exponential image of an interval in the Lie algebra. There is a right G-invariant measure on $\Gamma\backslash G$ given by

$$\int_{\Gamma\backslash G}f=\int_{\Omega}f(g)\,dg.$$

We denote by $L^p(\Gamma\backslash G)$ the usual L^p-spaces with respect to this measure.

For $g, h\in G$, we define

$$\varrho(g,h)=\inf_{\gamma\in\Gamma}|h^{-1}\gamma g|.$$

It is clear from (2. 3) that, for any $g, h, l \in G$,

$$\varrho(g, l) \leqq \varkappa(\varrho(g, h) + \varrho(h, l)).$$

Also, $\varrho(\gamma g, \delta h) = \varrho(g, h)$ for $\gamma, \delta \in \Gamma$, so ϱ can be regarded as a pseudodistance on $\Gamma \diagdown G$.

Given k and k^* as in Theorem 2. 1, we define, for small $\varepsilon > 0$,

$$k^{*,\varepsilon}(g, h) = \begin{cases} k^*(g, h) & \text{if} \quad \varrho(g, h) \geqq \varepsilon \\ 0 & \text{if} \quad \varrho(g, h) < \varepsilon \end{cases}$$

This is a bounded measurable function on $\Gamma \diagdown G$, therefore the operator A^ε defined on $L^p(\Gamma \diagdown G)$ $(1 < p < \infty)$ by

$$(A^\varepsilon f)(g) = \int_{\Gamma \diagdown G} f(h) k^{*,\varepsilon}(g, h)\, dh$$

is bounded.

As in [7], we also define k^ε on G by

$$k^\varepsilon(g) = \begin{cases} k(g) & \text{if} \quad |g| \geqq \varepsilon \\ 0 & \text{if} \quad |g| < \varepsilon \end{cases}$$

Theorem 2. 2. *Let* $1 < p < \infty$. *For all* $f \in L^p(\Gamma \diagdown G)$, $Af = \lim_{\varepsilon \to 0} A^\varepsilon f$ *exists in* $L^p(\Gamma \diagdown G)$, *and* A *is a bounded operator on* $L^p(\Gamma \diagdown G)$.

Proof. Let

$$E = \{(g, h) \mid g \in \Omega, h \in \Omega g\}. \tag{2. 12}$$

Then $k^*(g, h) - k(h^{-1} g)$ is bounded on E; this follows from the easily proven fact that, for small $\varepsilon > 0$, $(g, h) \in E$ implies $|h^{-1} \gamma g| > \varepsilon$ for all $e \neq \gamma \in \Gamma$.

Define, for, for $g \in \Omega$,

$$(Tf)(g) = \int_{\Omega g} f(h) [k^*(g, h) - k(h^{-1} g)]\, dh$$

$$(T^\varepsilon f)(g) = \int_{\Omega g} f(h) [k^{*,\varepsilon}(g, h) - k^\varepsilon(h^{-1} g)]\, dh.$$

Since, for fixed g, Ωg is also a fundamental domain for Γ, it is clear by classical arguments that T, T^ε are bounded operators $L^p(\Gamma \diagdown G) \to L^p(\Omega)$ and $\lim_{\varepsilon \to 0} T^\varepsilon = T$ in the strong operator topology.

Let $A_1^\varepsilon = A^\varepsilon - T^\varepsilon$ regarded as an operator $L^p(\Gamma \diagdown G) \to L^p(\Omega)$. To prove the theorem it is enough to show that $\lim_{\varepsilon \to 0} A_1^\varepsilon$ exists strongly. For this, in turn, it is enough to show that if $\varepsilon, \eta \to 0$, then $A_1^{\varepsilon,\eta} f = A_1^{\varepsilon} f - A_1^{\eta} f \to 0$ in $L^p(\Omega)$ for all $f \in L^p(\Gamma \diagdown G)$. Note that we have, for $g \in \Omega$,

$$(A_1^{\varepsilon} f)(g) = \int_{\Omega g} f(h) k^{\varepsilon}(h^{-1})\, dh.$$

We define the function f_1 on G by

$$f_1(g) = \begin{cases} f(g) & \text{if} \quad g \in \Omega^2, \\ 0 & \text{if} \quad g \notin \Omega^2. \end{cases}$$

By compactness, Ω^2 is covered by finitely many translates of Ω; hence $\|f_1\|_{L^p(G)} \leqq M \|f\|_{L^p(\Gamma \diagdown G)}$ with some constant M.

For functions on G, let B^{ε} denote the operator of convolution by k^{ε}. By [7] (Theorem 5. 1), $\lim_{\varepsilon \to 0} B^{\varepsilon}$ exists strongly. So $B^{\varepsilon,\eta} f_1 = B^{\varepsilon} f_1 - B^{\eta} f_1 \to 0$ in $L^p(G)$ as $\varepsilon, \eta \to 0$. Hence, $B^{\varepsilon,\eta} f_1 \to 0$ also as an element of $L^p(\Omega)$.

Now, for $g \in \Omega$, we have, by a change of variable,

$$A_1^{\varepsilon,\eta} f(g) - B^{\varepsilon,\eta} f(g) = \int_{G - g^{-1}\Omega g} f_1(gl)\,[k^{\varepsilon}(l^{-1}) - k^{\eta}(l^{-1})]\, dl.$$

By compactness, there exists $r_0 > 0$ such that $B(r_0) \subset g^{-1}\Omega g$ for all $g \in \Omega$. Thus, the last integral is 0 for $\varepsilon, \eta < r_0$, finishing the proof.

Remark. It is possible to show that the singular integral in the sense of [7] converges also a.e., not only in $L^p(G)$. Using this, it is easy to show by the method of [3] that our $\lim_{\varepsilon \to 0} A^{\varepsilon} f(g)$ also exists for a.a. g. With this fact available, the proof of Theorem 2. 2. can be slightly simplified along the lines of [3].

§ 3. H^p-spaces

In the following, we consider the domain D of [7, § 6], but only in the special case of two variables. The reason for this restriction is that all significant features of our problem already appear in this special case. Everything that follows is trivially generalizable to n variables, the material of the present section even to any generalized halfplane.

So, let

$$D = \{(z_1, z_2) \in \mathbf{C}^2 \mid \operatorname{Im} z_1 - \tfrac{1}{2}|z_2|^2 > 0\}$$

and let B be its boundary. Let $\mathfrak{N}$ be the subgroup of the group of holomorphic automorphisms of D which as a set equals $\mathbf{R} \times \mathbf{C}$, an element (ξ, ζ) acting by

$$z_1 \mapsto z_1 + \xi + i\bar{\zeta} z_2 + \frac{i}{2}|\zeta|^2$$

$$z_2 \mapsto z_2 + \zeta$$

$\mathfrak{N}$ is simply transitive on B, which gives a natural identification $\mathfrak{N} \ni (\xi, \zeta) = g \mapsto$ $\mapsto g \cdot 0 = \left(\xi + \frac{i}{2}|\zeta|^2, \zeta\right) \in B$. Lebesgue measure on $\mathbf{R} \times \mathbf{C}$ corresponds to a Haar measure on $\mathfrak{N}$. The group $\{a(t)\}$ of automorphisms given by

$$a(t)(\xi, \zeta) = (t\xi, t^{1/2}\zeta)$$

has the properties (2. 1), (2. 2) will $q=2$ and we have a smooth homogeneous gauge on $\mathfrak{N}$ given by

$$|(\xi, \zeta)| = (\xi^2 + \tfrac{1}{4}|\zeta|^4)^{1/2}.$$

Let Γ be a discrete subgroup of $\mathfrak{N}$ such that $\Gamma \backslash \mathfrak{N}$ is compact. Then, $\Gamma \backslash D$ is a complex manifold with boundary $\Gamma \backslash \mathfrak{N}$ (since $\mathfrak{N}$ is identified with B). Again, we identity functions on $\Gamma \backslash D$ with functions f on D such that $f \circ \gamma = f$ for all $\gamma \in \Gamma$.

As in [6], for $t > 0$, we denote $f_t(z_1, z_2) = f(z_1 + it, z_2)$ and $\tilde{f}_t(g) = f_t(g \cdot 0)$. If $f \circ \gamma = f$, the same is true for f_t and $\tilde{f}_t$. So, we may define $H^p(\Gamma \backslash D)$ as the space of all holomorphic functions f on $\Gamma \backslash D$ for which

$$\|f\|_{H^p(\Gamma \backslash D)} = \sup_{t>0} \|\tilde{f}_t\|_{L^p(\Gamma \backslash \mathfrak{N})} < \infty.$$

Theorem 3. 1. *Let* $1 < p < \infty$. *If* $f \in H^p(\Gamma \backslash D)$, *then* $\tilde{f} = \lim_{t \to 0} \tilde{f}_t$ *exists in* $L^p(\Gamma \backslash \mathfrak{N})$, *and* $f \mapsto \tilde{f}$ *is an isometric imbedding.*

Proof. As known from [6], when f is bounded, continuous on $\bar{D}$ and holomorphic on D, then it has a Poisson integral representation

$$\tilde{f}_t(g) = \int_{\mathfrak{N}} \tilde{P}^t(h^{-1}g) f(h \cdot 0)\, dh \tag{3.1}$$

where, for $g = (\xi, \zeta)$,

$$\tilde{P}^t(g) = \frac{1}{2\pi^2} \frac{2}{[\xi^2 + (t + \frac{1}{2}|\zeta|^2)^2]^2}. \tag{3.2}$$

It is easy to see that for h, g in a compact set and for large $|\gamma|$,

$$|\tilde{P}^t(h^{-1}\gamma g)| \leq C \frac{t^2}{|\gamma|^4}. \tag{3.3}$$

This shows the normal convergence of the series

$$P_t^*(g, h) = \sum_{\gamma \in \Gamma} \tilde{P}^t(h^{-1}\gamma g)$$

and therefore also that, for f bounded continuous on $\Gamma \backslash \bar{D}$ and holomorphic on $\Gamma \backslash D$, we have

$$\tilde{f}_t(g) = \int_{\Gamma \backslash \mathfrak{N}} P_t^*(g, h) f(h \cdot 0)\, dh. \tag{3.4}$$

Now, let $f \in H^p(\Gamma \backslash D)$. For each fixed $t_0 > 0$, f_{t_0} has a representation (3. 4). (The boundedness of f_{t_0} follows from the usual subharmonicity argument for any generalized halfplane D; in our special case, it is obvious from the fact that holomorphic functions on $\Gamma \backslash D$ are automatically bounded at infinity, cf. § 4). The standard weak compactness argument shows now that f is the Poisson integral of some function in $L^p(\Gamma \backslash \mathfrak{N})$.

Next, we show that whenever f is the Poisson integral of some $\varphi \in L^p(\Gamma \backslash \mathfrak{N})$ we have $\lim_{t \to 0} \tilde{f}_t = \varphi$ in $L^p(\Gamma \backslash \mathfrak{N})$. In fact, using Jensen's inequality [12, Vol. I, p. 24].

$$\|\varphi - \tilde{f}_t\|^p \leqq \int_{\Omega} dg \int_{\Omega g} P_t^*(g, h) |\varphi(g) - \varphi(h)|^p \, dh \leqq$$

$$\leqq \int \int_{\substack{(g,h) \in E \\ |h^{-1}g| < \eta}} \tilde{P}^t(h^{-1}g) |\varphi(g) - \varphi(h)|^p \, dg \, dh +$$

$$+ \int \int_{\substack{(g,h) \in E \\ |h^{-1}g| < \eta}} [P_t^*(g, h) - \tilde{P}^t(h^{-1}g)] \, |\varphi(g) - \varphi(h)|^p \, dg \, dh +$$

$$+ \int \int_{\substack{(g,h) \in E \\ |h^{-1}g| > \eta}} P_t^*(g, h) |\varphi(g) - \varphi(h)|^p \, dg \, dh.$$

For η small enough, the first integral is small by the results of [6]; the second is small since $P_t^*(g, h) - \tilde{P}^t(h^{-1}g)$ is bounded on E. Once η is chosen, (3. 3) shows that the third integral can be made small by choosing t small enough.

We see now that, for $f \in H^p(\Gamma \backslash D)$, $\tilde{f} = \lim_{t \to 0} \tilde{f}_t$ exists in $L^p(\Gamma \backslash \mathfrak{N})$ and f is the Poisson integral of $\tilde{f}$. The latter fact and Jensen's inequality show $\|\tilde{f}_t\| \leqq \|\tilde{f}\|$ which implies $\|\tilde{f}\|_{L^p(\Gamma \backslash \mathfrak{N})} = \|f\|_{H^p(\Gamma \backslash D)}$, finishing the proof.

Corollary. *$H^2(\Gamma \backslash D)$ is a Hilbert space.*

§ 4. An orthonormal system in $H^2(\Gamma \backslash D)$

In this section, we specify Γ. Let k be a positive integer and τ a complex number such that $\operatorname{Im} \tau > 0$. Let

$$\Gamma_\tau^k = \left\{ (a \operatorname{Im} \tau, b + c\tau) \mid \; b, c, \frac{k}{2}(a + bc) \in \mathbf{Z} \right\}.$$

Using some arguments of Brezin [2, p. 614], it is not hard to show that we get all possible complex manifolds $\Gamma \backslash D$ up to isomorphism by taking $\Gamma = \Gamma_\tau^k$. Different values of τ give isomorphic complex manifolds if they are equivalent under the modular group. (It is well known [1] that the nilmanifolds $\Gamma_\tau^k \backslash \mathfrak{N}$ are isomorphic for any two values of τ.)

The inequalities

$$|\xi| < \frac{\operatorname{Im}\tau}{k}, \quad |\operatorname{Re}\zeta| < \tfrac{1}{2}, \quad |\operatorname{Im}\zeta| < \frac{\operatorname{Im}\tau}{2}$$

determine a fundamental domain Ω for Γ_τ^k in $\mathfrak{N}$.

It is immediate [9, p. 137] that every function f holomorphic on $\Gamma_\tau^k \backslash D$ has a Fourier expansion

$$f(z_1, z_2) = \sum_{m \equiv 0 \ (\mathrm{mod}\ k)} e^{(\operatorname{Im}\tau)^{-1} m\pi i z_1} \psi_m(z_2)$$

with ψ_m holomorphic and satisfying

$$\psi_m(z+\zeta) = \psi_m(z) e^{(\operatorname{Im}\tau)^{-1}\pi m(\bar{\zeta} z + \frac{1}{2}|\zeta|^2)} \tag{4.1}$$

for all $\zeta = b + c\tau$ $(b, c \in \mathbf{Z})$. Defining χ_m by

$$\chi_m(z) = e^{-\frac{\pi}{2}(\operatorname{Im}\tau)^{-1} m z^2} \psi_m(z)$$

(4. 1) is seen to be equivalent with the pair of equations

$$\chi_m(z+1) = \chi_m(z), \quad \chi_m(z+\tau) = e^{-m\pi i(\tau + 2z)} \chi_m(z).$$

These are the standard functional equations of θ-functions. They have holomorphic solutions only for $m \geqq 0$. As first observed by PJATECKIĬ-ŠAPIRO [9, p. 140], from this it is immediate that every holomorphic f is bounded as $\operatorname{Im} z_1 \to \infty$. For $m=0$, the only solutions are the constants; for $m>0$, a basis of the space of solutions is given by

$$\chi_{ml}(z) = \sum_{j \equiv l \ (\mathrm{mod}\ m)} e^{\frac{\pi i \tau}{m} j^2} e^{2\pi i j z} \tag{4.2}$$

$(0 \leqq l \leqq m-1)$, cf. [5], [10] where the several variable case is also handled, showing how to extend our results to n-dimensional D.

By Theorem 3. 1., the computation of inner products in H^2 reduces to simple integrations on Ω. An easy computation shows that

$$\begin{cases} \varphi_{00} \equiv 1, \\ \varphi_{ml}(z_1, z_2) = e^{(\operatorname{Im}\tau)^{-1}\pi m(i z_1 + \frac{1}{2} z_2^2)} \chi_{ml}(z_2) \end{cases} \tag{4.3}$$

$(0 \leqq l < m,\ m \equiv 0 \ (\mathrm{mod}\ k))$ is an orthogonal basis of $H^2(\Gamma_\tau^k \backslash D)$. An application of the Parseval identity gives the norms:

$$\|\varphi_{ml}\|^2_{H^2(\Gamma \backslash D)} = \begin{cases} \dfrac{2(\operatorname{Im}\tau)^2}{k} & \text{if} \quad (m, l) = (0, 0) \\ \dfrac{2^{1/2}(\operatorname{Im}\tau)^{3/2}}{k m^{1/2}} & \text{if} \quad (m, l) \neq (0, 0). \end{cases} \tag{4.4}$$

$L^2(\Gamma_\tau^k\backslash\mathfrak{N})$ carries a unitary representation of $\mathfrak{N}$, the "right regular representation", whose harmonic analysis is well known [8], [2]. We wish to elucidate the position of $H^2(\Gamma_\tau^k\backslash D)$ and of the system (4. 3) in this context. We have [8]

$$L^2(\Gamma_\tau^k\backslash\mathfrak{N}) = \bigoplus_{m\equiv 0 \pmod k} L(m)$$

where $L(0)$ consists of constants and, for $m\neq 0$, $L(m)$ is the sum of $|m|$ copies of the irreducible representation of $\mathfrak{N}$ determined by the character m of the center.

Let $\{X, Y, Z\}$ be the basis of the Lie algebra such that $\exp tX=(t, 0)$, $\exp Y=(it, 0)$, $\exp tZ=(0, t)$. It is known [4] that for an irreducible representation π of $\mathfrak{N}$, the space of solutions f of $d\pi(X+iY)f = 0$ (sometimes called the vacuum subspace) is one-dimensional.

In the present case, it is clear that $\tilde{\varphi}_{ml}\in L(m)$, also we have $\left.\frac{\partial\varphi_{ml}}{\partial\bar{z}_2}\right|_0 = 0$. An easy computation gives $\left.\frac{\partial f}{\partial\bar{z}_2}\right|_0 = dR(X+iY)\tilde{f}|_e$ for any f differentiable on $\bar{D}$. Since X, Y are left invariant and since $\mathfrak{N}$ acts holomorphically, it follows that $dR(X+iY)\tilde{\varphi}_{ml} = 0$.

Thus, the space spanned by $\{\tilde{\varphi}_{m,0}, \dots, \tilde{\varphi}_{m,m-1}\}$ $(m>0)$ is exactly the "vacuum subspace" of $L(m)$. If we denote by L_{ml} the space spanned by all right translates of $\tilde{\varphi}_{ml}$, we have

$$L(m) = \bigoplus_{l=0}^{m-1} L_{ml}$$

a decomposition of $L(m)$ into irreducible subspaces. By the method of BREZIN [2] one can now construct a concrete orthogonal basis of each L_{ml} $(m>0)$ which contains $\tilde{\varphi}_{ml}$ as its first element.

As a curiosity, we mention also the following. If $\tau' = -\frac{1}{\tau}$, then $\Gamma_\tau^1\backslash D$ and $\Gamma_{\tau'}^1\backslash D$ are isomorphic as complex manifolds under the map $\iota:(z_1, z_2)\mapsto(|\tau|^2 z_1, \tau z_2)$ and the map $f\mapsto |\tau|^2 f\circ\iota$ is an isomorphism of the corresponding H^2-spaces. Since $H^2(\Gamma_\tau^1\backslash D)\cap L(1)$ is spanned by $\tilde{\varphi}_{10}$, we have therefore $c|\tau|^2\|\varphi_{10}\|^{-1}\varphi_{10}\circ\iota = \|\varphi'_{10}\|^{-1}\varphi_{10}$, where φ'_{10} is constructed from τ' the way φ_{10} is from τ, and $|c|=1$. After a short computation this reduces to

$$c|\tau|^{-1/2} e^{\pi\frac{z^2}{i\tau}} \chi'_{10}\left(\frac{z}{\tau}\right) = \chi_{10}(z)$$

where χ'_{10} is given by (4. 2) with τ' instead of τ. This equation, except for the exact value of the argument of c, is the fundamental transformation formula of θ-functions.

§ 5. The Szegő kernel of $\Gamma\backslash D$

In this section, we fix one of the groups Γ_τ^k and denote it briefly by Γ. The Szegő kernel S of D is given by [6]

(5.1) $$S(z, w) = \frac{1}{2\pi^2}[i(z_1-\bar{w}_1)-z_2\bar{w}_2]^{-2}$$

We define

(5.2) $$c^\Gamma = \frac{k}{4(\operatorname{Im}\tau)^2} + \lim_{r\to\infty} \sum_{0\neq|\gamma|<r} S(\gamma\cdot 0, 0).$$

It will be seen in the proof of Theorem 3.1 that this limit exists. Further, we define

(5.3) $$S^*(z, w) = c^\Gamma + S(z, w) + \sum_{e\neq\gamma}[S(\gamma z, w) - S(\gamma\cdot 0, 0)].$$

Lemma 5.1. *For z, w in any bounded subset of $\mathbf{C}^2\times\mathbf{C}^2$, the sum in* (5.3) *converges normally after the omission of finitely many terms.*

Proof. From (5.1) a simple explicit computation gives

$$|S(\gamma z, w) - S(\gamma\cdot 0, 0)| \leqq C|\gamma|^{-5/2}$$

for $|\gamma|$ large. Therefore, except for finitely many terms, the series is majorized by $\sum|\gamma|^{-5/2}$ which converges by (2.4) and the argument of Theorem 2.1.

Theorem 5.2. *S^* is the Szegő kernel of $\Gamma\backslash D$, i.e. the reproducing kernel of the Hilbert space $H^2(\Gamma\backslash D)$.*

Proof. By the same argument as in Theorem 2.1 one shows that $S^*(\gamma z, \delta w) = S(z, w)$ for all $\gamma, \delta\in\Gamma$. Thus S^* can be regarded as a function on $\Gamma\backslash D\times\Gamma\backslash D$.

For fixed $w\in D$, we introduce the usual notation $S_w^*(z) = S^*(z, w)$. S_w^* is a holomorphic function on $\bar{D}$ (and on $\Gamma\backslash\bar{D}$); this follows since, by Lemma 5.1, S_w^* is meromorphic everywhere and since none of the terms in (5.3) have any pole on $\bar{D}$. It follows now that S_w^* is bounded at infinity (cf. § 4), and hence that $S_w^*\in H^2(\Gamma\backslash D)$.

From the Poisson integral representation (Theorem 3.1), it follows that point evaluations are continuous functionals on $H^2(\Gamma\backslash D)$. Therefore to prove the Theorem, it is enough to prove that $(f, S_w^*) = f(w)$ for all w and for a system of functions that span $H^2(\Gamma\backslash D)$. We will show this for the orthogonal basis $\{\varphi_{ml}\}$.

First, we have

$$(\varphi_{00}, S_w^*) = \int_\Omega S^*(g\cdot 0, w)\,dg =$$

$$= c^\Gamma\frac{2(\operatorname{Im}\tau)^2}{k} + \lim_{r\to\infty}\left[\int_{|g|<r} S(g\cdot 0, w)\,dg - \frac{2(\operatorname{Im}\tau)^2}{k}\sum_{0\neq|\gamma|<r} S(\gamma\cdot 0, 0)\right].$$

By [7, Lemma 6.2], the limit of the integral on the right is 1/2. This shows that the limit in (5.2) exists, and also that $(\varphi_{00}, S_w^*) = 1 = \varphi_{00}(w)$.

Now, let $(m, l) \neq (0, 0)$. To compute (φ_{ml}, S_w^*) we note that by normal convergence of (5. 3) the series can be rearranged in any way and integrations can be performed term by term. Since clearly we have $\int_\Omega \varphi_{ml} = 0$, it follows that

$$(5.4)\qquad (\varphi_{m,l}, S_w^*) = \sum_{\gamma \in \Gamma} \int_\Omega S(w, \gamma u)\varphi_{ml}(u)\, d\beta(u)$$

where, as in [6], we denote $u = \left(x_1 + \frac{i}{2}|z_2|^2, z_2\right)$, $z_2 = x_2 + iy_2$, and $d\beta(u) = dx_1\, dx_2\, dy_2$.

We have

$$S(w, z) = 2\int_0^\infty e^{2\pi\lambda[i(z_1 - \bar{w}_1) - z_2 \bar{w}_2]} \lambda\, d\lambda$$

which shows that $S(w, u)$ as a function of x_1 is the Fourier transform of a continuous L^2-function. By [11, Theorem 58], the Fourier—Plancherel inversion formula is applicable pointwise (not only a.e.); thus, using (4. 3) and (4. 1),

$$\lim_{N\to\infty} \sum_{|\xi| \leq N} \int_{-\frac{\operatorname{Im}\tau}{k}}^{\frac{\operatorname{Im}\tau}{k}} S(w, (\xi, \zeta)\cdot u)\varphi_{ml}(u)\, dx_1 =$$

$$= \frac{m}{\operatorname{Im}\tau} e^{(\operatorname{Im}\tau)^{-1}\pi m[iw_1 - |z_2+\zeta|^2 + w_2(\bar{z}_2 + \bar{\zeta}) + \frac{1}{2}(z_2+\zeta)^2]} \chi_{ml}(z_2 + \zeta).$$

To find (5. 4), we have to integrate this with respect to x_2 and y_2, and sum over ζ. This gives

$$(5.5)\qquad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{m}{\operatorname{Im}\tau} e^{(\operatorname{Im}\tau)^{-1}\pi m[iw_1 - |z_2|^2 + \bar{z}_2 w_2 + \frac{1}{2} z_2^2]} \chi_{ml}(z_2)\, dx_2\, dy_2$$

To justify this step, it is enough to show that the integral exists absolutely. This follows by considering the series (4. 2) which defines χ_{ml}; for each term separately, the absolute integral exists and their sum is convergent.

This also shows that (5. 5) can be computed by substituting (4. 2) and integrating term-by-term. This is a direct computation which shows that (5.5) equals $\varphi_{ml}(w)$, finishing the proof.

Remark. From (4. 3), (4. 4) it follows by general principles that

$$(5.6)\qquad S^*(z, w) = \frac{k}{2(\operatorname{Im}\tau)^2} + \frac{k}{2^{1/2}(\operatorname{Im}\tau)^{3/2}} \sum_{\substack{m \equiv 0 \ (\mathrm{mod}\ k) \\ m > 0}} m^{1/2} \sum_{l=0}^{m-1} \varphi_{ml}(z)\overline{\varphi_{ml}(w)}$$

§ 6. The generalized M. Riesz theorem

Theorem 5. 2 shows that for $f \in L^2(\Gamma \backslash \mathfrak{N})$

$$Pf(z) = \int_{\Gamma \backslash \mathfrak{N}} S^*(z, g \cdot 0) f(g) \, dg \tag{6. 1}$$

gives the orthogonal projection of f onto $H^2(\Gamma \backslash D)$. We will show that (6. 1) also defines a bounded projection $L^p(\Gamma \backslash \mathfrak{N}) \to H^p(\Gamma \backslash D)$ for all $1 < p < \infty$. In the one-variable case this is an equivalent formulation of the classical M. Riesz theorem on the conjugate function.

As in [7], we denote for $t > 0$, $g = (\xi, \zeta)$.

$$k_t(g) = \frac{1}{2\pi^2} (t + \tfrac{1}{2}|\zeta|^2 - i\xi)^{-2}$$

$$k(g) = \frac{1}{2\pi^2} (\tfrac{1}{2}|\zeta|^2 - i\xi)^{-2}.$$

As shown in [7, § 6], k is a kernel satisfying the conditions of Theorem 2. 2.

We define k^ε, k^*, $k^{*,\varepsilon}$ as in § 2. Furthermore, we define

$$k_t^*(g, h) = c^\Gamma + \sum_{\gamma \in \Gamma} [k_t(h^{-1} \gamma g) - k(\gamma)]$$

Clearly, we have $k_t^*(g, h) = S^*(g \cdot 0, h(it, 0))$, and therefore

$$(Pf)_t^\sim(g) = \int_{\Gamma \backslash \mathfrak{N}} k_t^*(g, h) f(h) \, dh. \tag{6. 2}$$

Lemma 6. 1. *Let $1 < p < \infty$. Then, for all $f \in L^p(\Gamma \backslash \mathfrak{N})$,*

$$\lim_{\varepsilon \to 0} \int_{\Gamma \backslash \mathfrak{N}} f(h) [k_\varepsilon^*(g, h) - k^{*,\varepsilon}(g, h)] \, dh = \tfrac{1}{2} f(g)$$

in the sense of convergence in $L^p(\Gamma \backslash \mathfrak{N})$.

Proof. We give a sketch, omitting much tedious detail. Let E be the set (2. 12). Then, there exists $\varepsilon_0 > 0$ such that the series

$$k_\varepsilon^*(g, h) - k_\varepsilon(h^{-1} g) = \sum_{\gamma \neq e} [k_\varepsilon(h^{-1} \gamma g) - k(\gamma)]$$

is normally convergent for $(g, h) \in E$, $0 < \varepsilon \leqq \varepsilon_0$. This is a slight extension of Lemma 5. 1 and can be verified by some explicit computation. Using this, one shows next that

$$\sum_{\gamma \neq e} [k_\varepsilon(h^{-1} \gamma g) - k^\varepsilon(h^{-1} \gamma g)]$$

is bounded for $(g, h) \in E$, $0 < \varepsilon \leqq \varepsilon_0$, and tends to 0 as $\varepsilon \to 0$. Thus, the integral operator defined on $L^p(\Omega)$ by this kernel tends to 0. This reduces the proof of the lemma to proving that

$$\lim_{\varepsilon\to 0}\int_{\Omega g} f(h)\,[k_\varepsilon(h^{-1}g)-k^\varepsilon(h^{-1}g)]\,dh = \tfrac{1}{2}f(g).$$

Making a change of variable, we have the identity

$$\int_{\Omega g} f(h)\,[k_\varepsilon(h^{-1}g)-k^\varepsilon(h^{-1}g)\,dh = I_1+I_2+I_3,$$

$$I_1 = f(g)\int_{g^{-1}\Omega g}[k_\varepsilon(h^{-1})-k^\varepsilon(h^{-1})]\,dh,$$

$$I_2 = \int_{\substack{g^{-1}\Omega g\\ |h|>\varepsilon}}[f(gh)-f(h)]\,[k_\varepsilon(h^{-1})-k^\varepsilon(h^{-1})]\,dh,$$

$$I_3 = \int_{\substack{g^{-1}\Omega g\\ |h|<\varepsilon}}[f(gh)-f(g)]\,k_\varepsilon(h^{-1})\,dh.$$

By compactness, there exists $r>0$ such that $B(r)\subset g^{-1}\Omega g$ for all $g\in\Omega$. I_1 can be written as the sum of an integral on $B(r)$ and one on its complement. The first tends to $\frac{1}{2}f(g)$ as $\varepsilon\to 0$, by the corollary of [7, Lemma 6. 2]. The second tends to 0 since the integrand tends to 0 uniformly. One shows that I_2, I_3 tend to 0 as $\varepsilon\to 0$ by using Minkowski's integral inequality in the same way as in the proof of [7, Lemma 6. 3].

Theorem 6. 2. *Let* $1<p<\infty$. *The operator* P, *defined by* (6. 1) *for all* $f\in L^p(\Gamma\diagdown\mathfrak{N})$, *is a bounded projection onto* $H^p(\Gamma\diagdown D)$. *The boundary function of* Pf *is given by*

$$(6.\ 3)\qquad (Pf)^\sim(g) = \tfrac{1}{2}f(g)+\lim_{\varepsilon\to 0}\int_{\Gamma\diagdown\mathfrak{N}} k^{*,\varepsilon}(g,h)\,f(h)\,dh.$$

Proof. Pf is holomorphic since S^* is holomorphic. By Theorem 2. 2, the limit (6. 3) exists in $L^p(\Gamma\diagdown\mathfrak{N})$ and defines a bounded operator. Lemma 6. 1 shows that the boundary function of Pf is given by (6. 3). It follows that the L^p-norm of $(Pf)_t^\sim$ is bounded for small $t>0$; it is also bounded for large t since Pf, being holomorphic, is bounded at infinity. So, $Pf\in H^p(\Gamma\diagdown D)$.

To see that P is a projection, we have to show $P^2=P$. Now, $P^2f=Pf$ for $f\in L^2(\Gamma\diagdown\mathfrak{N})$ by Theorem 5. 2, and $L^2(\Gamma\diagdown\mathfrak{N})$ is dense in each $L^p(\Gamma\diagdown\mathfrak{N})$. To see that the range of P is all of $H^p(\Gamma\diagdown D)$, we note that $Pf=f$ for all $H^2(\Gamma\diagdown D)$ and $H^2(\Gamma\diagdown D)$ is dense in each $H^p(\Gamma\diagdown D)$ since $f=\lim_{t\to 0} f_t$ for all $f\in H^p(\Gamma\diagdown D)$, by Theorem 3. 1.

§ 7. Remarks on the Bergman kernel

Let $\mathscr{L}^2(\Gamma\backslash D)$ be the space of holomorphic functions square integrable on $\Gamma\backslash D$. A fundamental domain for the action of Γ on D is given by $(\operatorname{Re} z_1, z_2)\in\Omega$, $\operatorname{Im} z_1 > \frac{1}{2}|z_2|^2$ where Ω is as in § 4. Inner products in $\mathscr{L}^2(\Gamma\backslash D)$ are computed by integrating on this domain.

It is easy to see that the system (4. 3) with the omission of φ_{00} is an orthogonal basis in $\mathscr{L}^2(\Gamma\backslash D)$ and one has

(7. 1) $$\|\varphi_{ml}\|^2_{\mathscr{L}^2(\Gamma\backslash D)} = \frac{\operatorname{Im}\tau}{2\pi m}\|\varphi_{ml}\|^2_{H^2(\Gamma\backslash D)}.$$

Denoting by K^* the Bergman kernel of $\Gamma\backslash D$, we have by (4. 3), (4. 4),

(7. 2) $$K^*(z, w) = \frac{2^{1/2}\pi k}{(\operatorname{Im}\tau)^{5/2}} \sum_{\substack{m\equiv 0 \,(\mathrm{mod}\, k)\\ m>0}} m^{3/2} \sum_{l=0}^{m-1} \varphi_{ml}(z)\overline{\varphi_{ml}(w)}.$$

By (4. 3) and (5. 6), this is equal to $-2i\frac{\partial}{\partial z_1}S^*(z, w)$. Let K denote the Bergman kernel of D. From the explicit formulas in [6], we see that $K(z, w) = -2i\frac{\partial}{\partial z_1}S(z, w)$. Thus, by (5. 3), it follows that

(7. 3) $$K^*(z, w) = \sum_{\gamma\in\Gamma} K(\gamma z, w).$$

References

[1] L. AUSLANDER, L. GREEN et al., *Flows on homogeneous spaces*, Ann. of Math. Studies 53, Princeton University Press (Princeton, 1963).

[2] J. BREZIN, Harmonic analysis on nilmanifolds, *Trans. Amer. Math. Soc.*, **150** (1970), 611—618.

[3] A. P. CALDERÓN and A. ZYGMUND, Singular integrals and periodic functions, *Studia Math.*, **14** (1954), 249—271.

[4] P. CARTIER, Quantum mechanical commutation relations and Theta functions, *Proceedings of Symposia in Pure Mathematics*, Amer. Math. Soc., **9** (1966), 361—383.

[5] F. CONFORTO, *Abelsche Funktionen und algebraische Geometrie*, Springer-Verlag (1956).

[6] A. KORÁNYI, The Poisson integral for generalized half-planes and bounded symmetric domains, *Ann. of Math.*, **82** (1965), 332—350.

[7] A. KORÁNYI and S. VÁGI, Singular integrals on homogeneous spaces and some problems of classical analysis, *Ann. Scuola Norm. Sup. Pisa, Classe di Scienze*, **25** (1971), 575—648.

[8] C. C. MOORE, Decomposition of unitary representations defined by discrete subgroups of nilpotent groups, *Ann. of Math.*, **82** (1965), 146—182.

[9] I. I. PJATECKIĬ-ŠAPIRO, *Geometry of classical domains and theory of automorphic functions* (Moscow, 1961) (in Russian).

[10] C. L. SIEGEL, *Analytic functions of several complex variables*. Lecture notes, The Institute of Advanced Study, Princeton 1948, reprinted 1962.
[11] E. C. TITCHMARSH, *Introduction to the theory of Fourier integrals*, Clarendon Press (Oxford, 1937).
[12] A. ZYGMUND, *Trigonometric series*, Cambridge University Press (1959).

(Received June 6, 1972)

Über die Q-Funktion eines π-hermiteschen Operators im Raume $\Pi_\varkappa$

Von M. G. KREĬN in Odessa, und H. LANGER in Dresden

Herrn Professor Dr. Dr. h. c. Szőkefalvi-Nagy zum 60. Geburtstag am 29. VII. 1973 in Verehrung gewidmet

Der Begriff der Q-Funktion eines hermitischen Operators im Hilbertraum mit gleichen endlichen Defektzahlen n wurde von einem der Verfasser in den Arbeiten [1] ($n=1$) und [2] ($n<\infty$) eingeführt. Später verallgemeinerte Š. N. SAAKJAN [3] diese Definition auf den Fall beliebiger (endlicher oder unendlicher) gleicher Defektzahlen.

Für einen π-hermiteschen Operator im Pontrjaginschen Raume $\Pi_\varkappa$ haben wir die Q-Funktion bei beliebigen gleichen Defektzahlen in [4] eingeführt. Dabei erwies sie sich analog wie in den vorangegangenen Arbeiten als ein wichtiges Hilfsmittel zur Beschreibung aller verallgemeinerten Resolventen des gegebenen π-hermiteschen Operators.

In der vorliegenden Arbeit setzen wir die Untersuchung der Q-Funktion im Anschluß an [4] fort und zeigen insbesondere, daß sie sich durch verhältnismäßig einfache Eigenschaften vollständig charakterisieren läßt (Hauptsatz aus § 2. 4). Möglicherweise ist dieses Ergebnis auch im Falle des Hilbertraumes ($\varkappa=0$) neu.

Beim Beweis dieses Hauptsatzes spielen die Funktionen der Klassen $N_\varkappa^0(\mathfrak{G})$ und $N_\varkappa(\mathfrak{G})$ eine besondere Rolle. Einige Ergebnisse über solche Funktionen in den Paragraphen 3 und 4 kann man als Verallgemeinerungen bekannter Resultate zum sog. Nevanlinna—Pick-Problem [5], [6] und seinem operatortheoretischen Analogon [7] ansehen. Eventuell sind auch dabei gewisse Aussagen selbst für den positiv definiten Fall ($\varkappa=0$) neu, z. B. das Kriterium für die Selbstadjungiertheit des Operators A_Q^0 in § 4. 3.

In § 6 wird zwischen der Q-Funktion eines π-hermiteschen Operators und der charakteristischen Funktion seiner π-isometrischen Cayleytransformierten, die wir in [8] einführten, ein Zusammenhang hergestellt[1]). Daraus ergibt sich für den Hilbert-

[1]) Dieser war im Falle des Hilbertraumes den Verfassern von [1] und [9] klar, sobald diese Arbeiten erschienen; er wurde jedoch bis jetzt nirgends dargestellt.

raum sofort, daß ein einfacher hermitescher Operator durch seine Q-Funktion bis auf unitäre Äquivalenz eindeutig bestimmt ist. Das analoge Resultat für einfache π-hermitesche Operatoren beweisen wir in § 5 unmittelbar, ohne Benutzung der charakteristischen Funktion. Hierzu sei noch vermerkt (siehe § 1), daß die Abspaltung des einfachen Teiles eines π-hermiteschen Operators komplizierter ist als im definiten Fall; insbesondere läßt sich ein π-hermitescher Operator i.a. nicht als direkte Summe eines einfachen π-hermiteschen und eines π-selbstadjungierten Operators darstellen.

Beim Studium π-hermitescher Operatoren im Pontrjaginschen Raume treten spezifische Fragestellungen auf. Wir nennen hier insbesondere den wichtigen Begriff des rein hyperbolischen Operators[2]). Die Untersuchung der Q-Funktion gestattet es, Kriterien dafür aufzustellen, wann ein π-hermitescher Operator rein hyperbolisch ist (§ 7).

Die vorliegenden Untersuchungen stützen sich auf die Ergebnisse unserer Arbeiten [4], [8]. Im Unterschied zu diesen benutzen wir hier jedoch wesentlich die in [12] eingeführte Spektralzerlegung eines π-selbstadjungierten Operators im Pontrjaginschen Raum[3]).

§ 1. Der einfache Teil eines π-hermiteschen Operators

1. *Allgemeine Grundlagen.* Es sei A im folgenden stets ein π-hermitescher Operator[4]) im $\pi_\varkappa$-Raume $\Pi_\varkappa$, $0 \leqq \varkappa < \infty$, mit gleichen, endlichen oder unendlichen Defektzahlen: $\mathfrak{n}_+(A) = \mathfrak{n}_-(A) = \mathfrak{n}$. Wir setzen für einen beliebigen komplexen Punkt z:

$$\mathfrak{M}_z = (A - zI)\mathfrak{D}(A) \quad \text{und} \quad \mathfrak{N}_z = \mathfrak{M}_{\bar{z}}^{[\perp]}.$$

Dann ([4], § 2. 2) hat der Defektraum $\mathfrak{N}_z$ für alle Punkte z der offenen oberen Halbebene C_+ mit Ausnahme von höchstens $\varkappa$ Punkten die Dimension $\mathfrak{n}$; die erwähnten Ausnahmepunkte sind genau die in C_+ gelegenen Eigenwerte von A. Bezeichnen wir für einen solchen Eigenwert $z \in C_+ \cap \sigma_p(A)$ mit r_z seine algebraische Vielfachheit, so gilt $\dim \mathfrak{N}_z = \mathfrak{n} + r_z$ und $\sum r_z \leqq \varkappa$, wobei die Summation über alle $z \in \sigma_p(A) \cap C_+$

---

[2]) Die Existenz einer speziellen Klasse rein hyperbolischer π-hermitescher (π-isometrischer Operatoren wurde zuerst in [10] bemerkt. Bei ihr erweist sich das reelle Spektrum aller minimalen π-selbstadjungierten Erweiterungen als absolutstetig. In der vorliegenden Arbeit konstruieren wir jedoch rein hyperbolische π-hermitesche Operatoren, deren sämtliche kanonischen π-selbstadjungierten Erweiterungen diskretes Spektrum haben. Wir bemerken, daß rein hyperbolische Operatoren auch in der Streutheorie auftreten (vgl. [11]), worauf die Autoren von [10] an anderer Stelle eingehen werden.

[3]) Eine ausführliche Darstellung findet man z.B. in [13].

[4]) Wir benutzen die Bezeichnungen aus [4].

zu erstrecken ist. Entsprechende Aussagen gelten für die Punkte z der offenen unteren Halbebene C_-.

Wir bezeichnen im folgenden mit Δ_A die Menge derjenigen nichtreellen Punkte z, für die das π-Skalarprodukt auf $\mathfrak{N}_z$ entartet. Dann enthält Δ_A keine inneren Punkte ([4], Satz 3. 3), und das nichtreelle Spektrum jeder kanonischen[5]) π-selbstadjungierten Erweiterung von A liegt in Δ_A. Die Menge aller nichtreellen Punkte z, die nicht zu Δ_A gehören, bezeichnen wir mit Δ'_A:

$$\Delta'_A = (C_+ \cup C_-) \setminus \Delta_A.$$

Es sei weiter $\mathfrak{G}$ ein Hilbertraum der Dimension n. Wir wählen wie in [4] eine kanonische π-selbstadjungierte Erweiterung $\mathring{A}$ von A und konstruieren eine Operatorfunktion $\mathring{\Gamma}_z$, $z \in \varrho(\mathring{A})$, mit Werten in $[\mathfrak{G}, \Pi_\varkappa]$ und den folgenden Eigenschaften:

1) $\mathring{\Gamma}_z$ bildet $\mathfrak{G}$ eineindeutig und stetig auf $\mathfrak{N}_z$ ab;

2) für beliebige $z, \zeta \in \varrho(\mathring{A})$ gilt mit $\mathring{R}_z = (\mathring{A} - zI)^{-1}$:

$$(1.1) \qquad \frac{\mathring{\Gamma}_z - \mathring{\Gamma}_\zeta}{z - \zeta} = \mathring{R}_z \mathring{\Gamma}_\zeta \quad (= \mathring{R}_\zeta \mathring{\Gamma}_z).$$

Fixieren wir einen Punkt $z_0 \in \varrho(\mathring{A})$, so folgt aus (1. 1) leicht

$$\mathring{\Gamma}_z = (\mathring{A} - z_0 I)(\mathring{A} - zI)^{-1} \mathring{\Gamma}_{z_0} \qquad (z \in \varrho(\mathring{A})).$$

Dabei wählen wir der Einfachheit halber den Punkt z_0 so, daß der Defektraum $\mathfrak{N}_{z_0}$ positiv definit ist (vgl. [4], § 3. 1).

Die Operatorfunktion $\mathring{\Gamma}_z$ ist also durch (1. 1) eindeutig bestimmt bis auf eine Abbildung $\mathring{\Gamma}_z \to \mathring{\Gamma}_z L$, wobei L ein linearer Operator ist, der $\mathfrak{G}$ eineindeutig auf sich abbildet. Wir halten den Operator $\mathring{\Gamma}_{z_0}$ im folgenden stets fest. Aus (1. 1) folgt, daß $\mathring{\Gamma}_z$ eine holomorphe Operatorfunktion in $(C_+ \cup C_-) \setminus \sigma(\mathring{A})$ ist. Der zu $\mathring{\Gamma}_z$ adjungierte Operator $\mathring{\Gamma}_z^+ \in [\Pi_\varkappa, \mathfrak{G}]$ wird durch die Gleichung

$$[\mathring{\Gamma}_z \xi, f] = (\xi, \mathring{\Gamma}_z^+ f) \qquad (\xi \in \mathfrak{G},\ f \in \Pi_\varkappa)$$

definiert. Falls $\mathfrak{N}_z$ nicht entartet (d.h. $z \in \Delta'_A$), so bildet $\mathring{\Gamma}_z^+$ den Teilraum $\mathfrak{N}_z$ eineindeutig auf $\mathfrak{G}$ ab. Für solche Punkte z ist also der selbstadjungierte Operator $\mathring{\Gamma}_z^+ \mathring{\Gamma}_z$ in $\mathfrak{G}$ beschränkt invertierbar, und die Anzahl seiner negativen Eigenwerte stimmt — bei Berücksichtigung ihrer Vielfachheit — mit der Anzahl der negativen Quadrate von $\mathfrak{N}_z$ überein. Für die Punkte $z \in \Delta_A$ hat $\mathring{\Gamma}_z^+ \mathring{\Gamma}_z$ den isolierten, normal abspaltbaren Eigenwert $\lambda = 0$, und die Summe der Vielfachheiten aller seiner nichtpositiven Eigenwerte ist ebenfalls höchstens gleich $\varkappa$.

---

[5]) Eine Erweiterung von A heißt *kanonisch*, wenn sie im Ausgangsraum $\Pi_\varkappa$ wirkt.

2. *Der einfache Teil von A.* Wir nennen einen π-hermiteschen Operator A in $\Pi_\varkappa$ *einfach*, wenn die lineare Hülle seiner Defekträume $\mathfrak{N}_z$, $z\in\Delta'_A$, in $\Pi_\varkappa$ dicht liegt:

(1.2)
$$\Pi_\varkappa = \bigvee_{z\in\Delta'_A} \mathfrak{N}_z.$$

Es sei dem Leser überlassen, sich zu überlegen, daß für einen einfachen π-hermiteschen Operator A sogar $\Pi_\chi = \bigvee_{z\in\mathfrak{U}\setminus\Delta_A} \mathfrak{N}_z$ gilt, wenn $\mathfrak{U}$ eine beliebige offene Menge aus $C_+\cup C_-$ bezeichnet, deren Durchschnitt mit jeder Halbebene $C_\pm$ nicht leer ist. Die obige Definition der Einfachheit eines π-hermiteschen Operators unterscheidet sich von der in [4], § 2. 2 gegebenen. Wir zeigen jedoch am Ende dieses Abschnittes die Äquivalenz beider Definitionen.

Bekanntlich ist ein hermitescher Operator im Hilbertraum die direkte orthogonale Summe eines einfachen hermiteschen Operators und eines selbstadjungierten Operators [14]. Demgegenüber gilt in unserem Falle nur der

Satz 1. 1. *Es sei A ein π-hermitescher Operator in $\Pi_\varkappa$ mit gleichen Defektzahlen. Dann gestattet der Raum $\Pi_\varkappa$ eine Darstellung der Form*

(1.3)
$$\Pi_\varkappa = \mathfrak{N}_e[+](\mathfrak{N}^0\dotplus\mathfrak{N}')[+]\mathfrak{M}';$$ [6])

dabei sind $\mathfrak{N}^0$ und $\mathfrak{N}'$ neutrale, schiefverbundene Teilräume von $\mathfrak{D}(A)$, $\mathfrak{N}_e$ und $\mathfrak{M}'$ sind Pontrjaginräume mit $\varkappa_e$ bzw. $\varkappa'$ negativen Quadraten ($\varkappa = \dim\mathfrak{N}^0+\varkappa_e+\varkappa'$). Der Operator A besitzt bezüglich (1. 3) *die Matrixdarstellung*

(1.4)
$$\begin{pmatrix} A_e & 0 & A_{21}^+ & 0 \\ A_{21} & A_{22} & A_{23} & A_{24} \\ 0 & 0 & A_{22}^+ & 0 \\ 0 & 0 & A_{24}^+ & A' \end{pmatrix},$$

dabei ist A_e ein einfacher π-hermitescher, A' ein π-selbstadjungierter Operator, die Operatoren A_{2j}, $j=1, 2, 3, 4$, sind endlichdimensional, $A_{23}=A_{23}^+$, und es gilt

$$\mathfrak{N}^0 = \bigvee_{z\in\Delta'_{A_e}\setminus\sigma(A_{22})} (A_{22}-zI_{\mathfrak{N}^0})^{-1} A_{21}\mathfrak{N}_z(A_e).$$ [7])

Die Teilräume $\mathfrak{N}_e[+]\mathfrak{N}^0$ und $\mathfrak{N}^0[+]\mathfrak{M}'$ sind durch die angegebenen Eigenschaften eindeutig bestimmt.

Beweis. Wir setzen
$$\mathfrak{M} = \bigcap_{z\in\Delta'_A} \mathfrak{M}_z, \quad \mathfrak{N} = \bigvee_{z\in\Delta'_A} \mathfrak{N}_z;$$

[6]) Dabei können gewisse Komponenten nur aus dem Nullelement bestehen.

[7]) $\mathfrak{N}_z(A_e) = ((A_e-\bar{z}I_{\mathfrak{N}_e})\mathfrak{D}(A_e))^{[\perp]}\cap\mathfrak{N}_e$.

$\mathring{A}$ bezeichne wieder eine kanonische π-selbstadjungierte Erweiterung von A und $\mathring{\Gamma}_z$ die in Abschnitt 1 eingeführte Operatorfunktion mit Werten in $[\mathfrak{G}, \mathfrak{N}_z]$.

Ist $f\in\mathfrak{M}$, d.h. $f[\perp]\mathfrak{N}_z$, $z\in\Delta'_A$, so folgt aus $\mathring{\Gamma}_z-\mathring{\Gamma}_\zeta=(z-\zeta)\mathring{R}_z\mathring{\Gamma}_\zeta$ auch $\mathring{R}_z f[\perp]\mathfrak{N}_\zeta$ für alle $\zeta\in\Delta'_A$, d.h. $\mathring{R}_z f\in\mathfrak{M}$. Zu solchem f gibt es für beliebiges $z\in\Delta'_A$ ein $g\in\mathfrak{D}(A)$ mit $(A-zI)g=(\mathring{A}-zI)g=f$. Daraus folgt $\mathring{R}_z f=g$, also

$$\mathfrak{D}_1 := \mathring{R}_z\mathfrak{M}\subset\mathfrak{D}(A)\cap\mathfrak{M}. \tag{1.5}$$

Ist andererseits $g\in\mathfrak{D}(A)\cap\mathfrak{M}$, $f=(A-zI)g$ für ein $z\in\Delta'_A$, so gilt für $\zeta\in\Delta'_A$:

$$(A-\zeta I)^{-1}f=g+(\zeta-z)\mathring{R}_\zeta g\in\mathfrak{D}(A)\cap\mathfrak{M},$$

also $f=(A-\zeta I)(A-\zeta I)^{-1}f\in\mathfrak{M}_\zeta$. Deshalb gehört auch Ag zu $\mathfrak{M}_\zeta$, und wir erhalten $A(\mathfrak{D}(A)\cap\mathfrak{M})\subset\mathfrak{M}$. Außerdem ist $g=(A-zI)^{-1}f=\mathring{R}_z f\in\mathfrak{D}_1$, also

$$\mathfrak{D}_1=\mathfrak{D}(A)\cap\mathfrak{M}\quad\text{und}\quad(A-zI)\,\mathfrak{D}_1=\mathfrak{M}.$$

Mit der Menge $\mathfrak{M}$ läßt A auch $\mathfrak{N}=\mathfrak{M}^{[\perp]}$ invariant, genauer, es gilt $A(\mathfrak{D}(A)\cap\mathfrak{N})\subset\mathfrak{N}$.

Aus $\mathring{R}_z\mathfrak{M}\subset\mathfrak{M}$, $z\in\Delta'_A$, folgt weiter $\mathring{R}_z\mathfrak{N}\subset\mathfrak{N}$, also bildet $\mathring{R}_z$ insbesondere den endlichdimensionalen Teilraum $\mathfrak{N}^0=\mathfrak{M}\cap\mathfrak{N}$ eineindeutig auf sich ab. Dieser gehört auf Grund von (1.5) zu $\mathfrak{D}(A)$ und wird auch von $A-zI$, $z\in\Delta'_A$, auf sich abgebildet.

Wir wählen jetzt einen mit $\mathfrak{N}^0$ schiefverbundenen neutralen Teilraum $\mathfrak{N}'\subset\mathfrak{D}(A)$ und stellen $\mathfrak{M}$ und $\mathfrak{N}$ in der Form

$$\mathfrak{M}=\mathfrak{N}^0\dotplus\mathfrak{M}',\quad\mathfrak{N}=\mathfrak{N}^0\dotplus\mathfrak{N}_e$$

mit $\mathfrak{M}'[\perp]\mathfrak{N}'$, $\mathfrak{N}_e[\perp]\mathfrak{N}'$ dar. Dann ergibt sich für $\Pi_\varkappa$ die Zerlegung (1.3).

Aus der Invarianz der Teilräume $\mathfrak{M}$ und $\mathfrak{N}$ folgt für den π-hermiteschen Operator A leicht die Matrixdarstellung (1.4) mit π-hermiteschen Operatoren A_e in $\mathfrak{N}_e$ und A' in $\mathfrak{M}'$. Offensichtlich ist $\mathfrak{D}'_1=\mathfrak{D}_1\cap\mathfrak{M}'$ ein dichter Teil von $\mathfrak{M}'$. Außerdem gilt für $z\in\Delta'_A$

$$(A'-zI)\,\mathfrak{D}'_1=\mathfrak{M}',$$

also ist A' π-selbstadjungiert in $\mathfrak{M}'$.

Der Operator A_e ist offensichtlich auf einem dichten Teil von $\mathfrak{N}_e$ definiert, und man sieht leicht, daß f genau dann zum Defektraum $\mathfrak{N}_z$, $z\notin\sigma(A_{22})\cup\sigma(A')$, gehört, wenn es die Gestalt

$$f=f_1-(A_{22}-zI_{\mathfrak{N}^0})^{-1}A_{21}f_1\quad\text{mit}\quad f_1\in\mathfrak{N}_z(A_e) \tag{1.6}$$

hat. Deshalb stimmen die Defektzahlen von A und A_e überein, und es gilt $\Delta'_A \subset \Delta'_{A_e}$ sowie

$$\mathfrak{N}_e = \bigvee_{z \in \Delta'_A} \mathfrak{N}_z(A_e) = \bigvee_{z \in \Delta'_{A_e}} \mathfrak{N}_z(A_e),$$

$$(1.7) \quad \mathfrak{N} = \bigvee_{\substack{z \neq \bar{z} \\ z \notin \sigma(A_{22}) \cup \sigma(A')}} \mathfrak{N}_z; \quad \mathfrak{N}^0 = \bigvee_{z \in \Delta'_{A_e} \setminus \sigma(A_{22})} (A_{22} - zI_{\mathfrak{N}^0})^{-1} A_{21} \mathfrak{N}_z(A_e).$$

Also ist der Operator A_e einfach.

Um die letzte Aussage des Satzes zu beweisen, gehen wir von einer Zerlegung (1. 3) des Raumes $\Pi_\varkappa$ und einer Matrixdarstellung (1. 4) des Operators A mit den angegebenen Eigenschaften aus. Die Elemente f von $\mathfrak{N}_z$ $(z \in \Delta'_A)$ haben wieder die Gestalt (1. 6), also ist $\mathfrak{N} = \bigvee_{z \in \Delta'_A} \mathfrak{N}_z \subset \mathfrak{N}_e[+]\mathfrak{N}^0$. Wäre dabei $\mathfrak{N} \neq \mathfrak{N}_e[+]\mathfrak{N}^0$, so gäbe es Elemente $y_1 \in \mathfrak{N}_e$, $y_3 \in \mathfrak{N}'$ — nicht beide gleich dem Nullelement —, so daß

$$(1.8) \qquad [\mathfrak{N}_z(A_e), y_1 + A_{21}^+(A_{22}^+ - \bar{z}I_{\mathfrak{N}^0})^{-1} y_3] = \{0\}$$

für alle $z \in \Delta'_A$ gelten würde. Betrachten wir diese Beziehung für hinreichend großes $|z|$, so folgt aus der Einfachheit des Operators A_e leicht $y_1 = 0$, also erhält (1. 8) die Form

$$[(A_{22} - zI_{\mathfrak{N}^0})^{-1} A_{21} \mathfrak{N}_z(A_e), y_3] = \{0\} \quad \text{für alle} \quad z \in \Delta'_A.$$

Auf Grund der Voraussetzung $\mathfrak{N}^0 = \bigvee_{z \in \Delta'_{A_e} \setminus \sigma(A_{22})} (A_{22} - zI_{\mathfrak{N}^0})^{-1} A_{21} \mathfrak{N}_z(A_e)$ folgt daraus $y_3 = 0$. Somit gilt $\mathfrak{N} = \mathfrak{N}_e[+]\mathfrak{N}^0$, also ist $\mathfrak{N}_e[+]\mathfrak{N}^0$ eindeutig bestimmt als abgeschlossene lineare Hülle aller Defekträume $\mathfrak{N}_z$ von A, $z \in \Delta'_A$. Dann ist $\mathfrak{N}_0[+]\mathfrak{M}'$ eindeutig bestimmt als π-orthogonales Komplement von $\mathfrak{N}$. Damit ist der Satz bewiesen.

Wir bemerken, daß Satz 1. 1 auch für einen π-hermiteschen Operator A mit ungleichen Defektzahlen richtig bleibt. Dabei stimmen stets die Defektzahlen von A mit denen von A_e überein.

Der Operator A_e aus Satz 1. 1 heißt der *einfache Teil* des Operators A. Er ist durch die angegebenen Eigenschaften bis auf π-unitäre Äquivalenz eindeutig bestimmt: Man kann ihn auffassen als den durch die Einschränkung $A|_{\mathfrak{N}}$ im Faktorraum $\mathfrak{N}/\mathfrak{N}_0$ in natürlicher Weise erzeugten Operator.

Aus Satz 1.1 folgt, daß der π-hermitesche Operator A genau dann einfach ist, wenn er mit seinem einfachen Teil zusammenfällt. Die π-selbstadjungierten Erweiterungen $\tilde{A}$ von A ergeben sich, wenn man in der Darstellung (1. 4) den Operator A_e durch eine π-selbstadjungierte Erweiterung $\tilde{A}_e$ — evtl. mit Austritt in einen Oberraum $\tilde{\mathfrak{N}}_e$ von $\mathfrak{N}_e$ — ersetzt. Deshalb gehören die Punkte von $\sigma(A_{22}) \cup \cup \sigma(A_{22}^+) \cup \sigma(A')$ zum Spektrum jeder π-selbstadjungierten Erweiterung von A.

Schließlich enthält die folgende Aussage die Äquivalenz der hier und in [4], § 2. 2 gegebenen Definitionen der Einfachheit eines π-hermiteschen Operators.

Ein einfacher π-hermitescher Operator hat keinen Eigenwert; jeder π-hermitesche Operator A, der keinen nichtreellen Eigenwert hat und für den $\bigvee_{z \neq \bar{z}} \mathfrak{N}_z = \Pi_\varkappa$ gilt, ist einfach.

Für einen π-hermiteschen Operator A folgt nämlich aus $Af_0 - z_0 f_0 = 0$ leicht $[f_0, g] = 0$, $g \in \mathfrak{N}_z$, $z \in \Delta'_A$, also unter der Voraussetzung (1. 2) $f_0 = 0$. Sind die Voraussetzungen des zweiten Teiles der Aussage erfüllt, so führen wir die zu A gehörige Zerlegung (1. 3) des Raumes $\Pi_\varkappa$ durch. Dann enthält $\sigma(A_{22}) \cup \sigma(A')$ keinen nichtreellen Punkt, also gilt auf Grund der Voraussetzung und der ersten Beziehung von (1. 7) $\mathfrak{N} = \Pi_\varkappa$, d.h., der Operator A stimmt mit seinem einfachen Teil überein.

3. *Zerlegung eines π-selbstadjungierten Operators.* Wir benutzen später mehrmals die folgende Aussage:

Ist A ein π-selbstadjungierter Operator in $\Pi_\varkappa$, so gibt es eine Zerlegung von $\Pi_\varkappa$ als direkte π-orthogonale Summe zweier für A invarianter Teilräume $\Pi'_\varkappa \subset \subset \mathfrak{D}(A)$ und Π_0, $\Pi_\varkappa = \Pi'_\varkappa [+] \Pi_0$, mit den folgenden Eigenschaften: Die Einschränkung $A|_{\Pi'_\varkappa}$ ist ein beschränkter π-selbstadjungierter Operator, Π_0 ist ein Hilbertraum bezüglich des Skalarproduktes $[\cdot, \cdot]$ und die Einschränkung $A_0 = = A|_{\Pi_0}$ ist ein selbstadjungierter Operator in diesem Hilbertraum.

Zum Beweis dieser Aussage betrachten wir die Spektralfunktion E des Operators A ([12], [13]) und wählen ein offenes Intervall Δ, das alle kritischen Punkte von A enthält. Bezeichnet $\mathfrak{C}$ die lineare Hülle aller algebraischen Eigenräume von A zu nichtreellen Eigenwerten, so hat man nur $\Pi'_\varkappa = E(\Delta)\Pi_\varkappa [+] \mathfrak{C}$ und $\Pi_0 = (\Pi'_\varkappa)^{[\perp]}$ zu setzen.

4. *Spektralpunkte nichtpositiven Typs.* Für einen π-selbstadjungierten Operator A in $\Pi_\varkappa$ bezeichne $\sigma_0(A)$ die Menge derjenigen Eigenwerte von A, zu denen ein nichtpositives Eigenelement gehört. Die abgeschlossene lineare Hülle der algebraischen Eigenräume zu den in einem Gebiet $\mathfrak{A}$ der komplexen Ebene gelegenen Eigenwerten von $\sigma_0(A)$ bezeichnen wir mit $\mathfrak{C}_\mathfrak{A}(A)$; ist sign $\mathfrak{C}_\mathfrak{A}(A) = (l_+, l_0, l_-)$, so nennen wir $l_- + l_0$ den *Index* von $\mathfrak{A}$ (bezüglich A); den Index der Menge $\mathfrak{A} = \{\lambda\}$, $\lambda \in \sigma_0(A)$, nennen wir auch den *Index* des Punktes λ. Für nichtreelles $\lambda \in \sigma(A)$ ist sein Index also gleich der algebraischen Vielfachheit, für reelles λ stimmt er mit dem in [13] eingeführten Begriff des Index überein.

Entsprechend definieren wir für einen π-unitären Operator U die Menge $\sigma_0(U)$ sowie den Index einer Menge $\mathfrak{A}$ oder einer Punktes $\lambda \in \sigma_0(U)$ bezüglich U.

Satz 1. 2. *Es sei (U_n) eine Folge π-unitärer Operatoren in $\Pi_\varkappa$ mit $\|U_n - U_0\| \to 0$ $(n \to \infty)$, $\lambda_0 \in \sigma_0(U_0)$ und r_{λ_0} der Index von λ_0. Dann gibt es zu jeder hinreichend kleinen*

Umgebung $\mathfrak{A}$ von λ_0 eine natürliche Zahl n_0, so daß für $n \geqq n_0$ der Index von $\mathfrak{A}$ bezüglich U_n gleich r_{λ_0} ist.

Beweis. Für Punkte λ_0, die nicht auf der Einheitskreislinie liegen, folgt die Behauptung aus einem bekannten Ergebnis der Störungstheorie ([15], Satz 4. 3).

Wir betrachten einen Punkt $\lambda_0 \in \sigma_0(U_0)$ mit $|\lambda_0| = 1$ und eine Umgebung $\mathfrak{A}$ von λ_0, die von $\sigma_0(U_0) \setminus \{\lambda_0\}$ einen positiven Abstand hat. Angenommen, es gäbe eine Teilfolge (n_j) der Folge der natürlichen Zahlen, so daß der Index von $\mathfrak{A}$ bezüglich U_{n_j} kleiner als r_{λ_0} wäre. Den Raum $\Pi_\varkappa$ stellen wir in der Form

$$\Pi_\varkappa = \Pi_- [+] \Pi_+$$

mit einen $\varkappa$-dimensionalen negativen Teilraum Π_- dar; die zugehörige Zerlegung der Operatoren U_n sei

$$U_n = \begin{pmatrix} U_{11}^{(n)} & U_{12}^{(n)} \\ U_{21}^{(n)} & U_{22}^{(n)} \end{pmatrix}, \qquad n = 0, 1, 2, \ldots .$$

Ist $K_n \in [\Pi_-, \Pi_+]$, $\|K_n\| \leqq 1$, ein Winkeloperator für einen Pontrjaginschen $\varkappa$-dimensionalen nichtpositiven invarianten Teilraum von U_n, so gilt

$$\text{(1.9)} \qquad U_{21}^{(n)} + U_{22}^{(n)} K_n = K_n (U_{11}^{(n)} + U_{12}^{(n)} K_n), \qquad n = 1, 2, \ldots$$

Die Folge (K_{n_j}) enthält eine schwach konvergente Teilfolge, von der wir annehmen können, daß sie mit (K_{n_j}) zusammenfällt: $K_{n_j} \to K_0$ (schwach) für $j \to \infty$. Aus (1.9) folgt damit

$$U_{21}^{(0)} + U_{22}^{(0)} K_0 = K_0 (U_{11}^{(0)} + U_{12}^{(0)} K_0),$$

Die charakteristische Gleichung der Einschränkung von U_n auf den zu K_n gehörigen invarianten Teilraum lautet

$$p_n(z) \equiv \det(U_{11}^{(n)} + U_{12}^{(n)} K_n - zI) = 0,$$

deshalb konvergieren die charakteristischen Polynome $p_{n_j}(z)$ gegen $p_0(z)$. Da andererseits jeder Punkt $\lambda_0 \in \sigma_0(U_0)$ vom Index r_{λ_0} auch Nullstelle von $p_0(z)$ der Ordnung r_{λ_0} ist, müssen für hinreichend großes j in $\mathfrak{A}$ genau r_{λ_0} Nullstellen von $p_{n_j}(z)$ liegen. Das widerspricht der obigen Annahme.

Mit Hilfe der Cayleytransformation ergibt sich aus Satz 1.2 ohne Schwierigkeit die

Folgerung 1. 1. *Es sei (A_n) eine Folge π-selbstadjungierter Operatoren in $\Pi_\varkappa$, die in dem Sinne gegen den π-selbstadjungierten Operator A_0 konvergiert, daß für ein nichtreelles $z \in \varrho(A_0)$ auch $z \in \varrho(A_n)$ und $\|(A_n - zI)^{-1} - (A_0 - zI)^{-1}\| \to 0$ $(n \to \infty)$ gilt, $\lambda_0 \in \sigma_0(A_0)$ und r_{λ_0} der Index von λ_0. Dann gibt es zu jeder hinreichend kleinen Umgebung $\mathfrak{A}$ von λ_0 eine natürliche Zahl n_0, so daß für $n \geqq n_0$ der Index von $\mathfrak{A}$ bezüglich A_n gleich r_{λ_0} ist.*

§ 2. Die Q-Funktion eines π-hermiteschen Operators

1. *Definition der Q-Funktion.* Unter einer *Q-Funktion* $Q(z)=Q_A(z)$ des π-hermiteschen Operators A mit gleichen Defektzahlen verstehen wir eine Operatorfunktion mit Werten in $[\mathfrak{G}, \mathfrak{G}]$, die der folgenden Gleichung genügt:

$$\frac{Q(z)-Q^*(\zeta)}{z-\bar{\zeta}} = \mathring{\Gamma}_\zeta^+ \mathring{\Gamma}_z \qquad (z, \zeta \in \varrho(\mathring{A})). \tag{2.1}$$

Die Funktion $Q(z)$ hängt offensichtlich ab von der speziellen gewählten Erweiterung $\mathring{A}$, wir sprechen deshalb mitunter auch von einer zur kanonischen Erweiterung $\mathring{A}$ gehörigen Q-Funktion.

Ist $z_0 \in \varrho(\mathring{A})$ und setzen wir in (2. 1) $z=\zeta=z_0$, so ergibt sich $Q(z_0)=iy_0\mathring{\Gamma}_{z_0}^+\mathring{\Gamma}_{z_0}+C$ mit $y_0=\operatorname{Im} z_0$, $C=C^* \in [\mathfrak{G}, \mathfrak{G}]$. Aus (2. 1) folgt dann für $\zeta=z_0$

$$\begin{aligned} Q(z) &= C-iy_0\mathring{\Gamma}_{z_0}^+\mathring{\Gamma}_{z_0}+(z-\bar{z}_0)\mathring{\Gamma}_{z_0}^+\mathring{\Gamma}_z = \\ &= C-iy_0\mathring{\Gamma}_{z_0}^+\mathring{\Gamma}_{z_0}+(z-\bar{z}_0)\mathring{\Gamma}_{z_0}^+(\mathring{A}-z_0I)(\mathring{A}-zI)^{-1}\mathring{\Gamma}_{z_0} \qquad (z\in\varrho(\mathring{A})). \end{aligned} \tag{2.2}$$

Ohne Schwierigkeit überzeugt man sich davon, daß diese Funktion wirklich der Gleichung (2. 1) genügt. Die Funktion $Q(z)$ ist durch (2. 1) also bis auf einen von z unabhängigen selbstadjungierten Operator $C \in [\mathfrak{G}, \mathfrak{G}]$ bestimmt.

Aus (2. 2) ergibt sich für $z \in \varrho(\mathring{A})$ leicht

$$Q^*(\bar{z}) = Q(z)$$

und

$$\operatorname{Im} Q(z) = \operatorname{Im} z \cdot \mathring{\Gamma}_z^+ \mathring{\Gamma}_z. \tag{2.3}$$

Aus der letzten Gleichung folgt auf Grund der Bemerkungen am Ende von § 1. 1, daß $\lambda=0$ ein isolierter normal abspaltbarer Eigenwert endlicher Vielfachheit oder ein Punkt der Resolventenmenge von $Q(z)$ ist; letzteres trifft sicher dann zu, wenn der Teilraum $\mathfrak{N}_z$ positiv definit ist.

2. *Verallgemeinerte Resolventen.* Die Q-Funktion spielt eine wichtige Rolle bei der Beschreibung aller verallgemeinerten Resolventen des Operators A. Wir erinnern daran ([4]), daß eine *verallgemeinerte Resolvente* von A eine Operatorfunktion R_z mit Werten in $[\Pi_\varkappa, \Pi_\varkappa]$ der Gestalt

$$R_z = \tilde{P}(\tilde{A}-z\tilde{I})^{-1}|_{\Pi_\varkappa} \tag{2.4}$$

ist; dabei bezeichnet $\tilde{A}$ eine beliebige reguläre[8]) π-selbstadjungierte Erweiterung von A in einem Oberraum $\tilde{\Pi}_\varkappa \supset \Pi_\varkappa$ und $\tilde{P}$ den π-orthogonalen Projektor von $\tilde{\Pi}_\varkappa$ auf $\Pi_\varkappa$.

---

[8]) Eine π-selbstadjungierte Erweiterung $\tilde{A}$ von A heißt *regulär*, wenn sie in einem Oberraum $\Pi_{\varkappa 1}$ mit derselben Anzahl negativer Quadrate wie der Ausgangsraum wirkt.

Die verallgemeinerte Resolvente R_z aus (2. 4) heißt eine *kanonische Resolvente* von A, wenn die Erweiterung $\tilde{A}$ kanonisch, d.h. in $\Pi_\varkappa$ gewählt werden kann.

Mit $\mathfrak{T}(\mathfrak{G})$ bezeichnen wir die Menge aller „Operatorfunktionen" T der Gestalt

$$(2.5) \qquad T(z) = \hat{T}(z)\hat{P} + \infty(I-\hat{P}), \qquad z \in C_+ \cup C_-;$$

dabei bezeichnet $\hat{P}$ einen orthogonalen Projektor in $\mathfrak{G}$ und $\hat{T}$ eine Operatorfunktion mit $\hat{T}(z) = \hat{T}^*(\bar{z})$, deren Werte $\hat{T}(z)$ für $z \in C_+$ dicht definierte maximal dissipative Operatoren in $\hat{\mathfrak{G}} = \hat{P}\mathfrak{G}$ sind und deren Cayleytransformierte $V_{\hat{T}}(z) = (\hat{T}(z) - iI)(\hat{T}(z)+iI)^{-1}$ in C_+ holomorph von z abhängt. Gleichungen mit uneigentlichen Operatoren (2. 5) verstehen sich in dem Sinne, daß man in ihnen zunächst ∞ durch eine natürliche Zahl n ersetzt und anschließend $n \to \infty$ streben läßt. Für die Operatorfunktion (2. 5) gilt also z. B.

$$(T(z)+Q(z))^{-1} = \hat{P}(\hat{T}(z)+\hat{Q}(z))^{-1}\hat{P} \qquad (\hat{Q}(z) = \hat{P}Q(z)\hat{P}).$$

Ist $\hat{T}(z)$ aus (2. 5) ein von z unabhängiger selbstadjungierter Operator $\hat{T}$ in $\hat{\mathfrak{G}}$, so nennen wir

$$(2.6) \qquad T = \hat{T}\hat{P} + \infty(I-\hat{P})$$

einen uneigentlichen selbstadjungierten Operator; die Menge aller uneigentlichen selbstadjungierten Operatoren bezeichnen wir mit $\mathfrak{T}_0(\mathfrak{G})$.

Es sei jetzt $\mathring{A}$ eine kanonische π-selbstadjungierte Erweiterung von A. In [4], Satz 5. 1 wurde die folgende Aussage bewiesen:

Zwischen der Menge aller verallgemeinerten Resolventen (2. 4) *des Operators A und der Menge $\mathfrak{T}(\mathfrak{G})$ besteht eine eineindeutige Beziehung, vermittelt durch die Gleichung*

$$(2.7) \qquad R_z = (\mathring{A}-zI)^{-1} - \Gamma_z(T(z)+Q(z))^{-1}\Gamma_{\bar{z}}^+ \qquad (z \notin \sigma(\mathring{A}) \cup \sigma(\tilde{A}));$$

R_z ist genau dann kanonisch, wenn das zugehörige T zu $\mathfrak{T}_0(\mathfrak{G})$ gehört.

Eine π-selbstadjungierte Erweiterung $\tilde{A}$ von A nennen wir eine *reine Austrittserweiterung*, wenn aus $g \in (\mathfrak{D}(\tilde{A}) \cap \Pi_\varkappa) \setminus \mathfrak{D}(A)$ stets $\tilde{A}g \notin \Pi_\varkappa$ folgt. Wie man leicht sieht, ist das gleichbedeutend damit, daß für $f \in \Pi_\varkappa$ die Beziehungen $(\tilde{A}-z_0\tilde{I})^{-1}f \in \Pi_\varkappa$ und $f \in \mathfrak{M}_{z_0}$ äquivalent sind $(z_0 \in \varrho(\tilde{A}))$.

Hat der Operator A einen endlichen Defekt, d.h., ist der Raum $\mathfrak{G}$ endlichdimensional, so braucht man zu den Überlegungen in [4], § 5. 2 wenig hinzuzufügen, um die folgende Aussage zu erhalten:

Lemma 2. 1. *Genau dann ist die π-selbstadjungierte Erweiterung $\tilde{A}$ von A eine reine Austrittserweiterung, wenn für die Funktion $T \in \mathfrak{T}(\mathfrak{G})$ aus der Darstellung* (2. 7) *der zugehörigen verallgemeinerten Resolvente* R_z *von* A *gilt:* $\frac{\operatorname{Im} T(z)}{\operatorname{Im} z} > 0$ *für ein (und damit für alle)* $z \in C_+ \cup C_-$.

Gemäß dem Beweis von Satz 5. 1 aus [4] besteht nämlich für $\tilde{f} = R_z f - (\tilde{A} - z_0 \tilde{I})^{-1} f$ die Beziehung

$$[\tilde{f}, \tilde{f}] = \frac{1}{z_0 - \bar{z}_0} \{(T(z_0)\xi, \xi) - (\xi, T(z_0)\xi)\},$$

wobei wir $\xi = P(z_0)\Gamma_{\bar{z}_0}^+ f$ gesetzt haben[9]. Daraus folgt leicht die Behauptung.

3. *Beschreibung der Gesamtheit aller Q-Funktionen von A.* Bei der Definition der Q-Funktion vermerkten wir, daß diese (bei festem $\mathring{\Gamma}_{z_0}$) von der Wahl der kanonischen Erweiterung $\mathring{A}$ abhängt. Der folgende Satz beschreibt die Gesamtheit aller Q-Funktionen von A, die sich ergibt, wenn $\mathring{A}$ die Gesamtheit aller kanonischen π-selbstadjungierten Erweiterungen von A durchläuft; dabei setzen wir voraus, daß eine spezielle Q-Funktion $Q_0(z)$ des Operators A bekannt ist.

Satz 2. 1. *Ist* $Q_0(z)$ *eine Q-Funktion des π-hermiteschen Operators A, so ergibt sich die Gesamtheit aller Q-Funktionen von A aus der Beziehung*

$$(2.8) \qquad Q(z) = Q_0(z) - (Q_0(z) - Q_0^*(z_0))(T + Q_0(z))^{-1}(Q_0(z) - Q_0(z_0)) + C,$$

wenn T die Menge $\mathfrak{T}_0(\mathfrak{G})$ *aller uneigentlichen selbstadjungierten und C die Menge aller beschränkten selbstadjungierten Operatoren in* $\mathfrak{G}$ *durchläuft.*

Beweis. Die Q-Funktion $Q_0(z)$ gehöre zur kanonischen π-selbstadjungierten Erweiterung $\mathring{A}_0$ von A. Die Resolvente einer beliebigen kanonischen Erweiterung $\mathring{A}$ ergibt sich dann gemäß (2. 7) aus der Beziehung

$$(\mathring{A} - zI)^{-1} = (\mathring{A}_0 - zI)^{-1} - \Gamma_z (T + Q_0(z))^{-1} \Gamma_{\bar{z}}^+$$

mit einem uneigentlichen selbstadjungierten Operator T. Setzen wir diesen Ausdruck in (2. 2) ein und beachten (2. 1), so folgt die Behauptung.

Wir bemerken, daß die Beziehung (2. 8) für alle Punkte $z \in \varrho(\mathring{A}_0) \cap \varrho(\mathring{A})$ besteht. Ist T insbesondere ein beschränkter selbstadjungierter Operator in $\mathfrak{G}$ (d.h. $\hat{P} = I$ und $\hat{T} = \hat{T}^* \in [\mathfrak{G}, \mathfrak{G}]$ in (2. 6)) und setzen wir $C = -T - Q_0(z_0) - Q_0^*(z_0)$, so erhält (2. 8) die Form

$$Q(z) = -(T + Q_0^*(z_0))(T + Q_0(z))^{-1}(T + Q_0(z_0)).$$

Insbesondere ist also mit $Q_0(z)$ auch

$$Q(z) = -Q_0^*(z_0) Q_0^{-1}(z) Q_0(z_0)$$

und, falls speziell $\mathring{\Gamma}_{z_0}^+ \mathring{\Gamma}_{z_0} = I_{\mathfrak{G}}$ gilt, auch $-y_0^2 Q_0^{-1}(z)$ eine Q-Funktion des Operators A.

---

[9] $P(z) = (T(z) + Q(z))^{-1}$.

4. *Charakteristische Eigenschaften der Q-Funktion.* Auf Grund des folgenden Satzes können wir uns häufig auf die Betrachtung der Q-Funktionen einfacher π-hermitescher Operatoren beschränken.

Satz 2. 2. *Jede Q-Funktion des Operators A ist eine Q-Funktion seines einfachen Teiles A_e und umgekehrt.*

Beweis. Jede kanonische π-selbstadjungierte Erweiterung $\mathring{A}$ von A gestattet bezüglich einer Zerlegung (1. 3) von $\Pi_\varkappa$ eine Matrixdarstellung (1. 4), wobei man A_e durch eine π-selbstadjungierte Erweiterung $\mathring{A}_e$ in $\mathfrak{N}_e$ zu ersetzen hat, und umgekehrt definiert jede solche Matrix eine π-selbstadjungierte Erweiterung von A. Die Operatoren $\mathring{\Gamma}_z$ und $\mathring{\Gamma}_{\bar z}^+$ lassen sich bezüglich der Zerlegung (1. 3) folgendermaßen darstellen:

$$\mathring{\Gamma}_z = \begin{pmatrix} \mathring{\Gamma}_z^{(e)} \\ \mathring{\Gamma}_z^0 \\ 0 \\ 0 \end{pmatrix}; \qquad \mathring{\Gamma}_{\bar z}^+ = \left((\mathring{\Gamma}_{\bar z}^{(e)})^+ \ 0 \ (\mathring{\Gamma}_{\bar z}^0)^+ \ 0\right).$$

Offensichtlich hat dabei $\mathring{\Gamma}_z^{(e)} = P_{\mathfrak{N}_e} \mathring{\Gamma}_z$ die Eigenschaften 1) und 2) aus § 1. 1 für A_e an Stelle von A (mit $\mathring{A}_e$ an Stelle von $\mathring{A}$), und es gilt $\mathring{\Gamma}_{z_0}^+ \mathring{\Gamma}_{z_0} = (\mathring{\Gamma}_{z_0}^{(e)})^+ \mathring{\Gamma}_{z_0}^{(e)}$ sowie

$$\mathring{\Gamma}_{z_0}^+ (\mathring{A} - zI)^{-1} \mathring{\Gamma}_{z_0} = (\mathring{\Gamma}_{z_0}^{(e)})^+ (\mathring{A}_e - zI_{\mathfrak{N}_0})^{-1} \mathring{\Gamma}_{z_0}^{(e)}.$$

Damit folgt die Behauptung leicht aus der Beziehung (2. 2).

Wir erinnern ([8], § 1), daß ein in einer Teilmenge $\mathscr{Z}$ der komplexen Ebene definierter Kern $K(z, \zeta)$ mit Werten in $[\mathfrak{G}, \mathfrak{G}]$ ($\mathfrak{G}$-Hilbertraum) definitionsgemäß $\varkappa$ negative Quadrate (in $\mathscr{Z}$) hat, wenn folgendes gilt:

1) $K(z, \zeta) = K^*(\zeta, z)$ $(z, \zeta \in \mathscr{Z})$;
2) für eine beliebige natürliche Zahl n, beliebige $z_1, \dots, z_n \in \mathscr{Z}$ und $\xi_1, \dots, \xi_n \in \mathfrak{G}$ hat die Matrix
$$((K(z_i, z_j)\xi_i, \xi_j))_{i,j=1,2,\dots,n}$$
höchstens $\varkappa$ negative Eigenwerte und für mindestens eine solche Wahl von n; $z_1, \dots, z_n$; $\xi_1, \dots, \xi_n$ genau $\varkappa$ negative Eigenwerte.

Wir sagen, eine Funktion Q mit Werten in $[\mathfrak{G}, \mathfrak{G}]$ habe die *Eigenschaft* $(N_\varkappa^0)$ (bzw. $(N_\varkappa)$), $\varkappa$-nichtnegative ganze Zahl, wenn sie in C_+ meromorph ist (bzw. in $C_+ \cup C_-$ stückweise meromorph ist und $Q(z) = Q^*(\bar z)$ gilt) und der Kern

$$K(z, \zeta) = \frac{Q(z) - Q^*(\zeta)}{z - \bar\zeta} \tag{2. 9}$$

in $\mathscr{Z}_Q^0$ (bzw. $\mathscr{Z}_Q$) $\varkappa$ negative Quadrate hat; dabei bezeichnet $\mathscr{Z}_Q^0$ (bzw. $\mathscr{Z}_Q$) die Menge derjenigen Punkte aus C_+ (bzw. $C_+ \cup C_-$), in denen Q regulär ist. Im Falle $z = \zeta$ hat man die rechte Seite in (2. 9) durch $Q'(z)$ zu ersetzen.

Offensichtlich hat die Funktion Q genau dann z. B. die Eigenschaft $(N_\varkappa^0)$, wenn für beliebiges n, beliebige $z_1, \dots, z_n \in \mathscr{Z}_Q^0$ und $\xi_1, \dots, \xi_n \in \mathfrak{G}$ die quadratische Form

$$(2.10) \qquad \sum_{j,k=1}^{n} \left(\frac{Q(z_j)-Q^*(z_k)}{z_j-\bar{z}_k}\xi_j, \xi_k\right)\alpha_j\bar{\alpha}_k, \quad \alpha_j \in C,$$

nicht mehr als $\varkappa$ negative Quadrate und für mindestens eine Wahl der Zahlen n, z_j und der Vektoren ξ_j genau $\varkappa$ negative Quadrate hat.

Eigentliches Anliegen dieser Arbeit ist der Beweis der folgenden Charakterisierung der Q-Funktionen.

Hauptsatz. *Eine Funktion $Q(z)$ mit Werten in $[\mathfrak{G}, \mathfrak{G}]$ ist genau dann Q-Funktion eines einfachen π-hermiteschen Operators A in $\Pi_\varkappa$, wenn sie den folgenden Bedingungen genügt:*

(I) $Q(z)$ *hat die Eigenschaft* $(N_\varkappa)$;

(II) $\underset{y\uparrow\infty}{w\text{-lim}} \dfrac{Q(iy)}{y} = O$,[10] d.h. $\lim\limits_{y\uparrow\infty} \dfrac{(Q(iy)\xi, \xi)}{y} = 0$ *für alle* $\xi \in \mathfrak{G}$;

(III) $\lim\limits_{y\uparrow\infty} y(\operatorname{Im} Q(iy)\xi, \xi) = \infty$ *für alle* $\xi \in \mathfrak{G}$, $\xi \neq 0$;

(IV) *für mindestens ein nichtreelles z ist* $\operatorname{Im} Q(z)$ *gleichmäßig positiv.*

Wir bemerken, daß im Falle eines endlichdimensionalen Raumes $\mathfrak{G}$ mit Bedingung (III) stets auch Bedingung (IV) erfüllt ist.

Die einzelnen Aussagen dieses Satzes werden — zumeist in schärferer Form — in den folgenden Paragraphen bewiesen. Die Notwendigkeit der Bedingungen (I) und (II) ergibt sich aus § 3. 3 und die Notwendigkeit der Bedingung (III) aus Satz 3. 2. Die Notwendigkeit von (IV) folgt aus (2. 3) und den Bemerkungen am Ende von § 1. 1. Schließlich folgt die Hinlänglichkeit der angegebenen Bedingungen aus Satz 5. 1.

§ 3. Die Funktionenklasse $N_\varkappa^0(\mathfrak{G})$

1. *Der Raum $\Pi_\varkappa^0(Q)$.* Es sei jetzt $\mathfrak{G}$ ein beliebiger Hilbertraum, Q eine in der offenen oberen Halbebene C_+ meromorphe Funktion mit Werten in $[\mathfrak{G}, \mathfrak{G}]$. Hat diese Funktion Q überdies die Eigenschaft $(N_\varkappa^0)$, so erzeugt sie in natürlicher Weise einen $\pi_\varkappa$-Raum $\Pi_\varkappa^0(Q)$. Um das zu sehen, ordnen wir jedem Punkte z des Holomorphiegebietes $\mathscr{Z}_0^Q$ von Q ein Symbol e_z zu und bilden die lineare Menge $\mathfrak{L}^0(Q)$ aller formalen Summen

$$(3.1) \qquad f = \Sigma e_z \xi_z, \quad \xi_z \in \mathfrak{G}, \quad z \in \mathscr{Z}_Q^0,$$

[10] w-lim bezeichnet den Grenzwert im schwachen Sinne.

wobei nur endlich viele „Koeffizienten" ξ_z vom Nullelement verschieden sein sollen. Für zwei solche Elemente f und

$$g = \Sigma e_z \eta_z \in \mathfrak{L}^0(Q)$$

definieren wir ein (möglicherweise entartendes) Skalarprodukt $[f, g]$ durch die Gleichung

$$(3.2) \qquad [f, g] = \sum_{z, \zeta \in \mathscr{Z}_Q^0} \left(\frac{Q(z) - Q^*(\zeta)}{z - \bar{\zeta}} \xi_z, \eta_\zeta\right).$$

Auf Grund der Bedingung $(N_\varkappa^0)$ hat dieses Skalarprodukt genau $\varkappa$ negative Quadrate. Durch Faktorraumbildung $\mathfrak{L}^0(Q)/\mathfrak{J}$ nach der Menge $\mathfrak{J} = \{f: f \in \mathfrak{L}^0(Q), f[\perp]\mathfrak{L}^0(Q)\}$ aller isotropen Elemente von $\mathfrak{L}^0(Q)$ und Vervollständigung erhalten wir einen $\pi_\varkappa$-Raum $\Pi_\varkappa^0(Q)$, in den $\mathfrak{L}^0(Q)$ kanonisch eingebettet ist ([4], § 1. 2). Wir erinnern daran, daß bei dieser kanonischen Einbettung das Skalarprodukt $[f, g]$ zweier Elemente $f, g \in \mathfrak{L}^0(Q)$ sowie die Konvergenz einer Folge von Elementen aus $\mathfrak{L}^0(Q)$[11]) erhalten bleiben; dabei liegt das Bild von $\mathfrak{L}^0(Q)$ dicht in $\Pi_\varkappa^0(Q)$.

2. *Der Operator* A_Q^0. Für das Element $f = \sum e_z \xi_z \in \mathfrak{L}^0(Q)$ setzen wir

$$\chi(f) = \sum_{z \in \mathscr{Z}_Q^0} \xi_z.$$

Dann ist χ ein linearer Operator von $\mathfrak{L}^0(Q)$ auf $\mathfrak{G}$. Wir führen die Menge

$$\hat{\mathfrak{D}} = \{f : f \in \mathfrak{L}^0(Q), \chi(f) = 0\}$$

ein und definieren auf $\hat{\mathfrak{D}}$ einen Operator $\hat{A}$ durch die Gleichung

$$\hat{A} f = \sum_{z \in \mathscr{Z}_Q^0} z e_z \xi_z \quad \text{für} \quad f = \sum_{z \in \mathscr{Z}_Q^0} e_z \xi_z \in \hat{\mathfrak{D}},$$

Man zeigt leicht, daß $\hat{A}$ π-hermitesch ist:

$$[\hat{A} f, g] = [f, \hat{A} g] \qquad (f, g \in \hat{\mathfrak{D}}).$$

Wesentlich mehr läßt sich über den Operator $\hat{A}$ aussagen, falls die Funktion Q außer $(N_\varkappa^0)$ noch die folgende Eigenschaft $(\mathfrak{D})$ hat:

$(\mathfrak{D})$ *Es gibt eine Punktfolge* $(z_n) \subset \mathscr{Z}_Q^0$ *mit*

$$(3.3) \qquad \lim_{n \to \infty} \operatorname{Im} z_n = \infty \quad \textit{und} \quad w\text{-}\lim_{n \to \infty} \frac{Q(z_n)}{\operatorname{Im} z_n} = 0.$$

Da für $\xi, \eta \in \mathfrak{G}$ und $z \in \mathscr{Z}_Q^0$

$$[e_{z_n} \xi, e_{z_n} \xi] = \frac{\operatorname{Im}(Q(z_n)\xi, \xi)}{\operatorname{Im} z_n}, \quad [e_{z_n} \xi, e_z \eta] = \left(\frac{Q(z_n) - Q^*(z)}{z_n - \bar{z}} \xi, \eta\right)$$

[11]) Die Konvergenz $f_n \to g$ einer Folge $(f_n) \subset \mathfrak{L}^0(Q)$ gegen $g \in \mathfrak{L}^0(Q)$ wird dabei folgendermaßen erklärt (siehe [4], § 1.1): $[f_n, f_n] \to [g, g]$ und $[f_n, h] \to [g, h]$ für alle $h \in \mathfrak{L}^0(Q)$.

gilt, folgt aus (3. 3) für beliebiges $\xi \in \mathfrak{G}$

$$e_{z_n}\xi \to 0 \quad \text{für} \quad n \to \infty.$$

Daraus ergibt sich, daß $\hat{\mathfrak{D}}$ in $\mathfrak{L}^0(Q)$ dicht liegt: Für beliebiges $f \in \mathfrak{L}^0(Q)$ gilt nämlich $g_n = f - e_{z_n}\chi(f) \in \hat{\mathfrak{D}}$ und $g_n \to f$. Ist $h \in \mathfrak{I} \cap \hat{\mathfrak{D}}$, so gilt für alle $f \in \hat{\mathfrak{D}}$

$$[\hat{A}h, f] = [h, \hat{A}f] = 0,$$

also gehört auch $\hat{A}h$ zu $\mathfrak{I}$. Deshalb geht bei der kanonischen Einbettung von $\mathfrak{L}^0(Q)$ in $\Pi_\varkappa^0(Q)$ der Operator $\hat{A}$ in einen dicht definierten hermiteschen Operator in $\Pi_\varkappa^0(Q)$ über, dessen Abschließung wir mit A_Q^0 bezeichnen.

Wir überlegen uns, daß $\mathfrak{n}_+(A_Q^0) = 0$ gilt. Zu diesem Zweck betrachten wir einen Punkt $z_0 \in \mathscr{Z}_Q^0$. Aus der Stetigkeit der Funktion Q im Punkte z_0 folgt durch Betrachtung der Ausdrücke

$$[e_z\xi - e_{z_0}\xi, e_z\xi - e_{z_0}\xi] \quad \text{und} \quad [e_z\xi - e_{z_0}\xi, e_\zeta\eta] \quad (z, \zeta \in \mathscr{Z}_Q^0; \xi, \eta \in \mathfrak{G})$$

unmittelbar, daß $z \to z_0$ für beliebiges $\xi \in \mathfrak{G}$ auch $e_z\xi \to e_{z_0}\xi$ nach sich zieht. Andererseits gilt für $\sum e_z\xi_z \in \mathfrak{L}^0(Q)$ mit $\xi_{z_0} = 0$:

$$f = \sum \frac{e_z\xi_z - e_{z_0}\xi_z}{z - z_0} \in \hat{\mathfrak{D}}, \qquad (\hat{A} - z_0 I)f = \sum e_z\xi_z.$$

Die Menge $(\hat{A} - z_0 I)\hat{\mathfrak{D}}$ liegt somit dicht in $\mathfrak{L}^0(Q)$, also ist auch $(A_Q^0 - z_0 I)\mathfrak{D}(A_Q^0)$ dicht in $\Pi_\varkappa^0(Q)$ und fällt deshalb sogar mit $\Pi_\varkappa^0(Q)$ zusammen.

Wir wählen jetzt einen Punkt $z_0 \in \mathscr{Z}_Q^0 \setminus \sigma_p(A_Q^0)$ und definieren einen linearen Operator $\Gamma_{z_0}: \mathfrak{G} \to \mathfrak{L}^0(Q)$ durch die Gleichung

$$\Gamma_{z_0}\xi = e_{z_0}\xi \qquad (\xi \in \mathfrak{G}).$$

Da für beliebige $\xi, \eta \in \mathfrak{G}$ und $z \in \mathscr{Z}_Q^0$

$$(3.4) \qquad [\Gamma_{z_0}\xi, \Gamma_{z_0}\xi] = \frac{\operatorname{Im}(Q(z_0)\xi, \xi)}{\operatorname{Im} z_0}, \quad [\Gamma_{z_0}\xi, e_z\eta] = \left(\frac{Q(z_0) - Q^*(z)}{z_0 - \bar{z}}\xi, \eta\right)$$

gilt, ist der Operator Γ_{z_0} stetig. Offensichtlich können wir ihn auch auffassen als Operator von $\mathfrak{G}$ in $\Pi_\varkappa^0(Q)$, d.h. als Element von $[\mathfrak{G}, \Pi_\varkappa^0(Q)]$. Dann hat er einen (stetigen) π-adjungierten Operator $\Gamma_{z_0}^+ \in [\Pi_\varkappa^0(Q), \mathfrak{G}]$, definiert durch die Gleichung

$$[f, \Gamma_{z_0}\xi] = (\Gamma_{z_0}^+ f, \xi) \qquad (f \in \Pi_\varkappa^0(Q), \xi \in \mathfrak{G}).$$

Aus den Beziehungen (3. 4) folgt insbesondere mit $y_0 = \operatorname{Im} z_0$

$$(3.5) \qquad y_0 \Gamma_{z_0}^+ \Gamma_{z_0} = \operatorname{Im} Q(z_0), \quad \Gamma_{z_0}^+(e_z\xi) = \frac{Q(z) - Q^*(z_0)}{z - \bar{z}_0}\xi.$$

Da für $z\in\mathscr{X}_Q^0$, $z\neq z_0$ und $\xi\in\mathfrak{G}$

$$(3.6)\qquad (\hat{A}-zI)(e_z\xi-e_{z_0}\xi)=(z-z_0)e_{z_0}\xi$$

gilt, gehört $e_{z_0}\xi$ zu $\mathfrak{R}(\hat{A}-zI)$, also ist $\mathfrak{R}(\Gamma_{z_0})\subset\mathfrak{R}(A_Q^0-zI)$ und

$$(3.7)\qquad e_z\xi=e_{z_0}\xi+(z-z_0)(A_Q^0-zI)^{-1}(e_{z_0}\xi)=(A_Q^0-z_0I)(A_Q^0-zI)^{-1}(e_{z_0}\xi).$$

Wenden wir auf diese Beziehung den Operator $\Gamma_{z_0}^+$ an und beachten die zweite Gleichung von (3. 5), so folgt

$$(3.8)\qquad Q(z)=Q^*(z_0)+(z-\bar{z}_0)\Gamma_{z_0}^+(A_Q^0-z_0I)(A_Q^0-zI)^{-1}\Gamma_{z_0},$$

und zwar zunächst nur für $z\neq z_0$; für $z=z_0$ ist diese Beziehung aber nichts anderes als die erste Gleichung von (3. 5).

3. *Die Klasse* $N_\varkappa^0(\mathfrak{G})$. Mit $N_\varkappa^0(\mathfrak{G})$ bezeichnen wir die Gesamtheit aller in C_+ meromorphen Funktionen Q, deren Werte in $[\mathfrak{G},\mathfrak{G}]$ liegen und welche die Eigenschaften $(N_\varkappa^0)$ und $(\mathfrak{D})$ besitzen.

Wir nennen im folgenden einen π-selbstadjungierten (bzw. maximalen π-hermiteschen, $\mathfrak{n}_+(A)=0$) Operator A in $\Pi_\varkappa$ *eng verbunden* (bzw. *eng o-verbunden*[12])) mit einem Operator $\Gamma\in[\mathfrak{G},\Pi_\varkappa]$, wenn

$$\Pi_\varkappa=\bigvee_{z\in\varrho(A)}(A-zI)^{-1}\Gamma\mathfrak{G}\quad\Bigl(\text{bzw. }\Pi_\varkappa=\bigvee_{z\in\varrho(A)\cap C_+}(A-zI)^{-1}\Gamma\mathfrak{G}\Bigr)$$

gilt

Einen wichtigen Zusammenhang zwischen maximalen π-hermiteschen Operatoren und Funktionen der Klasse $N_\varkappa^0(\mathfrak{G})$ stellt der folgende Satz her.

Satz 3. 1. *Es sei A ein maximaler π-hermitescher Operator mit $\mathfrak{n}_+(A)=0$ in einem $\pi_\varkappa$-Raum $\Pi_\varkappa$, $\mathfrak{G}$ ein Hilbertraum, $S=S^*\in[\mathfrak{G},\mathfrak{G}]$, $\Gamma\in[\mathfrak{G},\Pi_\varkappa]$ und $z_0\in C_+\setminus\sigma_p(A)$. Dann definiert die Gleichung*

$$(3.9)\qquad Q(z)=S-iy_0\Gamma^+\Gamma+(z-\bar{z}_0)\Gamma^+(A-z_0I)(A-zI)^{-1}\Gamma$$

eine Funktion der Klasse $N_{\varkappa'}^0(\mathfrak{G})$ für ein gewisses $\varkappa'$ mit $0\leqq\varkappa'\leqq\varkappa$. Sind die Operatoren A und Γ eng o-verbunden, so ist $\varkappa'=\varkappa$. Umgekehrt gestattet jede Funktion $Q\in N_\varkappa^0(\mathfrak{G})$ die Darstellung (3. 9) *mit $\Pi_\varkappa=\Pi_\varkappa^0(Q)$, dem maximalen π-hermiteschen Operator $A=A_Q^0$ mit $\mathfrak{n}_+(A)=0$, $S=\operatorname{Re}Q(z_0)$ und $\Gamma=\Gamma_{z_0}$; dabei sind die Operatoren A_Q^0 und Γ_{z_0} eng o-verbunden.*

Beweis. Zum Beweis der zweiten Aussage bleibt nach den Betrachtungen in den Abschnitten 1 und 2 nur noch zu zeigen, daß die Operatoren A_Q^0 und Γ eng o-verbunden sind, d.h., daß die abgeschlossene lineare Hülle $\mathfrak{E}$ aller Vektoren $(A_Q^0-zI)^{-1}(e_{z_0}\xi)$ $\bigl(z\in C_+\setminus\sigma_p(A_Q^0),\ \xi\in\mathfrak{G}\bigr)$ mit $\Pi_\varkappa^0(Q)$ zusammenfällt.

12) d.h. eng verbunden bezüglich der oberen Halbebene.

Erweitert man A_Q^0 gegebenenfalls zu einem π-selbstadjungierten Operator $\tilde{A}$ in einem Oberraum $\tilde{\Pi}_\varkappa \supset \Pi_\varkappa^0(Q)$ und zerlegt diesen gemäß § 1. 3, so überzeugt man sich leicht von der Beziehung

$$\lim_{y \uparrow \infty} iy(iyI - A_Q^0)^{-1} f = f \qquad (f \in \Pi_\varkappa).$$

Deshalb gilt $e_{z_0}\xi \in \mathfrak{E}$, $\xi \in \mathfrak{G}$, und damit auch

$$e_z \xi = \big(I + (z - z_0)(A_Q^0 - zI)^{-1}\big)(e_{z_0}\xi) \in \mathfrak{E} \qquad (z \in \mathscr{Z}_Q^0,\ \xi \in \mathfrak{G}),$$

d.h., es ist $\mathfrak{E} = \Pi_\varkappa^0(Q)$.

Zum Beweis der ersten Aussage des Satzes setzen wir ohne Beschränkung der Allgemeinheit voraus, daß der Operator A in (3. 9) sogar π-selbstadjungiert ist; anderenfalls können wir ihn nämlich zu einem π-selbstadjungierten Operator $\tilde{A}$ in einem Oberraum $\tilde{\Pi}_\varkappa \supset \Pi_\varkappa$ erweitern, und die Gleichung (3. 9) besteht dann noch für $\tilde{A}$ an Stelle von A, wenn man nun auch Γ auffaßt als Abbildung von $\mathfrak{G}$ in $\tilde{\Pi}_\varkappa$ und Γ^+ als Abbildung von $\tilde{\Pi}_\varkappa$ in $\mathfrak{G}$: $\Gamma^+ f = 0$ für $f \in \tilde{\Pi}_\varkappa[-]\Pi_\varkappa$.

Wir extrapolieren die Funktion Q auf die Menge $C_- \setminus \sigma_p(A)$ durch die Festsetzung

$$Q(\bar{z}) = Q^*(z), \quad z \in C_+ \setminus \sigma_p(A).$$

Dann besteht die Gleichung (3. 9), wie man leicht verifiziert, für alle nichtreellen Punkte $z \notin \sigma_p(A)$. Für solche Punkte z schreiben wir sie in der Form

$$Q(z) = S + (z - x_0)\Gamma^+\Gamma + q(z)\Gamma^+ R_z \Gamma \qquad (z_0 = x_0 + iy_0)$$

mit $q(z) = (z - z_0)(z - \bar{z}_0)$, $R_z = (A - zI)^{-1}$, woraus

$$\frac{Q(z) - Q^*(\zeta)}{z - \bar{\zeta}} = \Gamma^+ \left(I + \frac{q(z) R_z - q(\bar{\zeta}) R_{\bar{\zeta}}}{z - \bar{\zeta}} \right) \Gamma \qquad (z, \zeta \in (C_+ \cup C_-) \setminus \sigma_p(A),\ z \neq \bar{\zeta})$$

folgt. Mit Hilfe der Beziehung $R_z - R_{\bar{\zeta}} = (z - \bar{\zeta}) R_z R_{\bar{\zeta}}$ können wir den in den runden Klammern stehenden Ausdruck folgendermaßen umformen:

$$I + \frac{q(z) - q(\bar{\zeta})}{z - \bar{\zeta}} R_z + q(\bar{\zeta}) R_z R_{\bar{\zeta}} = U_{z_0 \zeta}^+ U_{z_0 z},$$

wobei $U_{z_0 z} = I + (z - z_0) R_z = (A - z_0 I)(A - zI)^{-1}$ gesetzt wurde. Damit ergibt sich

$$(3.10) \qquad \frac{Q(z) - Q^*(\zeta)}{z - \bar{\zeta}} = \Gamma^+ U_{z_0 \zeta}^+ U_{z_0 z} \Gamma \qquad (z, \zeta \in (C_+ \cup C_-) \setminus \sigma_p(A),\ z \neq \bar{\zeta}).$$

Man sieht leicht, daß diese Beziehung auch für $z = \bar{\zeta}$ richtig bleibt, wenn man nur die linke Seite als $Q'(z)$ interpretiert.

Aus der Beziehung (3. 10) folgt das Vorliegen der Eigenschaft $(N_{\varkappa'}^0)$ für ein gewisses $\varkappa'$, $0 \leqq \varkappa' \leqq \varkappa$. In der Tat, für $z_j \in C_+ \setminus \sigma_p(A)$, $\xi_j \in \mathfrak{G}$ $(j = 1, 2, \ldots, n;\ n = 1, 2, \ldots)$

fällt die Form (2. 10) der komplexen Variablen α_j auf Grund von (3. 10) zusammen mit der Form

$$(3.11)\qquad \left[\sum_{j=1}^{n}\alpha_j f_j,\ \sum_{k=1}^{n}\alpha_k f_k\right] = \sum_{j,k=1}^{n}\alpha_j\bar{\alpha}_k[f_j, f_k], \qquad f_j = U_{z_0 z_j}\Gamma\xi_j,$$

die offensichtlich höchstens $\varkappa$ negative Quadrate hat.

Wir bemerken, daß sich in der gleichen Weise für die auf $(C_+\cup C_-)\setminus\sigma_p(A)$ extrapolierte Funktion Q sogar die Eigenschaft $(N_{\varkappa'})$ für ein gewisses $\varkappa'$, $0\leqq\varkappa'\leqq\varkappa$, ergibt.

Sind die Operatoren A und Γ eng o-verbunden, d.h., liegt die lineare Hülle der Vektoren $U_{z_0 z}\Gamma\xi$ $(\xi\in\mathfrak{G}, z\in C_+\setminus\sigma_p(A))$ dicht in $\Pi_\varkappa$, so gibt es offensichtlich stets solche $z_j\in C_+\setminus\sigma_p(A)$ und $\xi_j\in\mathfrak{G}$ $(j=1, 2, \ldots, n)$, daß die Form (3. 11) genau $\varkappa$ negative Quadrate hat.

Wir haben noch zu zeigen, daß die Funktion Q auch die Eigenschaft $(\mathfrak{D})$ besitzt. Statt dessen beweisen wir die folgende schärfere Aussage; dabei bezeichne W_Θ für beliebiges Θ, $0<\Theta<\frac{\pi}{2}$, denjenigen Winkelraum von C_+, dessen Punkte z durch die Ungleichung $\left|\arg z-\frac{\pi}{2}\right|\leqq\Theta$ $(z\in C_+)$ charakterisiert sind.

Für beliebiges Θ, $0<\Theta<\frac{\pi}{2}$, genügt die durch (3. 9) *definierte Operatorfunktion $Q(z)$ der Bedingung*

$$(3.12)\qquad w-\lim_{\operatorname{Im} z\uparrow\infty}\left.\frac{Q(z)}{\operatorname{Im} z}\right|_{z\in W_\Theta} = 0.$$

Gemäß (3. 9) ist die Beziehung (3. 12) äquivalent der folgenden:

$$w-\lim_{\operatorname{Im} z\uparrow\infty}\Gamma^+(A-z_0I)(A-zI)^{-1}\Gamma|_{z\in W_\Theta} = 0.$$

Wir setzen wieder voraus, daß der Operator A sogar π-selbstadjungiert ist, und zerlegen $\Pi_\varkappa$ gemäß § 1. 3 in die direkte π-orthogonale Summe zweier für A invarianter Teilräume $\Pi'_\varkappa$ und Π_0. Dann folgt

$$(3.13)\qquad \Gamma^+U_{z_0 z}\Gamma = \Gamma^+(A'-z_0I)(A'-zI)^{-1}P'\Gamma+\Gamma^+(A_0-z_0I)(A_0-zI)^{-1}P_0\Gamma,$$

wobei P' und $P_0 = I-P'$ die π-orthogonalen Projektoren von $\Pi_\varkappa$ auf $\Pi'_\varkappa$ bzw. Π_0 und A' bzw. A_0 die Einschränkungen von A auf $\Pi'_\varkappa$ bzw. Π_0 bezeichnen. Da der Operator A' beschränkt ist, verhält sich der erste Summand der rechten Seite von (3. 13) für $|z|\to\infty$ wie $O\left(\frac{1}{z}\right)$. Deshalb bleibt das Verhalten der Funktion

$$F(z;\xi,\eta) = [(A_0-z_0I)(A_0-zI)^{-1}P_0\xi, \Gamma\eta], \quad \xi,\eta\in\mathfrak{G},$$

im Winkelraum W_Θ zu betrachten.

Die Spektralzerlegung für die Resolvente von A_0:

(3.14) $$(A_0 - zI)^{-1} = \int_{-\infty}^{\infty} \frac{dE_\lambda}{\lambda - z}$$

gestattet es, die Funktion $F(z; \xi, \eta)$ in der Form

$$F(z; \xi, \eta) = \int_{-\infty}^{\infty} \frac{\lambda - z_0}{\lambda - z} d\sigma_{\xi,\eta}(\lambda)$$

darzustellen, wobei $\sigma_{\xi,\eta}(\lambda) = [E_\lambda P_0 \Gamma\xi, P_0 \Gamma\eta]$ eine Funktion von beschränkter Variation auf der reellen Achse ist.

Man sieht leicht, daß immer eine Konstante $\gamma_\Theta > 0$ existiert, so daß für $y = \operatorname{Im} z \geqq \gamma_\Theta$, $z \in W_\Theta$ gilt: $|(\lambda - z_0)(\lambda - z)^{-1}| < 1$. Andererseits ist für beliebiges $\gamma > 0$ offensichtlich

$$\lim_{\operatorname{Im} z \uparrow \infty} |(\lambda - z_0)(\lambda - z)^{-1}| |_{z \in W_\Theta, -\gamma \leqq \lambda \leqq \gamma} = 0.$$

Daraus folgt

$$\lim_{\operatorname{Im} z \uparrow \infty} F(z; \xi, \eta)|_{z \in W_\Theta} = 0,$$

womit die obige Aussage bewiesen ist.

Folgerung 3.1. *Jede Operatorfunktion* $Q \in N_\varkappa^0(\mathfrak{G})$ *hat die Eigenschaft* (3.12) *in einem beliebigen Winkelraum* W_Θ $\left(0 < \Theta < \frac{\pi}{2}\right)$.

Aus der Daratellung (3.9) der Funktion $Q \in N_\varkappa^0(\mathfrak{G})$ ergibt sich unmittelbar, daß diese höchstens $\varkappa$ Pole in der oberen Halbebene C_+ hat. Eine Untersuchung der zugehörigen Hauptteile erfolgt in § 4.3.

Satz 3.2. *In der Darstellung* (3.9) *der Operatorfunktion* $Q \in N_\varkappa^0(\mathfrak{G})$ *besteht die Beziehung*

$$\mathfrak{D}(A) \cap \Gamma\mathfrak{G} = \{0\}$$

genau dann, wenn für beliebiges $\xi \in \mathfrak{G}$, $\xi \neq 0$, *gilt:*

(3.15) $$\lim_{y \uparrow \infty} y \operatorname{Im}(Q(iy)\xi, \xi) = \infty.$$

Beweis. Wir setzen wieder A als π-selbstadjungiert voraus und benutzen die Zerlegung aus § 1.3: $\Pi_\varkappa = \Pi'_\varkappa [+] \Pi_0$, $A' = A|_{\Pi'_\varkappa}$, $A_0 = A|_{\Pi_0}$. Dann gilt für beliebiges $\xi \in \mathfrak{G}$, $\xi \neq 0$ und $z \in C_+ \setminus \sigma_p(A)$

(3.16) $$[U_{z_0 z}\Gamma\xi, U_{z_0 z}\Gamma\xi] = [U_{z_0 z}^{(0)} P_0 \Gamma\xi, U_{z_0 z}^{(0)} P_0 \Gamma\xi] + [U'_{z_0 z} P' \Gamma\xi, U'_{z_0 z} P' \Gamma\xi],$$

wobei $U_{z_0 z}^{(0)} = (A_0 - z_0 I)(A_0 - zI)^{-1}$, $U'_{z_0 z} = (A' - z_0 I)(A' - zI)^{-1}$ gesetzt wurde. Da der Operator A' beschränkt ist, verhält sich der zweite Summand auf der rechten Seite von (3.16) für $|z| \to \infty$ wie $O\left(\frac{1}{z}\right)$.

Bei Beachtung der Beziehung (3. 10) für $z=\zeta=iy$ und der Spektralzerlegung (3. 14) ergibt sich leicht

$$\operatorname{Im}(Q(iy)\xi,\xi)=y\int_{-\infty}^{\infty}\frac{|\lambda-z_0|^2}{\lambda^2+y^2}\,d\sigma_\xi(\lambda)+O\left(\frac{1}{y}\right)\qquad(y\uparrow\infty)$$

mit der nichtabnehmenden beschränkten Funktion $\sigma_\xi(\lambda)=[E_\lambda P_0\Gamma\xi,\,P_0\Gamma\xi]$. Deshalb ist die Gleichung (3. 15) für festes $\xi\in\mathfrak{G}$ äquivalent der Gleichung

$$\int_{-\infty}^{\infty}\lambda^2\,d\sigma_\xi(\lambda)=\infty. \tag{3.17}$$

Letztere besagt aber $P_0\Gamma\xi\notin\mathfrak{D}(A_0)$, und da $\mathfrak{D}(A')=\Pi'_\varkappa$, also $P'\Gamma\xi\in\mathfrak{D}(A')$ gilt, ist die Bedingung (3. 17) äquivalent mit $\Gamma\xi\notin\mathfrak{D}(A)$.

Bemerkung. Man sieht leicht, daß Satz 3. 2 richtig bleibt, wenn man darin die Bedingung (3. 15) durch die Bedingung

$$\lim_{\operatorname{Im} z\uparrow\infty}\operatorname{Im} z\cdot\operatorname{Im}(Q(z)\xi,\xi)\big|_{z\in W_\Theta}=\infty$$

mit einem gewissen festen Θ, $0<\Theta<\frac{\pi}{2}$, ersetzt.

4. *Eindeutigkeit der Darstellung* (3. 9). Zwei Operatoren A und A', die in $\pi_\varkappa$-Räumen $\Pi_\varkappa$ bzw. $\Pi'_\varkappa$ wirken, heißen *π-unitär äquivalent*, wenn eine Abbildung T existiert, die $\Pi_\varkappa$ π-unitär (d.h. unter Erhaltung des π-Skalarproduktes) auf $\Pi'_\varkappa$ abbildet[12]), so daß $T\mathfrak{D}(A)=\mathfrak{D}(A')$ und $A'Tf=TAf$ $(f\in\mathfrak{D}(A))$ gilt.

Wir haben nur wenige Bemerkungen hinzuzufügen, um die folgende wesentliche Ergänzung zu Satz 3. 1 zu erhalten.

Satz 3. 3. *In der Darstellung* (3. 9) *der Operatorfunktion $Q\in N^0_\varkappa(\mathfrak{G})^r$ mit eng o-verbundenen Operatoren A und Γ ist der maximale π-hermitesche Operator A bis auf π-unitäre Äquivalenz eindeutig bestimmt.*

Beweis. Ist Q in der Form (3. 9) mit eng o-verbundenen Operatoren A und Γ dargestellt, dann ist $Q(z)$ in $z=z_0$ regulär und es besteht auch die Darstellung (3. 8). Für beliebiges $z\in C_+\setminus\sigma_p(A)$, $\xi\in\mathfrak{G}$, setzen wir

$$T_0(A-z_0I)(A-zI)^{-1}\Gamma\xi=(A^0_Q-z_0I)(A^0_Q-zI)^{-1}\Gamma_{z_0}\xi\,(=e_z\xi).$$

---

[13]) Ein solcher Operator T bildet $\Pi_\varkappa$ sogar linear, stetig und eineindeutig auf $\Pi'_\varkappa$ ab. Man sieht auch leicht, daß jeder Operator T_0, der eine Menge $\mathfrak{E}\,(\subset\Pi_\varkappa)$ π-isometrisch (d.h. unter Erhaltung des π-Skalarproduktes) auf eine Menge $\mathfrak{E}'\,(\subset\Pi'_\varkappa)$ abbildet, sich zu einem linearen stetigen Operator T, der $\Pi_\varkappa$ π-unitär auf $\Pi'_\varkappa$ abbildet, fortsetzen läßt, wenn nur die linearen Hüllen von $\mathfrak{E}$ und $\mathfrak{E}'$ dicht in $\Pi_\varkappa$ bzw. $\Pi'_\varkappa$ liegen.

Auf Grund der Beziehungen (3. 2) und (3. 10) bildet T_0 die Elemente $U_{z_0 z}\xi$ π-isometrisch auf $e_z\xi$ ab, $z\in C_+\setminus\sigma_p(A)$, $\xi\in\mathfrak{G}$. Da die linearen Hüllen dieser Elemente in $\Pi_\varkappa$ bzw. $\Pi_\varkappa^0(Q)$ dicht liegen, läßt sich T_0 zu einem Operator T fortsetzen, der $\Pi_\varkappa$ π-unitär auf $\Pi_\varkappa^0(Q)$ abbildet. Dabei gilt

$$T(A-z_0I)(A-zI)^{-1}=(A_Q^0-z_0I)(A_Q^0-zI)^{-1}T \qquad (z\in C_+\setminus\sigma_p(A)),$$

woraus leicht $T\mathfrak{D}(A)=\mathfrak{D}(A_Q^0)$ und $TAf=A_Q^0Tf\,(f\in\mathfrak{D}(A))$ folgt.

5. *Die Extrapolationsaufgabe.* Es sei jetzt auf einer unendlichen Punktmenge $C_Q\subset C_+$ eine Funktion Q mit Werten in $[\mathfrak{G},\mathfrak{G}]$ definiert mit den folgenden Eigenschaften:

1) Der Kern $K(z,\zeta)=\dfrac{Q(z)-Q^*(\bar\zeta)}{z-\xi}$ hat $\varkappa$ negative Quadrate in C_Q.

2) Es gibt eine Punktfolge $(z_n)\subset C_Q$ mit der Eigenschaft (3. 3).

Dann kann man in der gleichen Weise wie in den Abschnitten 1 und 2 dieses Paragraphen den Raum $\Pi_\varkappa^0(Q)$ und den Operator A_Q^0 einführen. Obwohl letzterer i.a. nicht maximal π-hermitesch ist, erhält man für eine beliebige maximale π-hermitesche Erweiterung A von A_Q^0 mit $\mathfrak{n}_+(A)=0$ eine Darstellung (3. 8) der Funktion Q mit A an Stelle von A_Q^0. Daraus ergibt sich ohne Schwierigkeit der folgende

Satz 3. 4. *Zu jeder auf einer unendlichen Punktmenge $C_Q\subset C_+$ definierten Funktion Q mit Werten in $[\mathfrak{G},\mathfrak{G}]$ und den Eigenschaften* 1) *und* 2) *gibt es eine Funktion $Q\in N_\varkappa^0(\mathfrak{G})$, die — mit Ausnahme von höchstens $\varkappa$ Punkten aus C_Q — mit Q übereinstimmt.*

Die erwähnten Ausnahmepunkte gehören zum Punktspektrum des Operators A.

Ist speziell C_Q eine offene Menge und die Funktion Q mit den Eigenschaften 1) und 2) auf C_Q stetig, so läßt sich Q zu einer Funktion $\tilde Q\in N_\varkappa^0(\mathfrak{G})$ extrapolieren. Insbesondere ist dann also $\tilde Q$ meromorph in C_+ und damit auch Q holomorph in C_Q.

§ 4. Die Funktionenklasse $N_\varkappa(\mathfrak{G})$

1. *Der Raum $\Pi_\varkappa(Q)$ und der Operator A_Q.* Es sei Q wieder eine Funktion der Klasse $N_\varkappa^0(\mathfrak{G})$. Wir erweitern den maximalen π-hermiteschen Operator A_Q^0 in der Darstellung (3. 8) — falls er nicht bereits π-selbstadjungiert ist — zu einem π-selbstadjungierten Operator $\tilde A$ in einem Oberraum $\tilde\Pi_\varkappa\supset\Pi_\varkappa^0(Q)$. Dann besteht für Q die Darstellung (3. 8) mit $\tilde A$ an Stelle von A_Q^0:

$$Q(z)=Q^*(z_0)+(z-\bar z_0)\Gamma_{z_0}^+(\tilde A-z_0\tilde I)(\tilde A-z\tilde I)^{-1}\Gamma_{z_0}, \tag{4. 1}$$

wobei wir Γ_{z_0} wieder als Operator von $\mathfrak{G}$ in $\tilde\Pi_\varkappa$ auffassen. Extrapolieren wir Q auf das Spiegelbild $\bar{\mathscr{Z}}_Q^0$ von $\mathscr{Z}_Q^0$ bezüglich der reellen Achse durch die Festsetzung $Q(z)=$

$=Q^*(\bar{z})$, $z\in\bar{\mathscr{Z}}_Q^0$, dann gilt die Beziehung (4. 1) sogar für alle $z\in(\mathscr{Z}_Q^0\cup\bar{\mathscr{Z}}_Q^0)\setminus\sigma_p(\tilde{A})$. Aus der Bemerkung im Anschluß an Formel (3. 11) ergibt sich, daß die auf diese Weise extrapolierte Funktion Q sogar die Eigenschaft $(N_\varkappa)$ hat.

Wir überlegen uns jetzt, daß es andererseits die Eigenschaft $(N_\varkappa)$ — zusammen mit der Eigenschaft $(\mathfrak{D})$ — gestattet, eine natürliche π-selbstadjungierte Erweiterung A_Q des Operators A_Q^0 zu erhalten.

Zu diesem Zweck betrachten wir eine Funktion Q mit den Eigenschaften $(N_\varkappa)$ und (D), ordnen jedem $z\in\mathscr{Z}_Q$ ein Symbol e_z zu, bilden wieder die lineare Menge $\mathfrak{L}(Q)$ aller formalen Summen (3. 1) mit $\mathscr{Z}_Q^0$ ersetzt durch $\mathscr{Z}_Q$ und definieren für $f, g\in\mathfrak{L}(Q)$ das π-Skalarprodukt $[f, g]$ durch (3. 2). Auf der Menge

$$\check{\mathfrak{D}}=\{f: f\in\mathfrak{L}(Q), \chi(f)=0\},\ \chi(f)=\sum_{z\in\mathscr{Z}_Q}\xi_z \quad \text{für} \quad f=\sum_{z\in\mathscr{Z}_Q}e_z\xi_z,$$

definieren wir einen Operator $\check{A}$ durch die Gleichung $\check{A}f=\sum ze_z\xi_z$. Offensichtlich gilt dann $\mathfrak{L}^0(Q)\subset\mathfrak{L}(Q)$ und $\hat{A}\subset\check{A}$. Die kanonische Einbettung von $\mathfrak{L}^0(Q)$ in $\Pi_\varkappa^0(Q)$ kann man erweitern zu einer kanonischen Einbettung von $\mathfrak{L}(Q)$ in einen $\pi_\varkappa$-Raum $\Pi_\varkappa(Q)$. Dabei erzeugt $\check{A}$ einen Operator A_Q in $\Pi_\varkappa(Q)$, der sogar π-selbstadjungiert ist. Letzteres überlegt man sich ebenso, wie in § 3. 2 die Maximalität des π-hermiteschen Operators $\hat{A}_Q^0$ gezeigt wurde. Schließlich bleiben auch die Beziehungen (3. 6) und (3. 7) von § 3. 2 für $\check{A}$ und A_Q an Stelle von $\hat{A}$ und A_Q^0 für beliebige $z\in\mathscr{Z}_Q$ erhalten, und man überzeugt sich leicht davon, daß A_Q und Γ_{z_0} eng verbunden sind (vgl. den Anfang des Beweises von Satz 3. 1).

Mit $N_\varkappa(\mathfrak{G})$ bezeichnen wir die Klasse aller Funktionen Q mit Werten in $[\mathfrak{G}, \mathfrak{G}]$ und den Eigenschaften $(N_\varkappa)$ und $(\mathfrak{D})$. Die obigen Überlegungen lassen sich dann zusammenfassen in der folgenden Ergänzung von Satz 3. 1.

Satz 4. 1. *Es sei A ein π-selbstadjungierter Operator in einem $\pi_\varkappa$-Raum $\Pi_\varkappa$, $\mathfrak{G}$ ein Hilbertraum, $S=S^*\in[\mathfrak{G}, \mathfrak{G}]$, $\Gamma\in[\mathfrak{G}, \Pi_\varkappa]$ und $z_0\in C_+\setminus\sigma_p(A)$. Dann definiert die Gleichung*

$$Q(z)=S-iy_0\Gamma^+\Gamma+(z-\bar{z}_0)\Gamma^+(A-z_0I)(A-zI)^{-1}\Gamma \tag{4. 2}$$

eine Funktion der Klasse $N_{\varkappa'}(\mathfrak{G})$ für ein gewisses $\varkappa'$ mit $0\leqq\varkappa'\leqq\varkappa$. Sind die Operatoren A und Γ eng verbunden, so ist $\varkappa=\varkappa'$. Umgekehrt gestattet jede Funktion $Q\in N_\varkappa(\mathfrak{G})$ die Darstellung (4. 2) mit $\Pi_\varkappa=\Pi_\varkappa(Q)$, dem π-selbstadjungierten Operator $A=A_Q$, $S=\operatorname{Re}Q(z_0)$ und $\Gamma=\Gamma_{z_0}$; dabei sind die Operatoren A_Q und Γ_{z_0} eng verbunden.

Analog zu Satz 3. 3. gilt die folgende Aussage.

Satz 4. 2. *Gestattet die Funktion $Q\in N_\varkappa(\mathfrak{G})$ die Darstellung (4. 2) mit einem π-selbstadjungierten Operator A in $\Pi_\varkappa$, $S=S^*\in[\mathfrak{G}, \mathfrak{G}]$, $\Gamma\in[\mathfrak{G}, \Pi_\varkappa]$ und sind A und Γ eng verbunden, so ist A bis auf π-unitäre Äquivalenz eindeutig bestimmt, genauer, A ist dem Operator A_Q π-unitär äquivalent.*

2. *Spektralzerlegung der Funktion* $Q \in N_\varkappa(\mathfrak{G})$. Wir gehen jetzt aus von einer Darstellung (4. 2) der Funktion Q:

$$(4.3)\qquad Q(z) = \operatorname{Re} Q(z_0) - iy_0 \Gamma\Gamma^+ + (z-\bar{z}_0)\Gamma^+(A-z_0I)(A-zI)^{-1}\Gamma = $$
$$= C + z\Gamma^+\Gamma + (z-\bar{z}_0)(z-z_0)\Gamma^+(A-zI)^{-1}\Gamma$$

$\left(C = C^* = \operatorname{Re} Q(z_0) - x_0\Gamma^+\Gamma\right)$ mit einem π-selbstadjungierten Operator A in $\Pi_\varkappa$, der mit Γ eng verbunden ist. Für einen nichtreellen Eigenwert $\alpha \in \sigma_p(A)$ bezeichne E_α den Rieszschen Projektor auf den zugehörigen algebraischen Eigenraum; dann gilt bekanntlich $E_\alpha = E_{\bar{\alpha}}^+$ ([16]). Es sei im folgenden $\sigma_+(A) = \sigma(A) \cap C_+$. Wir setzen für $\alpha \in \sigma_+(A)$:

$$F_\alpha = E_\alpha + E_\alpha^+, \quad E_0 = I - \sum_{\alpha\in\sigma_+(A)} F_\alpha, \quad \Pi^{(\alpha)} = F_\alpha \Pi_\varkappa, \quad \Pi^{(0)} = E_0\Pi_\varkappa,$$
$$A_0 = A|_{\Pi^{(0)}}, \quad A_\alpha = A|_{\Pi^{(\alpha)}}.$$

Dann ist das Spektrum von A_0 reell, das von A_α besteht genau aus den Punkten α und $\bar{\alpha}$.

Die Räume $\Pi^{(\alpha)}$ sind $\pi_{\varkappa_\alpha}$-Räume der Dimension $2\varkappa_\alpha$ mit $\varkappa_\alpha = \dim \mathfrak{N}(E_\alpha)$, $\Pi^{(0)}$ ist ein $\pi_{\varkappa_0}$-Raum, $\varkappa_0 = \varkappa - \sum_{\alpha\in\sigma_+(A)} \varkappa_\alpha$, und es gilt

$$\Pi_\varkappa = \Pi^{(0)}[+]\sum_{\alpha\in\sigma_+(A)}^{[+]} \Pi^{(\alpha)}.$$

Die Resolvente $(A-zI)^{-1}$ gestattet bekanntlich die Darstellung

$$(4.4)\qquad (A-zI)^{-1} = (A_0 - zI_0)^{-1}E_0 + \sum_{\alpha\in\sigma_+(A)} H_\alpha(z)F_\alpha;$$

dabei ist $H_\alpha(z)$ die Summe der zu α und $\bar{\alpha}$ gehörigen Hauptteile von $(A-zI)^{-1}$. Setzen wir (4. 4) in (4. 3) ein, so folgt

$$(4.5)\qquad Q(z) = C + Q_0(z) + \sum_{\alpha\in\sigma_+(A)} Q_\alpha(z),$$

mit

$$(4.6)\qquad Q_0(z) = z\mathring{\Gamma}^+\mathring{\Gamma} + (z-\bar{z}_0)(z-z_0)\mathring{\Gamma}^+(A_0 - zI_0)^{-1}\mathring{\Gamma},$$

$$(4.7)\qquad Q_\alpha(z) = z\overset{\alpha}{\Gamma}{}^+\overset{\alpha}{\Gamma} + (z-\bar{z}_0)(z-z_0)\overset{\alpha}{\Gamma}{}^+(A_\alpha - zI_\alpha)^{-1}\overset{\alpha}{\Gamma}$$

($\mathring{\Gamma} = E_0\Gamma \in [\mathfrak{G}, \Pi^{(0)}]$, $\overset{\alpha}{\Gamma} = F_\alpha\Gamma \in [\mathfrak{G}, \Pi^{(\alpha)}]$). Man sieht leicht, daß der Operator A_α (bzw. A_0) mit $\overset{\alpha}{\Gamma}$ (bzw. $\mathring{\Gamma}$) wieder eng verbunden ist. Deshalb gilt gemäß dem ersten Teil von Satz 4. 1 $Q_\alpha \in N_{\varkappa_\alpha}(\mathfrak{G})$, $Q_0 \in N_{\varkappa_0}(\mathfrak{G})$.

Die Funktion Q_α, $\alpha \in \sigma_+(A)$, hat als einzige Singularitäten in der abgeschlossenen komplexen Ebene je einen Pol bei $z = \alpha$ und $z = \bar{\alpha}$, die Funktion Q_0 ist in $C_+ \cup C_-$ stückweise holomorph. Deshalb stimmt Q_α bis auf einen konstanten selbstadjungierten Summanden mit der Summe der Hauptteile der Funktion Q zu den Polen $z=\alpha$ und $z=\bar{\alpha}$ überein. Ist $H_\alpha(z) = \sum_{\nu=1}^{m_\alpha} \frac{E_{\alpha,\nu}}{(\alpha-z)^\nu} + \sum_{\nu=1}^{m_\alpha} \frac{E_{\alpha,\nu}^+}{(\bar{\alpha}-z)^\nu}$, so ergibt sich für Q_α aus (4. 7)

$$Q_\alpha(z) = zQ_{\alpha,0} + (z-\bar{z}_0)(z-z_0)\left\{\sum_{\nu=1}^{m_\alpha} \frac{Q_{\alpha,\nu}}{(\alpha-z)^\nu} + \sum_{\nu=1}^{m_\alpha} \frac{Q_{\alpha,\nu}^*}{(\bar{\alpha}-z)^\nu}\right\} \tag{4. 8}$$

mit $Q_{\alpha,\nu} = \overset{\alpha}{\Gamma}{}^+ E_{\alpha,\nu} \overset{\alpha}{\Gamma}$, $\nu = 1, 2, \ldots, m_\alpha$, $Q_{\alpha,0} = Q_{\alpha,1} + Q_{\bar{\alpha},1}^* = \overset{\alpha}{\Gamma}{}^+ \overset{\alpha}{\Gamma}$.

Die Resolvente $(A_0 - zI_0)^{-1}$ gestattet gemäß den Ergebnissen von [12], [13] die Darstellung

$$(A_0 - zI_0)^{-1} = \frac{1}{\varphi(z)}\left\{\int_{-\infty}^{\infty} \frac{dF(s)}{s-z} + P(A_0; z)\right\}; \tag{4. 9}$$

dabei ist $\varphi(z) = \frac{p^2(z)}{(z-z_0)^m (z-\bar{z}_0)^m}$, wenn $p(z)$ das Minimalpolynom der Einschränkung von A auf einen $\varkappa$-dimensionalen invarianten nichtpositiven Teilraum bezeichnet (m=Grad von p; bekanntlich (vgl. [17]) ist das Polynom p bis auf einen konstanten Faktor eindeutig bestimmt), $P(\zeta; z)$ ist die in der abgeschlossenen komplexen Ebene mit Ausnahme der Punkte z_0 und $\bar{z}_0$ in z und ζ holomorphe Funktion

$$P(\zeta; z) = \frac{\varphi(\zeta) - \varphi(z)}{\zeta - z}$$

und $F(s)$ eine auf der reellen Achse definierte Funktion mit Werten in $[\Pi^{(0)}, \Pi^{(0)}]$ und den Eigenschaften:

1) $F(0) = 0, \quad F(s) = \frac{F(s+0) + F(s-0)}{2}$;[14]

2) $[(F(s) - F(s'))g, g] \geqq 0$ für alle $g \in \Pi^{(0)}$, $s' \leqq s$;

3) $\int_{-\infty}^{\infty} d[F(s)g, g] < \infty$ für alle $g \in \Pi^{(0)}$.

Setzen wir (4. 9) in (4. 6) ein, so ergibt sich

$$Q_0(z) = zQ_{00} + \frac{(z-\bar{z}_0)(z-z_0)}{\varphi(z)}\left\{\int_{-\infty}^{\infty} \frac{d\Phi(s)}{s-z} + R(z)\right\} \tag{4. 10}$$

mit $Q_{00} (= Q_{00}^*) = \Gamma^+ \overset{\circ}{\Gamma} \in [\mathfrak{G}, \mathfrak{G}]$, $\Phi(s) = \overset{\circ}{\Gamma}{}^+ \overset{\circ}{F}(s) \overset{\circ}{\Gamma} \in [\mathfrak{G}, \mathfrak{G}]$, $\Omega(s') \leqq \Phi(s)$ für $s' \leqq s$ und $R(z) = \overset{\circ}{\Gamma}{}^+ P(\overset{\circ}{A}_0; z) \overset{\circ}{\Gamma} \in [\mathfrak{G}, \mathfrak{G}]$ ist eine auf der abgeschosse komplexen Ebene mit Ausnahme der Punkte z_0 und $\bar{z}_0$ holomorphe Operatorfunktion.

---

[14] Die Grenzwerte verstehen sich in der schwachen Operatorentopologie.

Bezeichnen β_j, $j=1, 2, \ldots, l$ die (notwendig reellen) Nullstellen des Polynoms p, $\beta_j \neq \beta_k$ für $j \neq k$, und setzen wir $\Phi_j = \Phi(\beta_j+0)-\Phi(\beta_j-0)$, so kann man (4. 10) in der äquivalenten Form

$$
\begin{aligned}
Q_0(z) = zQ_{00} + \int_{-\infty}^{\infty} \left\{ \frac{1}{t-z} + \frac{P(t;z)(z-z_0)(z-\bar{z}_0)}{\varphi(z)|t-z_0|^2} + \frac{z_0+\bar{z}_0-(t+z)}{|t-z_0|^2} \right\} d\Psi(t) \\
+ \frac{(z-z_0)(z-\bar{z}_0)}{\varphi(z)} \left(R(z) + \sum_{j=1}^{l} \frac{\Phi_j}{\beta_j - z} \right)
\end{aligned}
\tag{4.11}
$$

schreiben; dabei haben wir $d\Psi(t) = \dfrac{|t-z_0|^2}{\varphi(t)} d\Phi(t)$ gesetzt. Das Integral in (4. 11) existiert als uneigentliches Integral bezüglich der eventuellen Singularitäten des Maßes $d\Psi(t)$ in den Punkten $\beta_1, \ldots, \beta_l$ und ∞ in der starken Operatorentopologie.

Unter dem *Spektrum* $\sigma(Q)$ der Funktion $Q \in N_\varkappa(\mathfrak{G})$ verstehen wir die Menge aller Pole von Q sowie die Menge derjenigen reellen Punkte x, in die sich Q nicht so analytisch fortsetzen läßt, daß $Q(x)$ selbstadjungiert ist; außerdem sei $\sigma_+(Q) = \sigma(Q) \cap C_+$.

Satz 4. 3. *Jede Funktion $Q \in N_\varkappa(\mathfrak{G})$ gestattet eine Zerlegung*

$$
Q(z) = C + Q_0(z) + \sum_{\alpha \in \sigma_+(Q)} Q_\alpha(z) \tag{4.12}
$$

mit einem konstanten selbstadjungierten Operator C, $Q_\alpha \in N_{\varkappa_\alpha}(\mathfrak{G})$ $(\alpha \in \sigma_+(Q))$, $Q_0 \in N_{\varkappa_0}(\mathfrak{G})$; dabei gilt $\varkappa = \varkappa_0 + \sum_{\alpha \in \sigma_+(Q)} \varkappa_\alpha$. Die Operatorfunktion Q_α hat die Gestalt (4. 8) *und ist bis auf einen konstanten selbstadjungierten Summanden die Summe der zu α und $\bar{\alpha}$ gehörigen Hauptteile von Q; die Operatorfunktion Q_0 ist in C_+ und C_- stückweise holomorph und gestattet die Darstellungen* (4. 10) *und* (4. 11).

Die Gültigkeit dieses Satzes ergibt sich aus den vorangegangenen Überlegungen, wenn man nur beachtet, daß $\sigma_+(Q) = \sigma_+(A_Q)$ ist. Letzteres ist eine unmittelbare Konsequenz des folgenden Satzes.

Satz 4. 4. *Es gilt $\sigma(Q) = \sigma(A_Q)$, und diese Menge besteht genau aus den nichtreellen Polen von Q, den Nullstellen der Funktion φ sowie den Wachstumspunkten der Funktionen Φ oder Ψ aus den Darstellungen* (4. 10) *oder* (4. 11). *Jeder isolierte Punkt λ_0 von $\sigma(Q)$ ist ein Pol; seine Ordnung stimmt mit der Ordnung von λ_0 als Pol der Resolvente von A_Q überein.*

Beweis. Gemäß dem zweiten Teil von Satz 4. 1 gilt

$$
Q(z) = Q^*(z_0) + (z-\bar{z}_0)\Gamma_{z_0}^+(A_Q - \bar{z}_0 I)(A_Q - zI)^{-1}\Gamma_{z_0}, \tag{4.13}
$$

also ist Q in allen Punkten $z \notin \sigma(A_Q)$ regulär und nimmt in reellen Punkten $z=x \notin \sigma(A_Q)$ selbstadjungierte Werte an, d.h., es gilt $\sigma(Q) \subset \sigma(A_Q)$.

Um die inverse Inklusion zu beweisen, beachten wir zunächst die für beliebige $\zeta, \zeta' \notin \sigma(A_Q)$ aus (4. 13) folgende Beziehung

(4. 14)

$$\Gamma_{z_0}^+(A_Q-\bar{\zeta}I)^{-1}(A_Q-zI)^{-1}(A_Q-\zeta' I)^{-1}\Gamma_{z_0} = \frac{Q(z)}{(z-\zeta')(z-\bar{\zeta})(z-z_0)(z-\bar{z}_0)} + \cdots,$$

wobei die nicht aufgeschriebenen Summanden in allen Punkten $z \neq \bar{\zeta}, \zeta', z_0, \bar{z}_0$, holomorph von z abhängen. Ist λ_0 ein isolierter Punkt von $\sigma(A_Q)$ und gilt in einer Umgebung von λ_0

$$(A_Q-zI)^{-1} = \frac{A_{-n}}{(\lambda_0-z)^n} + \cdots + \frac{A_{-1}}{\lambda_0-z} + A_Q(z)(I-E_{\lambda_0}), \quad A_{-n} \neq 0, \; z \neq \lambda_0,$$

so hat auch Q im Punkte λ_0 einen Pol der Ordnung n, da anderenfalls aus (4. 14)

$$[A_{-n}(A_Q-\zeta' I)^{-1}\Gamma_{z_0}\xi, (A_Q-\zeta I)^{-1}\Gamma_{z_0}\eta] = 0$$

für alle $\xi, \eta \in \mathfrak{G}$, $\zeta, \zeta' \notin \sigma(A_Q)$ und wegen der engen Verbundenheit von A_Q und Γ_{z_0} daraus $A_{-n}=0$ folgen würde. Entsprechend ergibt sich aus (4. 14) mit Hilfe der Umkehrformel, daß die Spektralfunktion von A_Q in einem reellen Punkt $x \notin \sigma(Q)$ konstant ist. Damit ist der Satz bewiesen.

Wir bemerken schließlich, daß die oben eingeführten Operatoren A_Q^0 und A_Q zusammenfallen können. Man sieht leicht, daß dies genau dann der Fall ist, wenn A_Q^0 bereits π-selbstadjungiert ist oder wenn die Elemente $e_z\xi$ für $\xi \in \mathfrak{G}$, $\operatorname{Im} z > 0$ den ganzen Raum $\Pi_\varkappa(Q)$ aufspannen. Ist $\mathfrak{G}$ insbesondere ein eindimensionaler Raum, so läßt sich aus einem Ergebnis [18] (siehe auch [19]) eines der Verfasser folgern, daß dies genau dann der Fall ist, wenn

$$\int_{-\infty}^{\infty} \frac{\ln \Phi'(t)}{1+t^2}\, dt = -\infty$$

gilt; dabei bezeichnet $\Phi'(t)$ die Ableitung des absolutstetigen Anteiles der Funktion Φ aus der Darstellung (4. 10) von Q_0.

3. *Die Funktionen Q_α.* Es sei jetzt α eine komplexe Zahl, $\mathfrak{A}$ eine Umgebung des Punktes α in der komplexen Ebenen und $S(z)$ eine in allen Punkten $z \in \mathfrak{A}$, $z \neq \alpha$, erklärte und holomorphe Funktion mit Werten in $[\mathfrak{G}, \mathfrak{G}]$, die bei $z=\alpha$ einen Pol hat. Gemäß [20] definiert man die Polvielfachheit von α folgendermaßen.

Eine im Punkte α holomorphe Funktion $\chi(z)$ mit Werten aus $\mathfrak{G}$ und $\chi(\alpha) \neq 0$ heißt *Polfunktion* von $S(z)$ im Punkte α, wenn $\chi(z)$ die Darstellung $\chi(z) = S(z)\psi(z)$ mit einer in α holomorphen Funktion $\psi(z)$ gestattet, deren Werte ebenfalls in $\mathfrak{G}$

liegen und für die $\psi(\alpha)=0$ ist. Die Ordnung der Nullstelle α von $\psi(z)$ heißt die *Ordnung* der Polfunktion $\chi(z)$ und $\chi(\alpha)$ ein Polvektor von $S(z)$ zum Pol α. Die Polvektoren von $S(z)$ zum Pol α bilden zusammen mit dem Nullelement eine lineare Menge; deren Abschließung heißt der *Polkern* von $S(z)$ im Punkte α und wird mit Pol $S(\alpha)$ bezeichnet.

Unter dem *Rang* eines Polvektors $\chi_0 \in \operatorname{Pol} S(\alpha)$ versteht man das Maximum der Ordnungen aller Polfunktionen $\chi(z)$ mit $\chi(\alpha)=\chi_0$ zum Pole α; er werde mit rang χ_0 bezeichnet.

Es sei jetzt $\beta=\dim \operatorname{Pol} S(\alpha)<\infty$. Unter einem *kanonischen System* von Polvektoren von $S(z)$ im Punkte α versteht man eine Basis $\chi_1, \chi_2, \dots, \chi_\beta$ des Raumes Pol $S(\alpha)$ mit der Eigenschaft

$$\operatorname{rang} \chi_j = \max_{\chi \in L_j} \operatorname{rang} \chi, \qquad j = 1, 2, \dots, \beta,$$

wobei L_j die lineare Hülle von $\chi_j, \chi_{j+1}, \dots, \chi_\beta$ bezeichnet. Setzt man für ein kanonisches System von Polvektoren $\chi_1, \chi_2, \dots, \chi_\beta$ noch $p_j=\operatorname{rang} \chi_j$, $j=1, 2, \dots, \beta$, so sind die Zahlen $p_1, p_2, \dots, p_\beta$ durch die Funktion $S(z)$ eindeutig bestimmt; ihre Summe heißt die *Polvielfachheit* des Poles α der Funktion $S(z)$.

Ist $S(z)$ insbesondere die Resolvente eines abgeschlossenen linearen Operators B in $\mathfrak{G}$, für den $z=\alpha$ ein normaler[15]) Eigenwert ist, so stimmt dessen algebraische Vielfachheit mit der Polvielfachheit des Poles α von $S(z)=(B-zI)^{-1}$ überein.

Lemma 4. 1. *Gestattet die Funktion $S(z)$ bei $z=\alpha$ die Darstellung*

$$S(z) = \sum_{\nu=0}^{m-1} \frac{S_{m-\nu+1}}{(\alpha-z)^{\nu+1}} + S_0(z), \quad z \neq \alpha, \quad S_1 \neq O,$$

mit $0<m<\infty$ und einer bei $z=\alpha$ holomorphen Funktion $S_0(z)$, so stimmt die Polvielfachheit des Poles α von $S(z)$ überein mit der Dimension des Wertebereiches des Operators

$$\mathfrak{S} = \begin{pmatrix} S_1 & 0 & 0 & \dots & 0 \\ S_2 & S_1 & 0 & \dots & 0 \\ S_3 & S_2 & S_1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ S_m & S_{m-1} & S_{m-2} & \dots & S_1 \end{pmatrix}$$

im Raume $\mathfrak{G}^m = \mathfrak{G}_1 \oplus \mathfrak{G}_2 \oplus \dots \oplus \mathfrak{G}_m$ mit $\mathfrak{G}_1=\mathfrak{G}_2=\dots=\mathfrak{G}_m=\mathfrak{G}$.

---

[15]) D.h., α ist ein isolierter Punkt des Spektrums von B, dessen zugehöriger Rieszscher Projektor endlichdimensional ist (vgl. [21]).

Beweis. Aus der Definition des Polvektors folgt leicht, daß χ_0 genau dann ein Polvektor von $S(z)$ zum Pol α ist, wenn Elemente $\psi_1, \psi_2, \ldots, \psi_m$ existieren, so daß

$$\mathfrak{S}\vec{\psi} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \chi_0 \end{pmatrix}, \quad \vec{\psi} = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_{m-1} \\ \psi_m \end{pmatrix} \tag{4.16}$$

gilt; hat dabei χ_0 den Rang $p_0 > 0$, so kann $\psi_1 = \psi_2 = \cdots = \psi_{p_0-1} = 0$ gewählt werden.

Mit $\mathfrak{T}$ bezeichnen wir den folgenden Verschiebungsoperator im Raum $\mathfrak{G}^m$:

$$\mathfrak{T}\begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_{m-1} \\ \psi_m \end{pmatrix} = \begin{pmatrix} \psi_2 \\ \psi_3 \\ \vdots \\ \psi_m \\ 0 \end{pmatrix}.$$

Es sei $\chi_1, \chi_2, \ldots, \chi_\beta$ ein kanonisches System von Polvektoren von $S(z)$ im Punkte α. Wir ordnen χ_j einen Vektor

$$\vec{\psi}_j = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_{m-1} \\ \psi_m \end{pmatrix}$$

zu, so daß (4. 16) mit $\vec{\psi} = \vec{\psi}_j$, $\chi_0 = \chi_j$ besteht und dabei die ersten $p_j - 1$ Komponenten von $\vec{\psi}_j$ verschwinden, wenn p_j wieder den Rang von χ_j bezeichnet ($j = 1, 2, \ldots, \beta$). Man sieht leicht, daß $p_1 = m$ gilt und die Vektoren

$$\begin{gathered} \mathfrak{S}\vec{\psi}_1, \mathfrak{S}\mathfrak{T}\vec{\psi}_1, \ldots, \mathfrak{S}\mathfrak{T}^{m-1}\vec{\psi}_1, \\ \mathfrak{S}\vec{\psi}_2, \mathfrak{S}\mathfrak{T}\vec{\psi}_2, \ldots, \mathfrak{S}\mathfrak{T}^{p_2-1}\vec{\psi}_2, \\ \vdots \\ \mathfrak{S}\vec{\psi}_\beta, \mathfrak{S}\mathfrak{T}\vec{\psi}_\beta, \ldots, \mathfrak{S}\mathfrak{T}^{p_\beta-1}\vec{\psi}_\beta \end{gathered} \tag{4.17}$$

linear unabhängig sind. Also ist die Polvielfachheit des Poles α von $S(z)$ mindestens gleich der Dimension des Wertebereiches von $\mathfrak{T}$.

Wir müssen noch zeigen, daß umgekehrt auch jedes Element $\vec{\varphi} = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_m \end{pmatrix}$ des Wertebereiches von $\mathfrak{S}$ als Linearkombination von Elementen aus (4. 17) darstellbar ist. Zunächst gibt es ein $\vec{\psi}_1$ mit $\mathfrak{S}\mathfrak{T}^{m-1}\vec{\psi}_1 = \begin{pmatrix}\varphi_1\\ \vdots\end{pmatrix}$, also können wir annehmen, daß die erste Komponente φ_1 von $\vec{\varphi}$ gleich Null ist. Dann gibt es Elemente $\psi_1, \psi_2, \ldots$ $\ldots, \psi_m \in \mathfrak{G}$ mit

$$\mathfrak{S}\begin{pmatrix}\psi_1\\ \psi_2\\ \vdots\\ \psi_m\end{pmatrix} = \begin{pmatrix}0\\ \varphi_2\\ \vdots\\ \varphi_m\end{pmatrix}, \text{ also gilt auch } \mathfrak{S}\begin{pmatrix}0\\ 0\\ \vdots\\ \psi_1\\ \psi_2\end{pmatrix} = \begin{pmatrix}0\\ 0\\ \vdots\\ 0\\ \varphi_2\end{pmatrix},$$

d.h., φ_2 ist eine Linearkombination von Polvektoren vom Rang größer oder gleich $m-1$. Wir finden somit eine Linearkombination von Elementen aus (4. 17), deren erste Komponente verschwindet und deren zweite Komponente gleich φ_2 ist. Also können wir annehmen, daß die ersten beiden Komponenten von $\vec{\varphi}$ verschwinden. Wiederholung dieser Überlegungen liefert schließlich die Behauptung.

Der Raum $\mathfrak{G}$ sei jetzt endlichdimensional, die (Matrix-) Funktion $S(z)$ in einem Gebiet $\mathfrak{B}$ der komplexen Ebene meromorph, in mindestens einem Punkte von $\mathfrak{B}$ invertierbar und habe dort nur endlich viele charakteristische Zahlen (siehe z. B. [20]) und Pole. Die Summe der Nullvielfachheiten ([20], [22]) bzw. Polvielfachheiten aller in $\mathfrak{B}$ gelegenen charakteristischen Zahlen bzw. Pole von $S(z)$ nennen wir die *Nullvielfachheit* bzw. *Polvielfachheit* von $\mathfrak{B}$ bezüglich $S(z)$ und bezeichnen sie mit n bzw. p. Dann gilt ([20], [22])

$$n-p = \frac{1}{2\pi i}\, sp \int_{\mathscr{C}} S'(z) S^{-1}(z)\,dz \left(= \frac{1}{2\pi i}\int_{\mathscr{C}} d\log\det S(z)\right); \tag{4.18}$$

dabei ist $\mathscr{C}$ eine ganz in $\mathfrak{B}$ gelegene Kontur, die jeden Pol und jede charakteristische Zahl von $S(z)$ in $\mathfrak{B}$ genau einmal in positivem Sinne umschließt.

Die Funktion $Q_\alpha \in N_{\varkappa_\alpha}(\mathfrak{G})$ aus der Zerlegung (4. 5) oder (4. 12) läßt sich folgendermaßen schreiben:

$$(Q_\alpha(z)=)\; S(z) = S_0 + \sum_{\nu=1}^{m} \frac{S_{m-\nu+1}}{(\alpha-z)^\nu} + \sum_{\nu=1}^{m} \frac{S^*_{m-\nu+1}}{(\bar{\alpha}-z)^\nu}; \tag{4.19}$$

dabei ist $\alpha\neq\bar{\alpha}$, $m=m_\alpha$, $S_0=S_0^*\in[\mathfrak{G},\mathfrak{G}]$, $S_1\neq 0$ und die Operatoren $S_\nu\in[\mathfrak{G},\mathfrak{G}]$, $\nu=1, 2, \ldots, m$, sind endlichdimensional.

Wir zeigen jetzt, daß umgekehrt für beliebige Operatoren S_ν, $\nu=0, 1, \ldots, m$, mit diesen Eigenschaften die Funktion (4. 19) zu einer Klasse $N_\varkappa(\mathfrak{G})$ gehört. Genauer, es gilt der

Satz 4. 5. *Ist* $\alpha\neq\bar{\alpha}$, $S_0=S_0^*\in[\mathfrak{G},\mathfrak{G}]$ *und sind die Operatoren* $S_1, S_2, \ldots, S_m \in$ $\in[\mathfrak{G},\mathfrak{G}]$ *endlichdimensional,* $S_1\neq 0$, *dann gehört die Funktion* $S(z)$ *aus* (4. 19) *zur*

Klasse $N_\varkappa(\mathfrak{G})$, wenn $\varkappa$ die Polvielfachheit des Poles $z=\alpha$ (oder $z=\bar{\alpha}$) von $S(z)$ bezeichnet. Der Raum $\Pi_\varkappa(S)$ hat die Dimension $2\varkappa$.

Beweis. Die Bedingung ($\mathfrak{D}$) ist für $S(z)$ offensichtlich erfüllt. Zum Nachweis der Bedingung ($\mathfrak{N}_\varkappa$) gehen wir aus von der Beziehung

$$\frac{S(z)-S^*(\zeta)}{z-\bar{\zeta}} = \sum_{\nu=1}^{m}\Bigg(S_\nu \sum_{\sigma=0}^{\nu-1} \frac{1}{(\alpha-z)^{\sigma+1}}\,\frac{1}{(\alpha-\bar{\zeta})^{\nu-\sigma}} + \\ + S_\nu^* \sum_{\sigma=0}^{\nu-1} \frac{1}{(\bar{\alpha}-z)^{\sigma+1}}\,\frac{1}{(\bar{\alpha}-\bar{\zeta})^{\nu-\sigma}}\Bigg).$$

Für beliebige n; $\alpha_j\in C$, $z_j\in(C_+\cup C_-)\setminus\{\alpha,\bar{\alpha}\}$, $z_j\neq\bar{z}_k$ und $\xi_j\in\mathfrak{G}$, $j,k=1,2,\dots,n$, folgt daraus mit $\varphi_\sigma = \sum_{j=1}^{n}\frac{\alpha_j\xi_j}{(\alpha-z_j)^\sigma}$, $\bar{\varphi}_\sigma = \sum_{j=1}^{n}\frac{\alpha_j\xi_j}{(\bar{\alpha}-z_j)^\sigma}$ $(\sigma=1,2,\dots,m)$:

$$\text{(4.20)}\qquad \sum_{j,k=1}^{m}\Big(\frac{S(z_j)-S^*(z_k)}{z_j-\bar{z}_k}\xi_j,\xi_k\Big)\alpha_j\bar{\alpha}_k = \left(\begin{pmatrix} O & \mathfrak{S}^* \\ \mathfrak{S} & O\end{pmatrix}\begin{pmatrix}\vec{\varphi}\\ \vec{\bar{\varphi}}\end{pmatrix}, \begin{pmatrix}\vec{\varphi}\\ \vec{\bar{\varphi}}\end{pmatrix}\right);$$

dabei hat $\mathfrak{S}$ wieder die Gestalt (4.15) und wir haben

$$\vec{\varphi} = \begin{pmatrix}\varphi_1\\ \varphi_2\\ \vdots\\ \varphi_m\end{pmatrix}, \qquad \vec{\bar{\varphi}} = \begin{pmatrix}\bar{\varphi}_1\\ \bar{\varphi}_2\\ \vdots\\ \bar{\varphi}_m\end{pmatrix}$$

gesetzt. Die Anzahl der negativen (positiven) Quadrate der Form in (4.20) stimmt folglich überein mit der Anzahl der negativen (positiven) Eigenwerte von $\begin{pmatrix} O & \mathfrak{S}^* \\ \mathfrak{S} & O\end{pmatrix}$, die ihrerseits gleich der Dimension des Wertebereiches von $\mathfrak{S}$ ist. Mit Lemma 4.1 folgt daraus die Behauptung.

Eine Ergänzung zu Satz 4.4 bildet die

Folgerung 4.1. *Für einen nichtreellen Pol α von Q stimmt seine Polvielfachheit mit dem Index $\varkappa_\alpha$ des Punktes α bezüglich A_Q überein, m.a.W., α ist ein Pol von Q_α der Polvielfachheit $\varkappa_\alpha$.*

§ 5. Der Operator $\hat{A}_Q$

Wir zeigen in diesem Paragraphen, daß jede Operatorfunktion Q mit Werten in $[\mathfrak{G},\mathfrak{G}]$ und den Eigenschaften (I)—(IV) des Hauptsatzes die Q-Funktion eines einfachen π-hermiteschen Operators ist.

Es sei zunächst $Q\in N_\varkappa(\mathfrak{G})$. Dann hat für beliebiges $z\neq\bar{z}$ der Operator $\frac{\operatorname{Im} Q(z)}{\operatorname{Im} z}$ höchstens $\varkappa$ negative Eigenwerte. Genügt Q der Bedingung (III) des Hauptsatzes, so entartet der Imaginärteil von Q nicht in dem Sinne, daß kein Element $\xi \in \mathfrak{G}$ mit $(\operatorname{Im} Q(z))\xi = 0$ für alle $z \in \mathscr{Z}_Q$ existiert. Hat Q überdies die Eigenschaft (IV) des Hauptsatzes und ist z. B. für $z_0 \neq \bar{z}_0$: $\operatorname{Im} Q(z_0) \geqq \gamma I$, $\gamma > 0$, so ist der in § 3.2 eingeführte Operator Γ_{z_0} beschränkt invertierbar, denn dann gilt für beliebiges $\xi \in \mathfrak{G}$:

$$\|\Gamma_{z_0}\xi\|^2 \geqq [\Gamma_{z_0}\xi, \Gamma_{z_0}\xi] \geqq \gamma \|\xi\|^2.$$

Der Teilraum $\Gamma_{z_0}\mathfrak{G}$ ist also insbesondere abgeschlossen in $\Pi_\varkappa(Q)$.

Für beliebiges $z \in \mathscr{Z}_Q$ betrachten wir den Operator Γ_z:

$$\Gamma_z \xi = e_z \xi \qquad (\xi \in \mathfrak{G}).$$

Dann gilt gemäß (3. 7) $\Gamma_z = (A_Q - z_0 I)(A_Q - zI)^{-1}\Gamma_{z_0}$, also ist auch jeder Teilraum $\Gamma_z\mathfrak{G}$ $(z \in Z_Q)$ abgeschlossen.

Für die Funktion $Q \in N_\varkappa(\mathfrak{G})$ mit den Eigenschaften (III) und (IV) bezeichne $\mathfrak{D}$ die Menge aller $f \in \mathfrak{D}(A_Q)$, für die

$$[(A_Q - z_0 I)f, e_{\bar{z}_0}\eta] = 0 \quad \text{für alle} \quad \eta \in \mathfrak{G}$$

gilt. Die Menge $\mathfrak{D}$ ist unabhängig von der Wahl des Punktes z_0 $(z_0 \neq \bar{z}_0,\ z_0 \in \mathscr{Z}_Q)$. Für einen beliebigen Punkt $z \neq \bar{z}$, $z \notin \sigma_p(A_Q)$ und $f \in \mathfrak{D}$ gilt nämlich auch

$$[(A_Q - zI)f, e_{\bar{z}}\eta] = \left[(A_Q - z_0 I)f + (z_0 - z)f, (A_Q - \bar{z}_0 I)\frac{e_{\bar{z}_0}\eta - e_{\bar{z}}\eta}{\bar{z}_0 - \bar{z}}\right] =$$

$$= \left[(A_Q - z_0 I)f, (A_Q - \bar{z}_0 I)\frac{e_{\bar{z}_0}\eta - e_{\bar{z}}\eta}{\bar{z}_0 - \bar{z}}\right] - [(A_Q - z_0 I)f, e_{\bar{z}}\eta] = 0.$$

Die Menge $\mathfrak{D}$ liegt dicht in $\Pi_\varkappa(Q)$. Um das zu sehen, betrachten wir ein $f_0 \in \Pi_\varkappa(Q)$ mit $f_0 [\perp] \mathfrak{D}$. Für $g_0 = (A_Q - z_0 I)^{-1} f_0$ gilt dann $g_0 [\perp] (A_Q - \bar{z}_0 I)\mathfrak{D}$, d.h. $g_0 \in \Gamma_{z_0}\mathfrak{G} \cap$ $\cap\, \mathfrak{D}(A_Q)$, was der Aussage von Satz 3. 2 widerspricht.

Wir schränken jetzt den Operator A_Q ein auf die Menge $\mathfrak{D}$: $\hat{A}_Q = A_Q|_\mathfrak{D}$. Es sei vermerkt, daß genau diejenigen Elemente $f = \sum e_z \xi_z \in \mathfrak{L}(Q)$[16] zum Definitionsbereich von $\hat{A}_Q$ gehören, für die $\chi(f) = \sum_{z \in \mathscr{Z}_Q} \xi_z = 0$ und $\sum_{z \in \mathscr{Z}_Q} Q(z)\xi_z = 0$ ist. In der Tat, unter der Voraussetzung $\chi(f) = 0$ sind die Beziehungen

$$[(A_Q - zI)f, e_{\bar{z}_0}\eta] = 0 \text{ für alle } \eta \in \mathfrak{G}$$

[16] Genauer, die Bilder dieser Elemente bei der kanonischen Einbettung von $\mathfrak{L}(Q)$ in $\Pi_\varkappa(Q)$.

und $\sum_{z\in\mathscr{Z}_Q} Q(z)\xi_z=0$ gleichbedeutend, denn es gilt

$$[(A_Q-z_0I)\sum e_z\xi_z, e_{\bar z_0}\eta] = \sum(z-z_0)[e_z\xi_z, e_{\bar z_0}\eta] =$$

$$= \sum(z-z_0)\left(\frac{Q(z)-Q(z_0)}{z-z_0}\xi_z, \eta\right) = \left(\sum Q(z)\xi_z, \eta\right).$$

Im Falle eines endlichdimensionalen Raumes $\mathfrak{G}$ ist der Operator $\hat{A}_Q$ die Abschließung der Einschränkung von $\breve{A}$ (siehe § 4. 1) auf die Menge

$$\mathscr{L}_0 = \{f = \sum e_z\xi_z \in \mathscr{L}(Q): \sum\xi_z = 0, \sum Q(z)\xi_z = 0\}.$$

Das folgt leicht aus einem Vergleich der Dimensionen des Faktorraumes $\mathfrak{L}(Q)/\mathfrak{L}_0$ und des Defektraumes $\Gamma_{z_0}\mathfrak{G}$.

Satz 5. 1. *Die Funktion $Q\in N_\varkappa(\mathfrak{G})$ genüge den Voraussetzungen* (III) *und* (IV) *des Hauptsatzes. Dann ist $\hat{A}_Q$ ein einfacher π-hermitescher Operator, seine Defektzahlen sind gleich, sie stimmen mit der Dimension von $\mathfrak{G}$ überein, und es ist Q eine Q-Funktion von $\hat{A}_Q$.*

Beweis. Die Einfachheit des offensichtlich π-hermiteschen Operators $\hat{A}_Q$ ergibt sich aus der Tatsache, daß der Defektraum $\mathfrak{N}_z$ von $\hat{A}_Q$ genau aus den Elementen $e_z\xi$, $\xi\in\mathfrak{G}$, besteht ($z\in\mathscr{Z}_Q$); die Gleichung $\mathfrak{n}_+(\hat{A}_Q)=\mathfrak{n}_-(\hat{A}_Q)=\dim(\Gamma_{z_0}\mathfrak{G})$ folgt unmittelbar aus der Definition von $\hat{A}_Q$. Der Operator A_Q ist eine kanonische π-selbstadjungierte Erweiterung von $\hat{A}_Q$. Deshalb ist Q auf Grund des zweiten Teiles von Satz 4. 1 eine Q-Funktion des Operators $\hat{A}_Q$, wenn man nur Γ_{z_0} gemäß § 3. 2 festlegt.

Bemerkung. Ist der Raum $\mathfrak{G}$ insbesondere endlichdimensional und die Funktion Q meromorph, dann gehört die komplexe Ebene zur Menge $\tilde{\varrho}(\hat{A}_Q)$ der Punkte regulären Typs von $\hat{A}_Q$ (d.h., für jede komplexe Zahl z ist $(\hat{A}_Q-zI)\mathfrak{D}(\hat{A}_Q)$ abgeschlossen und $z\notin\sigma_p(\hat{A}_Q)$).

Um das zu sehen, beachten wir, daß der Operator $\hat{A}_Q$ gemäß Satz 5. 1 keine Eigenwerte hat. Für seine π-selbstadjungierte Erweiterung A_Q ist der Wertebereich $(A_Q-zI)\mathfrak{D}(A_Q)$ bei beliebigem komplexem z abgeschlossen, da die Resolvente von A_Q nach Satz 4. 4 keine anderen Singularitäten als Pole hat. Die Wertebereiche $(A_Q-zI)\mathfrak{D}(A_Q)$ und $(\hat{A}_Q-zI)\mathfrak{D}(\hat{A}_Q)$ unterscheiden sich aber nur durch einen endlichdimensionalen Teilraum.

Satz 5. 2. *Ein einfacher π-hermitescher Operator A in $\Pi_\varkappa$ ist durch seine Q-Funktion bis auf π-unitäre Äquivalenz eindeutig bestimmt, genauer, A ist dem Operator $\hat{A}_Q$ π-unitär äquivalent.*

Beweis. Es sei $\tilde{A}$ die kanonische π-selbstadjungierte Erweiterung aus (2. 2). Dann sind die Operatoren A_Q und $\tilde{A}$ gemäß Satz 4. 2 π-unitär äquivalent. Wie im

Beweis von Satz 3. 3 bezeichnen wir mit T die zugehörige π-isometrische Abbildung von $\Pi_\varkappa$ auf $\Pi_\varkappa(Q)$:

$$T\mathring{\Gamma}_z\xi = e_z\xi \quad \left(z\in(C_+\cup C_-)\setminus\sigma_p(\mathring{A}),\ \xi\in\mathfrak{G}\right).$$

Aus der für beliebige $\xi\in\mathfrak{G}$ und $f\in\mathfrak{D}(\mathring{A})$ bestehenden Beziehung

$$[(\mathring{A}-z_0I)f,\ \mathring{\Gamma}_{\bar{z}_0}\xi] = [(A_Q-z_0I)\,Tf,\ e_{\bar{z}_0}\xi]$$

folgt damit leicht $T\mathfrak{D}(A)=\mathfrak{D}(\hat{A}_Q)$ und $TAf=\hat{A}_QTf$ $\left(f\in\mathfrak{D}(A)\right)$.

Folgerung 5. 1. *Die Q-Funktion eines π-hermiteschen Operators A bestimmt dessen einfachen Teil A_e bis auf π-unitäre Äquivalenz eindeutig.*

§ 6. Zusammenhang zwischen Q-Funktion und charakteristischer Funktion

Die Q-Funktion eines π-hermiteschen Operators A hängt eng mit der charakteristischen Funktion seiner π-isometrischen Cayleytransformierten V zusammen. Um diesen Zusammenhang darzustellen, beginnen wir mit den notwendigen Definitionen.

Es sei V ein π-isometrischer Operator, dessen Definitionsbereich $\mathfrak{D}(V)$ und Wertebereich $\mathfrak{R}(V)$ nichtentartete Teilräume von $\Pi_\varkappa$ sind. Die Defektzahlen von V sollen übereinstimmen, d.h., es gelte $\dim\mathfrak{D}(V)^{[\perp]}=\dim\mathfrak{R}(V)^{[\perp]}=\mathfrak{n}$. Wir wählen einen Hilbertraum $\mathfrak{G}$ der Dimension $\mathfrak{n}$ und eine eineindeutige stetige Abbildung Γ von $\mathfrak{G}$ auf $\mathfrak{R}(V)^{[\perp]}$. Weiter definieren wir einen Operator $T\in[\Pi_\varkappa,\Pi_\varkappa]$ durch die Festsetzung

$$Tf=\begin{cases} Vf & f\in\mathfrak{D}(V),\\ 0 & f\in\mathfrak{D}(V)^{[\perp]}.\end{cases}$$

Es sei $\mathring{U}$ eine π-unitäre Erweiterung von V in $\Pi_\varkappa$. Unter der (zu $\{\mathring{U},\Gamma,\mathfrak{G}\}$ gehörigen) *charakteristischen Funktion* des Operators V verstehen wir die für $|\lambda|<1$, $\lambda^{-1}\notin\sigma(T^+)$ definierte Funktion

$$X(\lambda) = \lambda\Gamma^+(I-\lambda T^+)^{-1}\mathring{U}^+\Gamma. \tag{6. 1}$$

Wählen wir speziell $\mathfrak{G}=\mathfrak{D}(V)^{[\perp]}$ und $\Gamma=\mathring{U}|_{\mathfrak{D}(V)^{[\perp]}}$, so geht die hier gegebene Definition der charakteristischen Funktion in die von [8], § 3 über. Wir überlassen es dem Leser, aus den dort erhaltenen Ergebnissen die wesentlichen Eigenschaften der charakteristischen Funktion (6. 1) abzuleiten, und vermerken nur, daß sie im Inneren des Einheitskreises meromorph ist.

Es sei jetzt V insbesondere die Cayleytransformierte des π-hermiteschen Operators A mit gleichen Defektzahlen:

$$V = (A-\bar{z}_0I)(A-z_0I)^{-1},\ z_0\,\bar{\in}\,\Delta_A,\ z_0\neq\bar{z}_0.$$

Wählen wir $\mathring{U} = (\mathring{A} - \bar{z}_0 I)(\mathring{A} - z_0 I)^{-1}$, $\Gamma = \mathring{\Gamma}_{z_0}$ mit dem in § 1. 1 eingeführten Operator $\mathring{\Gamma}_{z_0}$ und setzen noch

$$\lambda = \frac{z - \bar{z}_0}{z - z_0},$$

so besteht zwischen der Q-Funktion Q von A:

$$Q(z) = (z - x_0)\mathring{\Gamma}_{z_0}^{+} \mathring{\Gamma}_{z_0} + (z - z_0)(z - \bar{z}_0)\mathring{\Gamma}_{z_0}^{+}(\mathring{A} - zI)^{-1}\mathring{\Gamma}_{z_0}$$

und der charakteristischen Funktion X von V der Zusammenhang

$$Q(z) = -iy_0\big(\Gamma^+\Gamma + X(\lambda)\big)\big(\Gamma^+\Gamma - X(\lambda)\big)^{-1}\Gamma^+\Gamma. \tag{6. 2}$$

Bildet Γ also den Raum $\mathfrak{G}$ insbesondere π-isometrisch auf $\mathfrak{R}(V)^{[\perp]}$ ab, d.h., gilt

$$[\Gamma\xi, \Gamma\eta] = (\xi, \eta),$$

so folgt

$$Q(z) = -iy_0\big(I + X(\lambda)\big)\big(I - X(\lambda)\big)^{-1}.$$

Zum Beweis der Beziehung (6. 2) gehen wir aus von der Gleichung

$$(I - \lambda T^+)^{-1} - (I - \lambda\mathring{U}^+)^{-1} = -(I - \lambda\mathring{U}^+)^{-1}\lambda\mathring{U}^+ P(I - \lambda T^+)^{-1};$$

dabei bezeichnet P den π-orthogonalen Projektor auf $\mathfrak{R}(V)^{[\perp]}$. Multiplizieren wir diese Gleichung von links mit Γ^+ und von rechts mit $\lambda\mathring{U}^+\Gamma$, so folgt

$$X(\lambda) - \lambda\Gamma^+(I - \lambda\mathring{U}^+)^{-1}\mathring{U}^+\Gamma = -\lambda\Gamma^+(I - \lambda\mathring{U}^+)^{-1}\mathring{U}^+\overset{(-1)}{\Gamma}{}^+ X(\lambda),$$

also

$$X(\lambda) = \Gamma^+(I - \lambda\mathring{U}^+)^{-1}\lambda\mathring{U}^+\Gamma\big(I - \overset{(-1)}{\Gamma}\overset{(-1)}{\Gamma}{}^+ X(\lambda)\big),$$

$$\big(\Gamma^+\Gamma + X(\lambda)\big)\big(\Gamma^+\Gamma - X(\lambda)\big)^{-1}\Gamma^+\Gamma = \Gamma^+(\mathring{U} + \lambda I)(\mathring{U} - \lambda I)^{-1}\Gamma. \tag{6. 3}$$

Andererseits bestätigt man durch eine einfache Rechnung die Beziehung

$$-iy_0(\mathring{U} + \lambda I)(\mathring{U} - \lambda I)^{-1} = (z - x_0)I + (z - z_0)(z - \bar{z}_0)(\mathring{A} - zI)^{-1}. \tag{6. 4}$$

Aus (6. 3) und (6. 4) ergibt sich die Behauptung.

Wir bemerken, daß es dieser Zusammenhang zwischen der Q-Funktion und der charakteristischen Funktion gestattet, aus der Formel zur Beschreibung aller verallgemeinerten Resolventen eines π-isometrischen Operators in [8] die Formel (2. 7) zur Beschreibung aller verallgemeinerten Resolventen eines π-hermiteschen Operators herzuleiten.

§ 7. Rein hyperbolische Operatoren

1. *Die Menge* $\mathfrak{I}_A$. Wir betrachten in diesem Paragraphen einen π-hermiteschen Operator A mit gleichen und endlichen Defektzahlen, d.h. der Raum $\mathfrak{G}$ sei endlichdimensional. Mit $\mathfrak{I}_A$ bezeichnen wir die Menge derjenigen Punkte z aus $C_+ \cup C_-$, die Eigenwerte einer regulären π-selbstadjungierten Erweiterung von A sein können. Gemäß [4], Folgerung 3. 1, sind das genau diejenigen Punkte $z \neq \bar{z}$, für die $l_-(z) + l_0(z) > 0$ gilt, wenn sign $\mathfrak{N}_z = (l_+(z), l_0(z), l_-(z))$ die Signatur des Defektraumes $\mathfrak{N}_z$ bezeichnet. Die Beziehung (2. 3) besagt, daß für einen einfachen π-hermiteschen Operator A der Punkt z genau dann zu $\mathfrak{I}_A$ gehört, wenn für eine Q-Funktion Q von A (und damit für alle Q-Funktionen von A) $\dfrac{\operatorname{Im} Q(z)}{\operatorname{Im} z}$ nicht positiv ist.

Die Menge $\mathfrak{I}_A$ liegt im Streifen $|\operatorname{Im} z| \leqq h_A$ für eine geeignete positive Konstante h_A ([4], Satz 3. 2), sie enthält alle ihre nichtreellen Randpunkte, und diese gehören zu Δ_A.

Wir überlegen uns, daß die Menge $\mathfrak{I}_A$ beschränkt ist. Wäre dies nicht der Fall, so gäbe es eine Folge (λ_n) nichtreeller Randpunkte von $\mathfrak{I}_A$ mit $|\lambda_n| \to \infty$. Zu jedem λ_n existiert gemäß [4], Satz 3. 1 eine kanonische π-selbstadjungierte Erweiterung A_n von A mit $\lambda_n \in \sigma_0(A_n)$. Zu der π-selbstadjungierten Erweiterung A_n gehöre in der Resolventenformel (2. 7) der Operator $T_n \in \mathfrak{T}_0(\mathfrak{G})$. Definieren wir zwischen zwei Operatoren $T, T' \in \mathfrak{T}_0(\mathfrak{G})$ einen Abstand $d(T, T')$ durch die Beziehung

$$d(T, T') = \|V(T) - V(T')\|, \tag{7. 1}$$

wenn $V(T)$ die Cayleytransformierte $(T - iI)(T + iI)^{-1}$ von T bezeichnet, so sieht man leicht, daß $\mathfrak{T}_0(\mathfrak{G})$ eine kompakte Menge ist. Die Folge (T_n) enthält somit eine konvergente Teilfolge (T_{n_j}): $d(T_{n_j}, T_0) \to 0$ für $j \to \infty$ und ein $T_0 \in \mathfrak{T}_0(\mathfrak{G})$. Dann gilt aber für $z \in \varrho(A_0)$ auch $\|(A_{n_j} - zI)^{-1} - (A_0 - zI)^{-1}\| \to 0$ $(j \to \infty)$,[17]) wenn A_0 die zu T_0 gehörige kanonische π-selbstadjungierte Erweiterung von A bezeichnet. Aus Folgerung 1. 1 und der Tatsache, daß der Index der Menge $\sigma_0(A_0)$ gleich $\varkappa$ ist, ergibt sich jetzt leicht ein Widerspruch.

Wir führen weiter die Menge $\tilde{\mathfrak{I}}_A$ aller (reellen oder komplexen) Punkte z ein, die für eine reguläre π-selbstadjungierte Erweiterung $\tilde{A}$ von A zu $\sigma_0(\tilde{A})$ gehören. Offensichtlich gilt $\mathfrak{I}_A = \tilde{\mathfrak{I}}_A \cap (C_+ \cup C_-)$. Ist der Durchschnitt einer Komponente von $\tilde{\mathfrak{I}}_A$ mit C_+ nicht leer, so nennen wir diesen eine *Komponente* von $\mathfrak{I}_A$ in C_+. Durch Betrachtung der nichtreellen Randpunkte von $\mathfrak{I}_A$ ergibt sich aus Folgerung 1. 1 leicht der

[17]) Davon überzeugt man sich leicht z.B. mit Hilfe der in [8], § 4, gegebenen Darstellung der verallgemeinerten Resolventen.

Satz 7. 1. *Die Menge $\mathfrak{J}_A$ hat in C_+ höchstens $\varkappa$ Komponenten.*

Die Menge $\mathfrak{J}_A$ läßt sich genauer beschreiben, wenn mann den Begriff des definisierenden Polynoms (vgl. [12], [13]) heranzieht. Wir führen das im Falle $\varkappa=1$ aus.

Ein reelles Polynom $p(z) = (z-\alpha)(z-\bar{\alpha})$ zweiten Grades nennen wir *definisierend* für den π-hermiteschen Operator A in Π_1, wenn

$$(7.2) \qquad [(A-\alpha I)f, (A-\alpha I)f] \geqq 0 \quad \text{für alle} \quad f \in \mathfrak{D}(A)$$

gilt. Die Menge aller definisierenden Polynome von A werde mit $\mathfrak{P}(A)$ bezeichnet. Offensichtlich ist $\mathfrak{P}(A)$ konvex, d.h., aus $p_1, p_2 \in \mathfrak{P}(A)$ folgt auch $\gamma p_1 + (1-\gamma)p_2 \in \mathfrak{P}(A)$ für $0 \leqq \gamma \leqq 1$.

Wir überlegen uns, daß die Zahl α genau dann zu $\tilde{\mathfrak{J}}_A$ gehört, wenn sie Nullstelle eines Polynoms $p \in \mathfrak{P}(A)$ ist. In der Tat, aus $\alpha \in \tilde{\mathfrak{J}}_A$ folgt, daß das Polynom $p(z) = (z-\alpha)(z-\bar{\alpha})$ definisierend für die Erweiterung $\tilde{A}$ von A mit $\alpha \in \sigma_0(\tilde{A})$ ist. Dann ist es aber auch definisierend für A. Besteht umgekehrt die Beziehung (7. 2), so enthält $\mathfrak{N}_{\bar{\alpha}}$ ein nichtpositives Element $f_0 \neq 0$. Im Falle $\alpha \neq \bar{\alpha}$ folgt damit aus [4], Folgerung 3. 1, daß eine reguläre π-selbstadjungierte Erweiterung $\tilde{A}$ von A mit $\alpha \in \sigma_0(\tilde{A})$ existiert. Ist $\alpha = \bar{\alpha}$, so ergibt sich durch die Festsetzung $\tilde{A}f = Af$ für $f \in \mathfrak{D}(A)$ und $\tilde{A}f_0 = \alpha f_0$ — sofern f_0 nicht bereits zu $\mathfrak{D}(A)$ gehört — sowie anschließende Fortsetzung des erhaltenen Operators zu einem π-selbstadjungierten Operator ebenfalls eine π-selbstadjungierte Erweiterung $\tilde{A}$ von A, für die $\alpha \in \sigma_0(\tilde{A})$ gilt.

Wir setzen jetzt $\tilde{C}_+ = C_+ \cup R$ und führen in C_+ die Abbildung $z \to \varphi(z) = 2\,\mathrm{Re}\, z + i|z|^2$ ein. Aus der Konvexität der Menge $\mathfrak{P}(A)$ und der obigen Überlegung ergibt sich dann der

Satz 7. 2. *Für einen π-hermiteschen Operator A mit gleichen Defektzahlen im Raume Π_1 ist das Bild $\varphi(\tilde{\mathfrak{J}}_A \cap \tilde{C}_+)$ konvex.*

2. *Rein hyperbolische Operatoren.* Wir nennen einen π-selbstadjungierten Operator A in $\Pi_\varkappa$ *rein hyperbolisch*, wenn sein Spektrum in C_+ aus genau $\varkappa$ Eigenwerten — jeder entsprechend seiner Vielfachheit oft gezählt — besteht. In diesem Falle enthält $\sigma_0(A)$ keine reellen Punkte, und der Pontrjaginsche $\varkappa$-dimensionale nichtpositive Teilraum $\mathfrak{L}$ ist z. B. durch die Bedingung $\mathrm{Im}\,\sigma(A|\mathfrak{L})>0$ eindeutig bestimmt: Er wird von denjenigen algebraischen Eigenräumen von A aufgespannt, die zu den Eigenwerten von A in der oberen Halbebene gehören.

Ein π-hermitescher Operator A in $\Pi_\varkappa$ heiße *rein hyperbolisch*, wenn alle seine regulären π-selbstadjungierten Erweiterungen rein hyperbolisch sind.

Hauptergebnis dieses Paragraphen ist der folgende

Satz 7. 3. *Der π-hermitesche Operator A in $\Pi_\varkappa$ mit gleichen endlichen Defektzahlen ist genau dann rein hyperbolisch, wenn mindestens eine kanonische π-selbst-*

adjungierte Erweiterung rein hyperbolisch ist und die Menge $\mathfrak{J}_A$ einen positiven Abstand von der reellen Achse hat.

Zum Beweis der Notwendigkeit hat man nur zu zeigen, daß die Menge der Randpunkte von $\mathfrak{J}_A$ von der reellen Achse einen positiven Abstand hat. Das ergibt sich leicht aus der Folgerung 1. 1.

Die Hinlänglichkeit zeigen wir in mehreren Schritten; dabei seien für A stets die im Satz genannten Voraussetzungen erfüllt, und $\mathring{A}$ bezeichne eine rein hyperbolische kanonische π-selbstadjungierte Erweiterung von A.

1) *Jede kanonische π-selbstadjungierte Erweiterung von A ist rein hyperbolisch.* Um das zu sehen, wählen wir eine Kontur $\mathscr{C}$, die ganz in C_+ liegt und $\mathfrak{J}_A \cap C_+$ genau einmal im positiven Sinne umläuft. Ist R_z eine kanonische Resolvente von A und T der gemäß (2. 7) zugehörige Operator aus $\mathfrak{T}_0(\mathfrak{G})$, so wird durch die Formel

$$\varkappa(T) = \dim\left(-\frac{1}{2\pi i}\int\limits_{\mathscr{C}} R_z\,dz\right)\Pi_\varkappa$$

eine stetige Funktion auf dem zusammenhängenden kompakten Raum $\mathfrak{T}_0(\mathfrak{G})$, versehen mit der Metrik (7. 1), definiert. Deshalb ist $\varkappa(T)$ konstant. Da andererseits nach Voraussetzung $\varkappa(T_0) \doteq \varkappa$ für $T_0 = \infty \cdot I$ gilt, folgt $\varkappa(T) = \varkappa$ für alle $T \in \mathfrak{T}_0(\mathfrak{G})$, woraus sich leicht die Behauptung ergibt.

2) *Jede reine Austrittserweiterung von A ist rein hyperbolisch.* Um das zu zeigen, stellen wir A zunächst in der Form (1. 4) mit seinem einfachen Teil A_e dar. Dann gilt $\mathfrak{J}_{A_e} \subset \mathfrak{J}_A$, und man sieht leicht, daß genau dann alle reinen Austrittserweiterungen von A rein hyperbolisch sind, wenn alle reinen Austrittserweiterungen von A_e rein hyperbolisch sind. Deshalb brauchen wir die Aussage 2) nur für einen einfachen Operator A zu beweisen. Wir betrachten eine kanonische π-selbstadjungierte Erweiterung $\mathring{A}$ von A und die zugehörige Q-Funktion (2. 2); dabei sei $\mathfrak{G}$ der Defektraum $\mathfrak{N}_{z_0}$, $\operatorname{Im} z_0 > h_A$, versehen mit dem π-Skalarprodukt von $\Pi_\varkappa$, und Γ_{z_0} die identische Abbildung von $\mathfrak{G}$ auf $\mathfrak{N}_{z_0}$. Da $\mathring{A}$ und damit auch A_Q rein hyperbolisch sind, ist die Polvielfachheit von $\mathfrak{J}_A \cap C_+$ bezüglich $Q(z)$ gemäß Folgerung 4. 1 gleich $\varkappa$.

Nach der Bemerkung am Ende von § 2. 3 ist mit $Q(z)$ auch $-y_0^2 Q^{-1}(z)$ eine Q-Funktion des Operators A. Für diese hat $\mathfrak{J}_A \cap C_+$ ebenfalls die Polvielfachheit $\varkappa$, also ist auch die Nullvielfachheit von $\mathfrak{J}_A \cap C_+$ bezüglich $Q(z)$ gleich $\varkappa$.

Es sei jetzt $\tilde{A}$ eine reine Austrittserweiterung von A. Dann ist das zugehörige $T \in \mathfrak{T}(\mathfrak{G})$ aus der Resolventenformel (2. 7) eine eigentliche Funktion (siehe § 2. 2), d.h., in der Darstellung (2. 6) gilt $\hat{P} = I$. Wir betrachten den Ausdruck

$$\frac{1}{2\pi i}\operatorname{sp}\int\limits_{\mathscr{C}} Q'(z)+\varepsilon T'(z))\big(Q(z)+\varepsilon T(z)\big)^{-1}\,dz, \qquad 0 \leqq \varepsilon \leqq 1,$$

wenn $\mathscr{C}$ dieselbe Bedeutung wie im ersten Teil des Beweises hat. Er hängt in dem angegebenen Bereich stetig von ε ab, und sein Wert ist eine ganze Zahl. Da er für $\varepsilon=0$ auf Grund von (4. 18) und dem oben Gesagten den Wert Null hat, gilt dasselbe für $\varepsilon=1$, d.h., für $Q(z)+T(z)$ stimmt die Nullvielfachheit von $\mathfrak{J}_A \cap C_+$ mit der Polvielfachheit von $\mathfrak{J}_A \cap C_+$ überein. Letztere ist aber gleich $\varkappa$, da dies für $Q(z)$ gilt und $T(z)$ in $\mathfrak{J}_A \cap C_+$ holomorph ist. Also ist auch die Nullvielfachheit von $\mathfrak{J}_A \cap C_+$ bezüglich $Q(z)+T(z)$ gleich $\varkappa$.

Schreiben wir die Formel (2. 7) jetzt in der Form

$$R_z = (\mathring{A}-z_0 I)(\mathring{A}-zI)^{-1}[(\mathring{A}-z_0 I)^{-1}-\Gamma_{z_0}(Q(z)+T(z))^{-1}\Gamma_{\bar{z}}^+],$$

so sieht man (vgl. [20]), daß die Polvielfachheit von $\mathfrak{J}_A \cap C_+$ bezüglich R_z mindestens gleich $\varkappa$ ist. Da sie andererseits mit der Summe der algebraischen Vielfachheiten aller in $\mathfrak{J}_A \cap C_+$ gelegenen Eigenwerte von $\tilde{A}$ übereinstimmt, ist $\tilde{A}$ rein hyperbolisch.

3) Es sei jetzt $\tilde{A}$ eine beliebige reguläre π-selbstadjungierte Erweiterung von A. Dann gibt es eine echte oder unechte π-hermitesche Erweiterung $\hat{A}$ von A in $\Pi_\varkappa$, so daß $\tilde{A}$ eine reine Austrittserweiterung von $\hat{A}$ ist. Offensichtlich gilt $\mathfrak{J}_{\hat{A}} \subset \mathfrak{J}_A$, also hat $\mathfrak{J}_{\hat{A}}$ einen positiven Abstand von der reellen Achse. Da jede kanonische Erweiterung von $\hat{A}$ eine kanonische Erweiterung von A ist, hat $\hat{A}$ eine rein hyperbolische kanonische π-selbstadjungierte Erweiterung. Auf Grund des zweiten Teiles des Beweises ist somit $\tilde{A}$ rein hyperbolisch. Damit ist der Satz bewiesen.

Folgerung 7. 2. *Ein π-hermitescher Operator A mit gleichen endlichen Defektzahlen ist genau dann rein hyperbolisch, wenn alle seine kanonischen Erweiterungen rein hyperbolisch sind.*

Wir bemerken, daß sich der Beweis von Satz 7. 3 im Falle $\mathfrak{n}_\pm(A)=1$ wesentlich vereinfacht. Gehört unter dieser Voraussetzung die komplexe Ebene sogar zur Menge $\tilde{\varrho}(A)$ der Punkte regulären Typs von A, so ist A genau dann rein hyperbolisch, wenn die Menge $\mathfrak{J}_A$ einen positiven Abstand von der reellen Achse hat.

Um das zu zeigen nehmen wir an, es gäbe unter dieser Voraussetzung an $\mathfrak{J}_A$ eine kanonische Erweiterung A_0 von A mit einem reellen Eigenwert $\lambda \in \sigma_0(A_0)$. Das zugehörige Eigenelement sei z. B. negativ, d.h., λ sei von negativem Typ ([4], § 6. 2). Entspricht der Erweiterung A_0 in der Resolventenformel (2. 7) die Konstante $\tau_0 : T(z)\equiv\tau_0$, $A_0=A^{(\tau_0)}$, so hat gemäß [4], § 6. 2 für hinreichend nahe bei τ_0 gelegene $\tau>\tau_0$ der Operator $A^{(\tau)}$ einen Eigenwert $\lambda(\tau)$ negativen Typs, der sich mit wachsendem τ nach rechts bewegt. Da sich andererseits die Eigenwerte positiven Typs mit wachsendem τ nach links bewegen, gibt es einen Wert τ_1, so daß $\lambda(\tau_1)=\lim\limits_{\tau\uparrow\tau_1}\lambda(\tau)$ ein kritischer Eigenwert von $A^{(\tau_1)}$ ist. Für hinreichend nahe bei τ_1 gelegene $\tau>\tau_1$ hat dann der Operator $A^{(\tau)}$ gemäß [4], § 6. 2 nichtreelle Eigenwerte mit beliebig kleinem Imaginärteil, was der Voraussetzung über die Menge $\mathfrak{J}_A$ widerspricht.

Diese Bemerkung gibt die Möglichkeit, rein hyperbolische π-hermitesche Operatoren zu konstruieren. Man braucht nur eine komplexe Funktion $Q(z)$ mit den Eigenschaften (I)—(III) des Hauptsatzes aus § 2. 4 (für den eindimensionalen Raum $\mathfrak{G}$ der komplexen Zahlen) zu wählen, deren Singularitäten nur isolierte Pole sind und deren Menge $\mathfrak{J}_Q$ aller Punkte $z \in C_+ \cup C_-$ mit $\frac{\operatorname{Im} Q(z)}{\operatorname{Im} z} \leqq 0$ einen positiven Abstand von der reellen Achse hat. Dann ist der Operator $\hat{A}_Q$ rein hyperbolisch, denn auf Grund der Bemerkung im Anschluß an Satz 5. 3 gilt $\tilde{\varrho}(\tilde{A}_Q) = C$.

Im Falle $\varkappa = 1$ hat z. B. die Funktion

$$Q(z) = \sum_{j=1}^{\infty} \varrho_j \left\{ \frac{1}{\lambda_j - z} - \frac{1}{\lambda_j} \right\} + \frac{\varepsilon}{\alpha - z} + \frac{\varepsilon}{\bar{\alpha} - z}$$

mit $\varrho_j > 0$, $\sum_{j=1}^{\infty} \varrho_j = \infty$, $\sum_{j=1}^{\infty} \frac{\varrho_j}{\lambda_j^2} < \infty$, $\lambda_j = \bar{\lambda}_j$, $\lambda_j \neq 0$, $j = 1, 2, \ldots$, $\alpha \neq \bar{\alpha}$, für hinreichend kleines $\varepsilon > 0$ alle verlangten Eigenschaften.

Wie jeder π-hermitesche Operator A mit endlichen Defektzahlen und $\tilde{\varrho}(A) = C$ hat der Operator $\hat{A}_Q$ Erweiterungen mit Austritt aus Π_1, deren reelles Spektrum absolutstetig ist. Solche ergeben sich z. B., wenn man in der Formel (2. 7) $T(z) \equiv i$ $(z \in C_+)$ setzt.

Literatur

[1] М. Г. Крейн, Об эрмитовых операторах с дефектиными индексами, равными единице, *ДАН СССР*, **43** (8) (1944), 339—342.

[2] М. Г. Крейн, О резольвентах эрмитова оператора с индексом дефекта (m, m), *ДАН СССР*, **52** (8) (1946), 657—660.

[3] Ш. Н. Саакян, К теории резольвент симметрического оператора с бесконечными дефектными числами, *ДАН Арм. ССР*, **41** (4) (1965), 193—198.

[4] М. Г. Крейн — Г. К. Лангер, О дефектных подпространствах и обобщённых резольвентах эрмитова оператора в пространстве $\Pi_\varkappa$, *Функциональный анализ*, **5** (2) (1971), 61—73; **5** (3) (1971), 54—69.

[5] G. Pick, Über die Beschränkungen analytischer Funktionen, welche durch vorgegebene Funktionswerte bewirkt werden, *Math. Annalen*, **77** (1916), 7—23.

[6] R. Nevanlinna,, Über beschränkte analytische Funktionen, *Ann. Acad. Sci. Fenn. Ser. A*, *XXXII*, **7** (1929), 1—75.

[7] B. Sz.-Nagy — A. Korányi, Operatortheoretische Behandlung und Verallgemeinerug eines Problemkreises in der komplexen Funktionentheorie, *Acta Math.*, **100** (1958), 171—202.

[8] M. G. Kreĭn — H. Langer, Über die verallgemeinerten Resolventen und die charakteristische Funktion eines isometrischen Operators im Raume $\Pi_\varkappa$, *Colloquia Mathematica Societatis János Bolyai. 5. Hilbert Space Operators*, Tihany (Hungary), 1970, 353—399.

[9] М. С. Лившиц, Об одном классе линейных операторов в гильбертовом пространстве, *Матем. сб.*, **19** (61) (1946), 239—260.

[10] В. М. Адамян — Д. З. Аров — М. Г. Крейн, Аналитические свойства пар Шмидта — Ганкелева оператора и обобщённая задача Шура—Такаги, *Матем. сб.*, **86 (128)**, 1 (9) (1971), 34—75.

[11] P. D. LAX — R. S. PHILLIPS, *Scattering Theory*, Academic Press (New York—London, 1967).

[12] М. Г. Крейн — Г. К. Лангер, О спектральной функции самосопряжённого оператора в пространстве с индефинитной метрикой, *ДАН СССР*, **152** (1) (1963), 39—42.

[13] H. LANGER, *Spektraltheorie linearer Operatoren in J-Räumen und einige Anwendungen auf die Schar* $L(\lambda) = \lambda^2 I + \lambda B + C$, Habilitationsschrift, Technische Universität Dresden, 1965.

[14] М. Г. Крейн, Основные положения теории представления эрмитовых операторов с индексом дефекта (m, m), *Укр. матем. ж.*, **1**, **2** (1949), 3—66.

[15] И. Ц. Гохберг — М. Г. Крейн, Основные положения о дефектных числах, корневых числах и индексах линейных операторов, *Успехи матем. наук*, **12**, **2** (1957), 43—118.

[16] H. LANGER, Zur Spektraltheorie J-selbstadjungierter Operatoren, *Math. Annalen*, **146** (1962), 60—85.

[17] И. С. Иохвидов — М. Г. Крейн, Спектральная теория линейных операторов в пространствах с индефинитной метрикой. II, *Труды Моск. матем. общества*, **8** (1959), 413—496.

[18] М. Г. Крейн, Об одной экстраполяционной проблеме А. Н. Колмогорова, *ДАН СССР*, **46**, **8** (1945), 339—342.

[19] Н. И. Ахиезер, *Лекции по теории аппроксимации*, Наука (Москва, 1965).

[20] И. Ц. Гохберг, О некоторых вопросах спектральной теории конечномероморфных оператор-функций, *Известия Акад. Наук Арм. ССР*, **6** (1971), 160—181; см. также „Письмо в редакцию", *там же*, **7** (1972).

[21] И. Ц. Гохберг — М. Г. Крейн, *Введение в теорию несамосопряжённых операторов*, Наука (Москва, 1965).

[22] И. Ц. Гохберг — Е. И. Сигал, Операторное обобщение теоремы о логарифмическом вычёте и теоремы Руше, *Матем. сб.*, **84 (126)** (1971), 607—630.

(Eingegangen am 22. Juni 1972)

On imbedding theorems

By L. LEINDLER in Szeged

Dedicated to Professor B. Szőkefalvi-Nagy on his 60th birthday

Let $\varphi(x)\equiv\varphi_p(x)$ $(p\geqq 1)$ be a nonnegative increasing function on $[0,\infty)$ with the following properties:

(1) $$\frac{\varphi(x)}{x}\uparrow \quad\text{and}\quad \frac{\varphi(x)}{x^p}\downarrow \quad\text{as}\quad x\to\infty.$$

The set of the measurable functions $f(x)$ on $[0, 1]$ for which $\int\limits_0^1 \varphi(|f(x)|)\,dx<\infty$ will be denoted by $\varphi(L)$.

If $f(x)\in\varphi(L)$ then the "modulus of continuity of $f(x)$ with respect to φ" will be defined by

$$w_\varphi(\delta,f)=\sup_{0\leqq h\leqq\delta}\bar\varphi\Big(\int\limits_0^{1-h}\varphi(|f(x+h)-f(x)|)\Big)\,dx \qquad (0\leqq\delta\leqq 1),$$

where $\bar\varphi(x)$ denotes the inverse function of $\varphi(x)$.

If $\varphi(x)=x^p$ $(p\geqq 1)$ then $\varphi(L)$ and $\omega_\varphi(\delta,f)$ will be denoted, as usual, by L^p and $\omega_p(\delta,f)$, respectively.

P. L. UL'JANOV has proved imbedding theorems in several papers (see for instance [4] and [5]). Among other things, he gave conditions which assure that a function $f(x)\in L^p$ $(p\geqq 1)$ should belong to another space L^ν $(\nu>p)$. A sample theorem is as follows (see [4], Theorem 1): If $f(x)\in L^p$ $(p\geqq 1)$ and

(2) $$\sum_{n=1}^{\infty} n^{\frac{\nu}{p}-2}\,\omega_p^\nu\left(\frac{1}{n},f\right)<\infty,$$

then $f(x)\in L^\nu$.

In [1] we generalized some of Ul'janov's results and gave conditions assuring the transition from L^p to $L^p(\ln^+ L)^\beta$ (see Corollary 1, in case $p=1$ see also Theorem 2 in [4]) and from L^p to $\varphi_\nu(L)$ (see Theorem 2). The latter result states that if $f(x)\in L^p$ and

(3) $$\sum_{n=1}^{\infty}\frac{1}{n^2}\,\varphi_\nu\left(n^{1/p}\omega_p\left(\frac{1}{n}\right)\right)<\infty$$

then $f(x)\in\varphi_\nu(L)$.

It is clear that in the special case $\varphi_\nu(x)=x^\nu$ (3) reduces to (2).

In the present paper we are going to give conditions which assure the transition from an arbitrary collection $\varphi_p(L)$ to another $\varphi_\nu(L)$.

More precisely we prove the following theorems

Theorem 1. *Let $f(x)\in\varphi(L)$ $\left(\varphi(x)\equiv\varphi_p(x), p\geqq 1\right)$ and let $\{\lambda_k\}$ be a nonnegative monotonic sequence of numbers such that*

$$\sum_{k=m}^{\infty}\frac{\lambda_k}{k^{1+\varepsilon}}\leqq K(\lambda)\frac{\lambda_m}{m^\varepsilon},\text{ [1]} \tag{4}$$

where $\varepsilon=(4[p+1]+2)^{-1}$, [2] *and furthermore let $\Lambda(x)=\sum_{k=1}^{x}\frac{\lambda_k}{k}$.* [3]

Then $\sum_{n=1}^{\infty}\frac{\lambda_n}{n}\varphi\left(\omega_\varphi\left(\frac{1}{n},f\right)\right)<\infty$ implies $f(x)\in\varphi(L)\Lambda(L)$, and

$$\int_0^1\varphi(|f(x)|)\Lambda(|f(x)|)\,dx\leqq K(\varphi,\lambda)\left\{\sum_{n=1}^{\infty}\frac{\lambda_n}{n}\varphi\left(\omega_\varphi\left(\frac{1}{n},f\right)\right)+\int_0^1\varphi(|f(x)|)\,dx\right\}. \tag{5}$$

Theorem 2. *Set $\varphi(x)\equiv\varphi_p(x)$ and $\psi(x)\equiv\varphi_\nu(x)$ $(p,\nu\geqq 1)$. Let $\{\varrho_k\}$ be a nonnegative nondecreasing sequence of numbers with*

$$\sum_{k=m}^{\infty}\frac{\varrho_k}{k^2}\leqq K(\varrho)\frac{\varrho_m}{m}, \tag{6}$$

and denote by $\varrho(x)$ the continuous function which is linear between n and $n+1$, furthermore $\varrho(n)=\varrho_n$. Suppose that $f(x)\in\varphi(L)$ and

$$\sum_{n=1}^{\infty}\frac{\varrho_n}{n^2}\psi\left(\bar\varphi\left(n\varphi\left(\omega_\varphi\left(\frac{1}{n},f\right)\right)\right)\right)<\infty.$$

Then $f(x)\in\psi(L)\varrho(L)$, and

$$\begin{aligned}&\int_0^1\psi(|f(x)|)\varrho(|f(x)|)\,dx\leqq\\ &\leqq K(\varphi,\psi,\varrho)\left\{\sum_{n=1}^{\infty}\frac{\varrho_n}{n^2}\psi\left(\bar\varphi\left(n\varphi\left(\omega_\varphi\left(\frac{1}{n},f\right)\right)\right)\right)+\psi\left(\int_0^1\varphi(|f(x)|)\,dx\right)\right\}.\end{aligned} \tag{7}$$

---

[1]) K and K_i, denote either absolute constants or constants depending on certain functions and numbers which are not necessary to specify; $K(\alpha,\beta,\ldots)$ and $K_i(\alpha,\beta,\ldots)$ denote positive constants depending only on the indicated parameters. These constants are not necessarily the same at each occurrence.

[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.

The methods of proof of these theorems are similar to those of Theorem 1 and 2 given in [1], but new lemmas, using the modulus of continuity $\omega_\varphi\left(\frac{1}{n}, f\right)$ instead of $\omega_p\left(\frac{1}{n}, f\right)$, have to be introduced. The proofs of these lemmas run similarly to those of the old ones which were proved partly by P. L. UL'JANOV [4] and me [1].

These theorems have the following two corollaries, which have so far been proved only for $\varphi(x) \equiv x^p$ and $\varrho_k \equiv 1$ (see in [1] Corollary 1 and Theorem 2).

Corollary 1. *If* $f(x) \in \varphi(L)$, $\beta > -1$ *and*

$$\sum_{n=2}^{\infty} \frac{1}{n} (\ln n)^\beta \varphi\left(\omega_\varphi\left(\frac{1}{n}, f\right)\right) < \infty,$$

then $f(x) \in \varphi(L)(\ln^+ L)^{\beta+1}$.

Corollary 2. *If* $f(x) \in L^p$ $(p \geqq 1)$ *and*

$$\sum_{n=1}^{\infty} \frac{\varrho_n}{n^2} \psi\left(n^{1/p} \omega_p\left(\frac{1}{n}, f\right)\right) < \infty,$$

then $f(x) \in \psi(L)\varrho(L)$, *where* $\psi(x)$, ϱ_k *and* $\varrho(x)$ *have the same meanings as in Theorem* 2.

From Corollary 2, by choosing special ϱ_k and $\psi(x)$, we obtain two other corollaries.

Corollary 3. *If* $f(x) \in L^p$ $(p \geqq 1)$, $0 \leqq \alpha < 1$, $\beta \geqq 0$ *and*

$$\sum_{n=2}^{\infty} \frac{(\ln n)^\beta}{n^{2-\alpha}} \psi\left(n^{1/p} \omega_p\left(\frac{1}{n}, f\right)\right) < \infty,$$

then $f(x) \in \psi(L) L^\alpha (\ln^+ L)^\beta$.

Corollary 4. *If* $f(x) \in L^p$ $(p \geqq 1)$, $\nu \geqq p$, $0 \leqq \alpha < 1$, $\beta \geqq 0$ *and*

$$\sum_{n=2}^{\infty} (\ln n)^\beta n^{\frac{\nu}{p}+\alpha-2} \omega_p^\nu\left(\frac{1}{n}, f\right) < \infty,$$

then $f(x) \in L^{\nu+\alpha}(\ln^+ L)^\beta$.

§ 1. Lemmas

We require the following lemmas.

Lemma 1. ([4], Lemma 7'.) *If* $f(x)\in L(0,1)$ *and* $F(z)$ *is a nonnegative nonincreasing function equidistributed with* $|f(x)|$, *that is,*

$$\text{mes}\{x: x\in[0,1], |f(x)|>y\} = \text{mes}\{z: z\in[0,1], F(z)>y\},$$

then

$$\sup_{\substack{E\subset[0,1]\\|E|=\alpha}} \int_E |f(x)|\,dx = \int_0^\alpha F(z)\,dz$$

for any $0\leqq\alpha\leqq 1$; *furthermore if* $0\leqq\alpha\leqq\frac{1}{2}$ *and*

$$\sup_{\substack{E\subset[0,1]\\|E|=\alpha}} \int_E |f(x)|\,dx = \int_{E_0} |f(x)|\,dx,$$

then

$$\sup_{\substack{E\subset[0,1]-E_0\\|E|=\alpha}} \int_E |f(x)|\,dx = \int_\alpha^{2\alpha} F(z)\,dz.$$

Lemma 2. ([5], Remark 1.) *If* $f(x)\in\varphi(L)$ *and* $0\leqq a<b\leqq 1$, *then*

$$\int_a^b\int_a^b \varphi(|f(x)-f(y)|)\,dx\,dy = 2\int_0^{b-a}\Big\{\int_a^{b-u}\varphi(|f(u+y)-f(y)|)\,dy\Big\}\,du.$$

Lemma 3. *Let* $\varphi(x)\equiv\varphi_p(x)$. *If* $u(x)$ *and* $v(x)$ *are nonnegative measurable functions on the interval* I, *then we have*

$$\varphi\left(\frac{\int_I u(x)\,v(x)\,dx}{\int_I u(x)\,dx}\right) \leqq 2^p\,\frac{\int_I u(x)\,\varphi(v(x))\,dx}{\int_I u(x)\,dx}.$$

This lemma immediately follows from results of H. P. Mulholland [2] (see Theorem 1 and Remark 2. 34).

Lemma 4. *If* $a_n\geqq 0$ *and* $\lambda_n>0$, *then*

$$\sum_{n=1}^{\infty}\lambda_n\,\varphi\left(\sum_{i=1}^{n}a_i\right) \leqq K(\varphi)\sum_{n=1}^{\infty}\lambda_n\,\varphi\left(\frac{a_n}{\lambda_n}\sum_{k=n}^{\infty}\lambda_k\right).$$

This lemma is a part of Theorem of J. Németh [3] (see the inequality (8) of Theorem).

Lemma 5. *If $f(x)\in\varphi(L)$ and*

$$\psi_n(x)=n\int_{\frac{k}{n}}^{\frac{k+1}{n}} f(t)\,dt \quad \text{for} \quad x\in\left[\frac{k}{n},\frac{k+1}{n}\right), \qquad 0\leqq k<n,$$

then

$$\int_0^1 \varphi(|f(t)-\psi_n(t)|)\,dt \leqq K(\varphi)\,\varphi\left[\omega_\varphi\left(\frac{1}{n},f\right)\right]. \tag{1.1}$$

Proof. Using Lemma 3, we obtain that

$$\int_0^1 \varphi(|f(t)-\psi_n(t)|)\,dt=\sum_{k=0}^{n-1}\int_{\frac{k}{n}}^{\frac{k+1}{n}} \varphi\left(\left|f(t)-n\int_{\frac{k}{n}}^{\frac{k+1}{n}} f(x)\,dx\right|\right)dt\leqq$$

$$\leqq\sum_{k=0}^{n-1}\int_{\frac{k}{n}}^{\frac{k+1}{n}} \varphi\left(n\int_{\frac{k}{n}}^{\frac{k+1}{n}} |f(t)-f(x)|\,dx\right)dt\leqq \tag{1.2}$$

$$\leqq K_1(\varphi)\sum_{k=0}^{n-1}\int_{\frac{k}{n}}^{\frac{k+1}{n}}\left(n\int_{\frac{k}{n}}^{\frac{n+1}{n}} \varphi(|f(t)-f(x)|)\,dx\right)dt\equiv I_1.$$

Next we use Lemma 2 and have

$$I_1\leqq 2K_1 n\sum_{k=0}^{n-1}\int_0^{\frac{1}{n}}\left\{\int_{\frac{k}{n}}^{\frac{k+1}{n}-u} \varphi(|f(u+y)-f(y)|)\,dy\right\}du\leqq$$

$$\leqq K_2 n\int_0^{\frac{1}{n}}\left\{\int_0^{1-u} \varphi(|f(u+y)-f(y)|)\,dy\right\}du\leqq K_2\,\varphi\left[\omega_\varphi\left(\frac{1}{n},f\right)\right].$$

From this and (1. 2) we obtain (1. 1).

Lemma 6. *If $f(x)\in\varphi(L)$ then*

$$\omega_\varphi\left(\frac{1}{n},f\right)\geqq K(\varphi)\,\bar\varphi\left[\frac{1}{n}\,\varphi\left(n\left(\int_0^{\frac{1}{n}} F(z)\,dz-\int_{\frac{1}{n}}^{\frac{2}{n}} F(z)\,dz\right)\right)\right],$$

where $\bar\varphi(x)$ denotes the inverse function of $\varphi(x)$ and $F(z)$ has the same meaning as in Lemma 1.

Proof. Set

$$\alpha(t) \equiv \alpha_n(t) = n \int\limits_{\frac{k}{n}}^{\frac{k+1}{n}} |f(u)|\,du = a_k \geqq 0 \quad \text{for} \quad t \in \left[\frac{k}{n}, \frac{k+1}{n}\right)$$

($k = 0, 1, \dots, n-1$). Denote by $0 \leqq b_0 \leqq b_1 \leqq \dots \leqq b_{n-1}$ the nondecreasing rearrangement of the sequence $\{a_k\}_0^{n-1}$. Then

$$\int\limits_0^{1-\frac{1}{n}} \varphi\left(\left|\alpha\left(t+\frac{1}{n}\right)-\alpha(t)\right|\right) dt = \sum_{k=0}^{n-2} \int\limits_{\frac{k}{n}}^{\frac{k+1}{n}} \varphi\left(\left|\alpha\left(t+\frac{1}{n}\right)-\alpha(t)\right|\right) dt \geqq \tag{1.3}$$

$$\geqq \sum_{k=0}^{n-2} \varphi(|a_{k+1}-a_k|)\frac{1}{n} \geqq \frac{1}{n}\varphi(b_{n-1}-b_{n-2}).$$

Set $B = \{t : t \in [0, 1],\ \alpha(t) = b_{n-1}\}$. Then it is clear that for arbitrary sets $B_1 \subset B$ with $|B_1| = \frac{1}{n}$ and $E \subset [0, 1]$ with $|E| = \frac{1}{n}$ we have

$$b_{n-1} - b_{n-2} = n\left(\int\limits_{B_1} \alpha(t)\,dt - \frac{b_{n-2}}{n}\right) \geqq n\Big(\int\limits_E \alpha(t)\,dt - \sup_{\substack{E_1 \subset [0,1]-E \\ |E_1| = \frac{1}{n}}} \int\limits_{E_1} \alpha(t)\,dt\Big),$$

and thus

$$b_{n-1} - b_{n-2} \geqq n \sup_{\substack{E \subset [0,1] \\ |E| = \frac{1}{n}}} \Big\{\int\limits_E \alpha(t)\,dt - \sup_{\substack{E_1 \subset [0,1]-E \\ |E_1| = \frac{1}{n}}} \int\limits_{E_1} \alpha(t)\,dt\Big\}, \tag{1.4}$$

Since $\varphi(|a|) \leqq \varphi(|a+b|+|b|) \leqq K(\varphi)\{\varphi(|a+b|)+\varphi(|b|)\}$, we obatin

$$\varphi(|a+b|) \geqq K_1\varphi(|a|) - \varphi(|b|). \tag{1.5}$$

Using (1.3), (1.4) and (1.5) we have

$$\varphi\left(\omega_\varphi\left(\frac{1}{n}, f\right)\right) \geqq \varphi\left(\omega_\varphi\left(\frac{1}{n}, |f|\right)\right) \geqq$$

$$\geqq \int\limits_0^{1-\frac{1}{n}} \varphi\left(\left|\left|f\left(t+\frac{1}{n}\right)\right| - \alpha\left(t+\frac{1}{n}\right) + \alpha\left(t+\frac{1}{n}\right) - \alpha(t) + \alpha(t) - |f(t)|\right|\right) dt \geqq$$

$$\geqq K_1 \int\limits_0^{1-\frac{1}{n}} \varphi\left(\left|\alpha\left(t+\frac{1}{n}\right)-\alpha(t)\right|\right) dt - K_2 \int\limits_0^1 \varphi(||f(t)|-\alpha(t)|)\,dt \geqq$$

$$\geqq \frac{K_1}{n}\varphi\Big(n \sup_{\substack{E \subset [0,1] \\ |E| = \frac{1}{n}}} \Big\{\int\limits_E \alpha(t)\,dt - \sup_{\substack{E_1 \subset [0,1]-E \\ |E_1| = \frac{1}{n}}} \int\limits_{E_1} \alpha(t)\,dt\Big\}\Big) - K_2\varphi\left(\omega_\varphi\left(\frac{1}{n}, f\right)\right).$$

Hence we obtain that

$$K_3\,\varphi\left(\omega_\varphi\left(\frac{1}{n},f\right)\right) \geqq \frac{1}{n}\,\varphi\Big(n \sup_{\substack{E\subset[0,1]\\ |E|=\frac{1}{n}}}\Big\{\int_E \alpha(t)\,dt - \sup_{\substack{E_1\subset[0,1]-E\\ |E_1|=\frac{1}{n}}} \int_{E_1} \alpha(t)\,dt\Big\},$$

which implies, by $\bar\varphi(ca)\leqq c\bar\varphi(a)$ $(a\geqq 0,\ c\geqq 1)$, that

$$(1.6)\qquad K_4\,\bar\varphi\left(n\varphi\left(\omega_\varphi\left(\frac{1}{n},f\right)\right)\right) \geqq n \sup_{\substack{E\subset[0,1]\\ |E|=\frac{1}{n}}}\Big\{\int_E \alpha(t)\,dt - \sup_{\substack{E_1\subset[0,1]-E\\ |E_1|=\frac{1}{n}}} \int_{E_1} \alpha(t)\,dt\Big\}.$$

An easy computation gives that

$$\sup_{\substack{E\subset[0,1]\\ |E|=\frac{1}{n}}}\Big\{\int_E \alpha(t)\,dt - \sup_{\substack{E_1\subset[0,1]-E\\ |E_1|=\frac{1}{n}}} \int_{E_1} \alpha(t)\,dt\Big\} \geqq$$

$$\geqq \sup_{\substack{E\subset[0,1]\\ |E|=\frac{1}{n}}}\Big\{\int_E |f(t)|\,dt - \sup_{\substack{E_1\subset[0,1]-E\\ |E_1|=\frac{1}{n}}} \int_{E_1} |f(t)|\,dt\Big\} - 2\sup_{\substack{A\subset[0,1]\\ |A|=\frac{1}{n}}} \int_A \big||f(t)|-\alpha(t)\big|\,dt.$$

This and (1. 6) imply

$$\varphi\Big(n \sup_{\substack{E\subset[0,1]\\ |E|=\frac{1}{n}}}\Big\{\int_E |f(t)|\,dt - \sup_{\substack{E_1\subset[0,1]-E\\ |E_1|=\frac{1}{n}}} \int_{E_1} |f(t)|\,dt\Big\}\Big) \leqq$$

$$\leqq \varphi\left(2n \sup_{\substack{A\subset[0,1]\\ |A|=\frac{1}{n}}} \int_A \big||f(t)|-\alpha(t)\big|\,dt + K_4\,\bar\varphi\left(n\varphi\left(\varphi_\varphi\left(\frac{1}{n},f\right)\right)\right)\right).$$

Consequently, by Lemma 3 and Lemma 5, we have

$$\varphi\Big(n \sup_{\substack{E\subset[0,1]\\ |E|=\frac{1}{n}}}\Big\{\int_E |f(t)|\,dt - \sup_{\substack{E_1\subset[0,1]-E\\ |E_1|=\frac{1}{n}}} \int_{E_1} |f(t)|\,dt\Big\}\Big) \leqq$$

$$\leqq K_5\left\{n \sup_{\substack{A\subset[0,1]\\ |A|=\frac{1}{n}}} \int_A \big(\big||f(t)|-\alpha(t)\big|\big)\,dt + n\varphi\left(\omega_\varphi\left(\frac{1}{n},f\right)\right)\right\} \leqq$$

$$\leqq K_6\, n\varphi\left(\omega_\varphi\left(\frac{1}{n},f\right)\right).$$

Hence, by Lemma 1, the statement of Lemma 6 immediately follows.

Lemma 7. *If $f(x)\in\varphi(L)$ then*

$$F(2^{-n}) \leqq K(\varphi)\left\{\int_0^1 \varphi(|f(x)|)\,dx+\sum_{k=1}^{n-1}\bar\varphi(2^k\varphi(\omega_\varphi(2^{-k},f)))\right\} \tag{1.7}$$

for any $n\geqq 1$, where $F(z)$ has the same meaning as in Lemma 1.

Proof. By Lemma 6 we have

$$K_1\bar\varphi(2^k\varphi(\omega_\varphi(2^{-k},f)))\geqq 2^k\left(\int_0^{2^{-k}}F(z)\,dz-\int_{2^{-k}}^{2^{-k+1}}F(z)\,dz\right)$$

and

$$K_1\bar\varphi(2^{k+1}\varphi(\omega_\varphi(2^{-k-1},f)))\geqq 2^{k+1}\left(\int_0^{2^{-k-1}}F(z)\,dz-\int_{2^{-k-1}}^{2^{-k}}F(z)\,dz\right),$$

whence

$$K_2\bar\varphi(2^k\varphi(\omega_\varphi(2^{-k},f)))\geqq 2^k\left(2\int_0^{2^{-k-1}}F(z)\,dz-\int_{2^{-k}}^{2^{-k+1}}F(z)\,dz\right)\geqq$$

$$\geqq F(2^{-k-1})-F(2^{-k}).$$

Summing up this inequality from 1 to $n-1$ we obtain (1.7).

Lemma 8. *If $f(x)\in\varphi_p(L)$ and $R=2^{2p+1}$, then*

$$\int_0^{\frac{1}{Rn}}\varphi(F(x))\,dx\leqq\int_{\frac{1}{Rn}}^{\frac{1}{n}}\varphi(F(x))\,dx+K(\varphi)\varphi\left[\omega_\varphi\left(\frac{1}{n},f\right)\right]$$

for any $n\geqq 1$.

Proof. Set

$$E_n^*=\left\{x:\ x\in[0,1],\ |f(x)|>F\left(\frac{1}{Rn}\right)\right\}$$

and

$$E_n^{**}=\left\{x:\ x\in[0,1],\ |f(x)|=F\left(\frac{1}{Rn}\right)\right\}.$$

If $|E_n^*|<\frac{1}{Rn}$, then define y so that

$$|E_n^{**}\cap(0,y)|=\frac{1}{Rn}-|E^*|,$$

furthermore let

$$E_n=(E_n^{**}\cap(0,y))\cup E_n^*$$

If $|E_n^*|=\frac{1}{Rn}$, then set $E_n=E_n^*$.

First we estimate the integral of $\varphi(|f(x)|)$ on E_n. Let $\psi_n(x)$ be the same function as in Lemma 5. By (1) we have

$$\varphi(|f(x)|) \leqq 2^p\{\varphi(|f(x)-\psi_n(x)|)+\varphi(|\psi_n(x)|)\}.$$

Hence

$$\int_{E_n} \varphi(|f(x)|)\,dx \leqq 2^p\Big\{\int_{E_n}\varphi(|f(x)-\psi_n(x)|)\,dx+\int_{E_n}\varphi(|\psi_n(x)|)\,dx\Big\}. \tag{1.8}$$

Since $\varphi(x)$ is increasing on $[0,\infty)$, in view of Lemma 1 we also have

$$\int_{E_n}\varphi(|f(x)|)\,dx=\int_0^{\frac{1}{Rn}}\varphi(F(x))\,dx$$

(see also [6], p. 29), and by Lemma 5

$$\int_0^1 \varphi(|f(x)-\psi_n(x)|)\,dx \leqq K(\varphi)\,\varphi\left(\omega_\varphi\left(\frac{1}{n},f\right)\right);$$

consequently (see (1.8))

$$\int_0^{\frac{1}{Rn}}\varphi(F(x))\,dx = K_1\,\varphi\left(\omega_\varphi\left(\frac{1}{n},f\right)\right)+\frac{2^p}{Rn}\varphi(\max|\psi_n(x)|). \tag{1.9}$$

But

$$\max|\psi_n(x)| \leqq n\int_0^{\frac{1}{n}}F(x)\,dx,$$

thus by Lemma 3

$$\varphi(\max|\psi_n(x)|) \leqq 2^p n\int_0^{\frac{1}{n}}\varphi(F(x))\,dx.$$

Therefore, by (1.9) and by the definition of R, we have

$$\int_0^{\frac{1}{Rn}}\varphi(F(x))\,dx \leqq K_1\varphi\left(\omega_\varphi\left(\frac{1}{n},f\right)\right)+\frac{2^{2p}}{R}\int_0^{\frac{1}{n}}\varphi(F(x))\,dx \leqq$$

$$\leqq K_1\varphi\left(\omega_\varphi\left(\frac{1}{n},f\right)\right)+\tfrac{1}{2}\int_0^{\frac{1}{Rn}}\varphi(F(x))\,dx+\tfrac{1}{2}\int_{\frac{1}{Rn}}^{\frac{1}{n}}\varphi(F(x))\,dx,$$

whence the statement of Lemma 8 follows.

Lemma 9. *If* $f(x)\in\varphi(L)$ $(\varphi(x)\equiv\varphi_p(x))$ *and* $\varepsilon=(4[p+1]+2)^{-1}$, *then*

$$\int_0^{\frac{1}{n}}\varphi(F(x))\,dx\leqq\frac{K(\varphi)}{n^{\varepsilon}}\left\{\sum_{k=1}^{n}k^{\varepsilon-1}\varphi\left(\omega_\varphi\left(\frac{1}{k},f\right)\right)+\int_0^1\varphi(F(x))\,dx\right\}$$

for any $n\geqq1$.

Proof. Set $\mu=2[p+1]+1$. Since $2^{\mu}>R=2^{2p+1}$, by Lemma 8 we have

$$(1.10)\qquad \int_0^{2^{-n}}\varphi(F(x))\,dx=\Big(\int_0^{2^{-\mu-n}}+\int_{2^{-\mu-n}}^{2^{-n}}\Big)\varphi(F(x))\,dx\leqq$$

$$\leqq 2\int_{2^{-\mu-n}}^{2^{-n}}\varphi(F(x))\,dx+K\varphi(\omega_\varphi(2^{-n},f)).$$

Furthermore it is clear that

$$(1.11)\qquad \sum_{k=n}^{\infty}\int_{2^{-k-\mu}}^{2^{-n}}\varphi(F(x))\,dx\leqq\mu\sum_{k=n}^{\infty}\int_{2^{-k-1}}^{2^{-k}}\varphi(F(x))\,dx=\mu\int_0^{2^{-n}}\varphi(F(x))\,dx.$$

Thus, by (1. 10) and (1. 11), we get

$$(1.12)\qquad \sum_{k=n}^{\infty}\int_{2^{-k-\mu}}^{2^{-k}}\varphi(F(x))\,dx\leqq2\mu\Big\{\int_{2^{-\mu-n}}^{2^{-n}}\varphi(F(x))\,dx+K_1\varphi(\omega_\varphi(2^{-n},f))\Big\}.$$

Let $N=2\mu$, and, furthermore, let us define for every $n\geqq1$

$$a_n=\int_{2^{-\mu-n}}^{2^{-n}}\varphi(F(x))\,dx,\qquad b_n=K_1\varphi(\omega_\varphi(2^{-n},f))$$

and

$$\alpha_n=\sum_{k=(n-1)N+1}^{nN}a_k,\qquad \beta_n=\sum_{k=(n-1)N+1}^{nN}b_k.$$

Considering (1. 12) we have

$$(1.13)\qquad \sum_{k=n}^{\infty}a_k\leqq N(a_n+b_n).$$

Since $a_n\geqq0$, by (1. 13) we obtain that, for any nonnegative integers m and j,

$$\sum_{i=m+1}^{\infty}\alpha_i=\sum_{k=mN+1}^{\infty}a_k\leqq\sum_{k=mN+1-j}^{\infty}a_k=N(a_{mN+1-j}+b_{mN+1-j}).$$

Hence, taking $j=1,2,\dots,N$ and summing up, we get

$$N\sum_{i=m+1}^{\infty}\alpha_i\leqq N\sum_{j=1}^{N}(a_{mN+1-j}+b_{mN+1-j})=N\sum_{k=(m-1)N+1}^{mN}(a_n+b_k)=N(\alpha_m+\beta_m).$$

Consequently

$$\sum_{i=m+1}^{\infty} \alpha_i \leqq \alpha_m + \beta_m.$$

Multiplying this inequality by $\max(2^{m-2}, 1)$ for all m, $1 \leqq m \leqq n$, by summing and simplifying we obtain

$$2^{n-1} \sum_{i=n+1}^{\infty} \alpha_i \leqq \alpha_1 + \beta_1 + \sum_{k=2}^{n} 2^{k-2} \beta_k. \tag{1.14}$$

Inserting the definitions of α_i and β_i, from (1. 14) it follows that

$$2^{n-1} \sum_{k=nN+1}^{\infty} a_k \leqq \sum_{k=1}^{N} (a_k + b_k) + \sum_{k=2}^{n} 2^{k-2} \sum_{i=(k-1)N+1}^{kN} b_i,$$

that is

$$\sum_{k=nN+1}^{\infty} \int_{2^{-\mu-k}}^{2^{-k}} \varphi(F(x))\,dx \leqq \tag{1.15}$$

$$\leqq 2^{-n} \left\{ 2 \sum_{k=1}^{N} \int_{2^{-\mu-k}}^{2^{-k}} \varphi(F(x))\,dx + 2K_1 \sum_{k=1}^{N} \varphi(\omega_\varphi(2^{-k}, f)) + \sum_{k=2}^{n} 2^k N \varphi(\omega_\varphi(2^{-(k-1)N}, f)) \right\}.$$

By (1. 15) it is clear that

$$\int_0^{2^{-(nN+1)}} \varphi(F(x))\,dx \leqq 2^{-n} \left\{ 2N \int_0^1 \varphi(F(x))\,dx + 2NK_1 \varphi(\omega_\varphi(\tfrac{1}{2}, f)) + \right. \tag{1.16}$$

$$\left. + N \sum_{k=2}^{n} 2^k \varphi(\omega_\varphi(2^{-(k-1)N}, f)) \right\} \leqq$$

$$\leqq K_1(\varphi) 2^{-n} \left\{ \int_0^1 \varphi(F(x))\,dx + \sum_{k=1}^{n-1} 2^k \varphi(\omega_\varphi(2^{-kN}, f)) \right\}.$$

If $2^{nN+1} \leqq m < 2^{(n+1)N+1}$, then from (1. 16) it follows with $\varepsilon = (4[p+1]+2)^{-1} = \dfrac{1}{N}$ that

$$\int_0^{\frac{1}{m}} \varphi(F(x))\,dx \leqq \frac{K_2(\varphi)}{m^\varepsilon} \left\{ \int_0^1 \varphi(F(x))\,dx + \sum_{k=1}^{\varepsilon \overset{2}{\smile}{\log} m} 2^k \varphi(\omega_\varphi(2^{-kN}, f)) \right\}. \tag{1.17}$$

The estimation of the sum of the right-hand side is very easy. Indeed, we have

$$\sum_{k=1}^{\varepsilon \overset{2}{\smile}{\log} m} 2^k \varphi(\omega_\varphi(2^{-kN}, f)) \leqq 2^N \sum_{k=1}^{\varepsilon \overset{2}{\smile}{\log} m} \frac{2^k}{2^{kN}} \sum_{i=2^{(k-1)N}+1}^{2^{kN}} \varphi\left(\omega_\varphi\left(\frac{1}{i}, f\right)\right) \leqq$$

$$\leqq 2^{N+1} \sum_{k=1}^{\varepsilon \overset{2}{\smile}{\log} m} \sum_{i=2^{(k-1)N}+1}^{2^{kn}} i^{\varepsilon-1} \varphi\left(\omega_\varphi\left(\frac{1}{i}, f\right)\right) \leqq 2^{N+1} \sum_{i=1}^{m} i^{\varepsilon-1} \varphi\left(\omega_\varphi\left(\frac{1}{i}, f\right)\right);$$

inserting this into (1. 17), we obtain the statement of Lemma 9.

§ 2. Proof of the theorems

Proof of Theorem 1. Let $F(x)$ be the same function as in Lemma 1. Since for any nonnegative nondecreasing function $\chi(u)$ on $[0, \infty)$

$$\int_0^1 \chi(|f(x)|)\,dx = \int_0^1 \chi(F(x))\,dx$$

(see [6], p. 29), it is sufficient to prove (5) for $F(x)$.

Set

$$E_n = \{x : x \in [0,1], \quad n \leqq F(x) < n+1\}, \qquad (n = 0, 1, \ldots).$$

It is clear that $E_n E_m = 0$ if $n \neq m$, $\sum_{n=0}^{\infty} E_n = [0, 1]$, and if $E_n = \{\alpha_{n+1}, \alpha_n\}$ then $\alpha_n \downarrow 0$ as $n \to \infty$. Put

$$A_n = \int_0^{\alpha_n} \varphi(F(x))\,dx.$$

Then $A_n \to 0$ as $n \to \infty$, and thus for $n \geqq n_0$

$$\varphi(1) \geqq A_n \geqq \alpha_n \varphi(n) \geqq \alpha_n n \varphi(1),$$

that is $\alpha_n \leqq \frac{1}{n}$ for $n \geqq n_0$, and therefore

$$A_n \leqq \int_0^{\frac{1}{n}} \varphi(F(x))\,dx \qquad (n \geqq n_0). \tag{2.1}$$

To prove (5) we split the integral into an infinite sum and then make an Abel-transformation; thus we obtain

$$I \equiv \int_0^1 \varphi(F(x)) \Lambda(F(x))\,dx = \sum_{n=0}^{\infty} \int_{E_n} \varphi(F(x))\,dx \sum_{k=1}^{n} \frac{\lambda_k}{k} \leqq$$

$$\leqq \sum_{k=1}^{\infty} \frac{\lambda_k}{k} \sum_{n=k}^{\infty} \int_{E_n} \varphi(F(x))\,dx = \sum_{k=1}^{\infty} \frac{\lambda_k}{k} A_k.$$

From this, (4), and (2.1), by using Lemma 9, we get that

$$I \leqq K(\varphi) \sum_{k=1}^{\infty} \frac{\lambda_k}{k^{1+\varepsilon}} \left\{ \sum_{n=1}^{k} n^{\varepsilon-1} \varphi\left[\omega_\varphi\left(\frac{1}{n}, f\right)\right] + \int_0^1 \varphi(F(x))\,dx \right\} \leqq$$

$$\leqq K(\varphi) \sum_{k=1}^{\infty} \frac{\lambda_k}{k^{1+\varepsilon}} \int_0^1 \varphi(F(x))\,dx + K(\varphi) \sum_{n=1}^{\infty} n^{\varepsilon-1} \varphi\left[\omega_\varphi\left(\frac{1}{n}, f\right)\right] \sum_{k=n}^{\infty} \frac{\lambda_k}{k^{1+\varepsilon}} \leqq$$

$$\leqq K(\varphi, \lambda) \int_0^1 \varphi(F(x))\,dx + K(\varphi, \lambda) \sum_{n=1}^{\infty} \frac{\lambda_n}{n} \varphi\left[\omega_\varphi\left(\frac{1}{n}, f\right)\right],$$

whence (5) obviously follows.

The proof is thus complete.

Proof of Theorem 2. To prove (7) we use Lemma 4 and 7. It is clear that

$$I \equiv \int_0^1 \psi(|f(x)|)\varrho(|f(x)|)\,dx = \int_0^1 \psi(F(x))\varrho(F(x))\,dx = \tag{2.2}$$

$$= \sum_{n=0}^{\infty} \int_{2^{-n-1}}^{2^{-n}} \psi(F(x))\varrho(F(x))\,dx \leqq \sum_{n=0}^{\infty} 2^{-n-1}\psi(F(2^{-n-1}))\varrho(F(2^{-n-1})).$$

Since, by (1), $\bar{\varphi}(x)x^{-1}$ decreases ($\bar{\varphi}(x)$ denotes the inverse of $\varphi(x)$), for any $c \geqq 1$ and $a \geqq 0$ we have

$$\bar{\varphi}(ca) \leqq c\bar{\varphi}(a).$$

This and (1. 7) imply

$$F(2^{-n}) \leqq K_1 2^n.$$

Using this inequality, (6) and (1. 7), by (2. 2) we have

$$I \leqq K_2 \sum_{n=0}^{\infty} 2^{-n}\varrho(2^n)\psi\Big(\int_0^1 \varphi(|f(x)|)\,dx + \sum_{k=1}^{n} \bar{\varphi}(2^k\varphi(\omega_\varphi(2^{-k},f))) \leqq$$

$$\leqq K_3 \sum_{n=1}^{\infty} 2^{-n}\varrho(2^n)\psi\Big(\int_0^1 \varphi(|f(x)|)\,dx\Big) +$$

$$+ K_3 \sum_{n=0}^{\infty} 2^{-n}\varrho(2^n)\psi\left(\sum_{k=1}^{n} \bar{\varphi}(2^k\varphi(\omega_\varphi(2^{-k},f)))\right) \leqq K_4\psi\Big(\int_0^1 \varphi(|f(x)|)\,dx +$$

$$+ K_5 \sum_{n=0}^{\infty} 2^{-n}\varrho(2^n)\psi\left(\sum_{k=1}^{n} \sum_{i=2^{k-1}+1}^{2^k} \frac{1}{i}\bar{\varphi}\left[i\varphi\left(\omega_\varphi\left(\frac{1}{i},f\right)\right)\right]\right) \leqq \tag{2.3}$$

$$\leqq K_4\psi\Big(\int_0^1 \varphi(|f(x)|)\Big)\,dx +$$

$$+ K_6 \sum_{m=1}^{\infty} \frac{\varrho(m)}{m^2}\psi\left(\sum_{i=1}^{m} \frac{1}{i}\bar{\varphi}\left[i\varphi\left(\omega_\varphi\left(\frac{1}{i},f\right)\right)\right]\right).$$

Applying Lemma 4, by (6), we obtain

$$\sum_{m=1}^{\infty} \frac{\varrho_m}{m^2}\psi\left(\sum_{i=1}^{m} \frac{1}{i}\bar{\varphi}\left[i\varphi\left(\omega_\varphi\left(\frac{1}{i},f\right)\right)\right]\right) \leqq K_6 \sum_{m=1}^{\infty} \frac{\varrho_m}{m^2}\psi\left(\bar{\varphi}\left[m\varphi\left(\omega_\varphi\left(\frac{1}{m},f\right)\right)\right]\right).$$

This and (2. 3) imply (7), and this completes the proof.

References

[1] L. LEINDLER, On embedding of classes H_p^ω, *Acta Sci. Math.*, **31** (1970), 13—31.
[2] H. P. MULHOLLAND, The generalization of certain inequality theorems involving powers, *Proc. London Math. Soc.*, **33** (1932), 481—516.
[3] J. NÉMETH, Generalizations of the Hardy—Littlewood inequality. II, *Acta Sci. Math.*, **35** (1973) (to appear).
[4] П. Л. Ульянов, Вложение некоторых классов функций H_p^ω, *Изв. АН СССР*, сер. матем., **32** (1968), 649—686.
[5] ———— Теоремы вложения и соотношения между наилучшими приближениями (модулями непрерывности) в разных метриках, *Мат. Сбор.*, **81** (1970), 104—131.
[6] A. ZYGMUND, *Trigonometric Series*, Vol. I (Cambridge, 1959).

(Received February 3, 1972)

О некоторых обобщениях теории сильно демпфированных пучков на случай пучков произвольного порядка

А. С. МАРКУС, В. И. МАЦАЕВ и Г. И. РУССУ (Кишинев и Москва, СССР)

Посвящается академику Б. С.-Надь к его шестидесятилетию

Введение

Решение методом Фурье различных задач математической физики естественно приводит к изучению разложений по собственным векторам, отвечающим части спектра самосопряженного операторного пучка, т. е. полинома

$$L(\lambda) = \sum_{k=0}^{n} \lambda^k A_k \qquad (A_k^* = A_k).$$

Общей спектральной теории такого рода пока не существует.

Необходимо отметить, что спектральная теория одного важного класса несамосопряженных пучков успешно развивается уже на протяжении более 20 лет, начиная с основополагающей работы М. В. Келдыша [1]. В то же время, построение общей теории самосопряженных пучков началось, по существу, только в 1964 году, когда появилась фундаментальная работа М. Г. Крейна и Г. Лангера [2], посвященная систематическому изучению самосопряженных квадратичных пучков. Одна из основных идей этой работы заключается в сопоставлении пучку $\lambda^2 I + \lambda B + C$ операторного квадратного уравнения $Z^2 + BZ + C = 0$ и в отыскании корня этого уравнения, спектр которого совпадает с некоторой частью спектра пучка.

Результаты работы [2] по спектральной теории самосопряженных квадратичных пучков получили дальнейшее развитие в работах Г. Лангера [3, 4]. В недавней работе Г. Лангера [5]*) были получены некоторые общие результаты по спектральной теории самосопряженных пучков произвольного порядка. Постановка задач в [5] близка к нашей.

---

*) Авторы благодарны Г. Лангеру за предоставленную им возможность ознакомиться с рукописью этой работы. Это позволило упростить некоторые доказательства в нашей статье.

Настоящая статья посвящена, в основном, обобщениям на пучки n-го порядка теории одного из классов квадратичных пучков, изученных в [2, 3], — сильно демпфированных пучков*). В отличие от работ [2—5], основной подход статьи заключается не в изучении корней операторного уравнения n-го порядка, а в факторизации пучка, т. е. представлении его в виде $L(\lambda) = L_+(\lambda)(Z-\lambda I)$, где спектр оператора Z совпадает с некоторой частью σ спектра $L(\lambda)$, а $L_+(\lambda)$ — пучок $n-1$-го порядка, обратимый на σ. Эти два подхода тесно связаны между собой (см. ниже замечание 1).

В работах [2—5] систематически используются методы и результаты теории операторов в пространстве с индефинитной метрикой, в особенности теоремы о существовании специальных инвариантных подпространств. В нашей статье используются не геометрические, а аналитические методы. Основным инструментом являются недавние результаты И. Ц. Гохберга и Ю. Лайтерера о факторизации оператор-функций.

Объект исследования этой статьи — *гиперболический пучок*. Так мы называем самосопряженный пучок $L(\lambda)$, квадратичная форма которого имеет при любом $f\neq 0$ простые вещественные корни, а старший коэффициент A_n — равномерно положительный оператор.

Оказывается, что, как и в случае $n=2$, корни квадратичной формы $(L(\lambda)f, f)$ образуют неперекрывающиеся промежутки — спектральные зоны. Установлению этого факта и некоторых других утверждений о спектральных зонах посвящен первый параграф.

Основным результатом статьи является доказанная в § 2 теорема 4, в которой утверждается, что всякой спектральной зоне гиперболического пучка, замыкание которой не пересекается с замыканиями других зон, отвечает факторизация пучка $L(\lambda)$, причем соответствующий оператор Z подобен самосопряженному. Отметим, что при $n=2$ эта теорема вытекает из теоремы Г. Лангера [3].

В этом же параграфе строится пример, показывающий существенность условий теоремы 4 для подобия Z самосопряженному оператору. Здесь же рассматриваются обобщения теоремы 4 на случай, когда A_n не является дефинитным.

В последнем параграфе указываются некоторые ограничения на коэффициенты пучка, обеспечивающие выполнение условий теорем § 2. В заключение рассматривается один класс пучков, не являющихся гиперболическими, а именно, самосопряженные пучки, полученные малым возмущением линейного пучка $A-\lambda I$.

Авторы выражают благодарность И. Ц. Гохбергу и М. Г. Крейну за ценные обсуждения.

---

*) В матричном случае теория сильно демпфированных пучков была построена Даффином [6].

§ 1. Свойства спектральных зон

1. Пусть $\mathfrak{H}$ — гильбертово пространство, $\mathfrak{R}$ — множество всех линейных ограниченных операторов, действующих в $\mathfrak{H}$, и $\mathfrak{S}_\infty$ — множество всех вполне непрерывных операторов из $\mathfrak{R}$. Если $A \in \mathfrak{R}$, то через $\operatorname{im} A$ обозначается множество значений оператора A, а через $\ker A$ — его ядро. Как обычно, $\sigma(A)$ обозначает спектр оператора A, а $W(A)$ — его числовой образ, т. е. $W(A) = \{(Af,f): \|f\| = 1\}$. Неравенство $A \gg 0$ будет обозначать, что оператор A равномерно положителен, т. е. существует число $\delta > 0$ такое, что $(Af,f) \geqq \delta(f,f)\,(f \in \mathfrak{H})$.

Полиномиальным операторным пучком называют операторный полином

$$L(\lambda) = \lambda^n A_n + \lambda^{n-1} A_{n-1} + \cdots + \lambda A_1 + A_0 \tag{1. 1}$$

с коэффициентами из $\mathfrak{R}$. Если $A_k^* = A_k$ $(k=0, 1, \ldots, n)$, то пучок $L(\lambda)$ называется *самосопряженным*.

Через $\sigma(L)$ обозначим *спектр пучка* $L(\lambda)$, т. е. множество всех комплексных чисел λ, для которых оператор $L(\lambda)$ не обратим. Если уравнение $L(\lambda_0)\varphi = 0$ имеет решение $\varphi_0 \neq 0$, то λ_0 будем называть *собственным числом* пучка $L(\lambda)$, а вектор φ_0 — соответствующим ему *собственным вектором*. Вектор φ_k называется *присоединенным вектором* пучка $L(\lambda)$, соответствующим собственному числу λ_0, если существуют векторы $\varphi_0(\neq 0)$, $\varphi_1, \ldots, \varphi_{k-1}$ такие, что

$$\sum_{j=0}^{m} \frac{1}{j!} L^{(j)}(\lambda_0)\varphi_{m-j} = 0 \qquad (m = 0, 1, \ldots, k).$$

2. Самосопряженный пучок (1. 1) будем называть *гиперболическим*, если $A_n \gg 0$ и при любом $f \neq 0$ все корни многочлена $(L(\lambda)f, f)$ вещественны и различны.

Корни многочлена $(L(\lambda)f,f)$, занумерованные в порядке убывания, обозначим $p_k(f)$ $(k=1, 2, \ldots, n)$. Так как $p_k(\alpha f) = p_k(f)$ $(f \neq 0, \alpha \neq 0)$, то можно рассматривать функционалы $p_k(f)$ на единичной сфере S пространства $\mathfrak{H}$. Очевидно, $p_k(f)$ является ограниченным непрерывным функционалом на S, и поэтому множество его значений Δ_k является непустым связным ограниченным подмножеством вещественной прямой, т. е. некоторым промежутком (или точкой). Этот промежуток Δ_k назовем k-ой *спектральной зоной* пучка $L(\lambda)$.

Установим вначале некоторые простые свойства гиперболических пучков.

1°. *Спектр гиперболического пучка* $L(\lambda)$ *вещественный, причем* $\sigma(L) \subset \subset \bigcup_{j=1}^{n} \bar{\Delta}_j$.

В самом деле, при $\|f\|=1$

(1.2)
$$\|L(\lambda)f\| \geqq |(L(\lambda)f,f)| = (A_n f,f)\prod_{j=1}^{n}|\lambda-p_j(f)| \geqq \|A_n^{-1}\|^{-1}\prod_{j=1}^{n}\varrho(\lambda,\Delta_j),$$

где $\varrho(\lambda, \Delta_j)$ — расстояние от λ до Δ_j.

Подставляя в неравенство (1.2) $\bar\lambda$ вместо λ, получим

$$\|[L(\lambda)]^* f\| \geqq \|A_n^{-1}\|^{-1}\prod_{j=1}^{n}\varrho(\lambda,\Delta_j).$$

Отсюда и из (1.2) непосредственно вытекает, что если $\lambda\notin\bigcup_{j=1}^{n}\bar\Delta_j$, то оператор $L(\lambda)$ обратим.

2°. *Если $L(\lambda)$ — гиперболический пучок, то у него нет присоединенных векторов.*

В самом деле, допустим что существуют число λ_0 (вещественное в силу 1°) и векторы $\varphi_0(\neq 0)$ и φ_1 такие, что

$$L(\lambda_0)\varphi_0=0,\quad L(\lambda_0)\varphi_1=-L'(\lambda_0)\varphi_0.$$

Умножая последнее равенство скалярно на φ_0, получим

$$(L'(\lambda_0)\varphi_0,\varphi_0)=-(L(\lambda_0)\varphi_1,\varphi_0)=-(\varphi_1,L(\lambda_0)\varphi_0)=0.$$

Так как $(L(\lambda_0)\varphi_0,\varphi_0)=0$, то отсюда следует, что λ_0 является кратным корнем многочлена $(L(\lambda)\varphi_0,\varphi_0)$, а это противоречит условию гиперболичности.

3. Здесь мы установим основное свойство спектральных зон гиперболического пучка, состоящее в том, что они не перекрываются. Для квадратичного пучка это свойство было установлено в случае $\dim\mathfrak{H}<\infty$ Даффином [6] и в общем случае М. Г. Крейном и Г. Лангером [2] (см. также [4, 7]).

Нам понадобится следующее простое предложение ([7], лемма 1. 1).

Лемма 1. *Пусть A и B — ограниченные самосопряженные операторы. Если для некоторых векторов f_1 и f_2*

$$(Af_1,f_1)=(Af_2,f_2)=0,\quad (Bf_1,f_1)>0,\quad (Bf_2,f_2)<0,$$

*то найдется вектор $f\neq 0$ такой, что $(Af,f)=(Bf,f)=0$**).

---

*) Как заметил Б. С. Митягин, эта лемма вытекает из теоремы Теплица—Хаусдорфа о выпуклости числового образа оператора. В самом деле, полагая $C=B+iA$ и $\hat f_k=f_k/\|f_k\|$ $(k=1,2)$, получаем, что $(C\hat f_1,\hat f_1)>0$, $(C\hat f_2,\hat f_2)<0$, и поэтому существует вектор f $(\|f\|=1)$, такой что $(Cf,f)=0$. Обратно, из леммы 1 с помощью аффинного преобразования без труда выводится теорема Теплица—Хаусдорфа.

Теорема 1. *Если $L(\lambda)$ — гиперболический пучок, то его различные спектральные зоны не пересекаются.*

Доказательство. Так как все корни многочлена $(L(\lambda)f, f)$ $(f\neq 0)$ вещественны и различны, то его производная в соседних корнях имеет противоположные знаки, и, следовательно,

$$(1.3)\qquad (-1)^{k-1}(L'(p_k(f))f, f) > 0 \qquad (k = 1, 2, \ldots, n; \quad f \neq 0).$$

Допустим, что утверждение теоремы неверно. Тогда $\Delta_k \cap \Delta_{k+1}$ непусто для некоторого k, т. е. найдутся вещественное число α и ненулевые векторы φ, ψ такие, что

$$(1.4)\qquad (L(\alpha)\varphi, \varphi) = (L(\alpha)\psi, \psi) = 0,$$

причем $\alpha = p_k(\varphi) = p_{k+1}(\psi)$. В силу (1. 3)

$$(1.5)\qquad (L'(\alpha)\varphi, \varphi)(L'(\alpha)\psi, \psi) < 0.$$

Из (1. 4) и (1. 5) согласно лемме 1 вытекает, что $(L(\alpha)g, g) = (L'(\alpha)g, g) = 0$ для некоторого $g\neq 0$, что противоречит простоте корней многочлена $(L(\lambda)g, g)$. Теорема доказана*).

4. Для получения основных результатов этой статьи свойство неперекрываемости спектральных зон оказывается недостаточным.

Будем говорить, что две спектральные зоны *отделены*, если их замыкания не пересекаются.

Следующая теорема идейно близка к теореме 1. Она показывает, что из существования равномерного зазора между соседними корнями многочлена $(L(\lambda)f, f)$ вытекает, что соответствующие спектральные зоны отделены.

Теорема 2. *Пусть $L(\lambda)$ — гиперболический пучок. Если для некоторого k $(1 \leqq k \leqq n-1)$ найдется положительное число ϱ такое, что $p_k(f) - p_{k+1}(f) \geqq \varrho$ для любого вектора $f\neq 0$, то спектральные зоны Δ_k и Δ_{k+1} отделены.*

Доказательство. Допустим, что утверждение теоремы не имеет места, т. е. что $\sup \Delta_{k+1} = \inf \Delta_k$. Обозначим это число через γ и докажем вначале существование положительных чисел ε, δ и μ таких, что из соотношений

$$|(L(\lambda)f, f)| < \varepsilon, \quad \|f\| = 1, \quad |\lambda - \gamma| < \mu$$

вытекает $|(L'(\lambda)f, f)| > \delta$.

В самом деле, если это не так, то найдутся нормированная последовательность векторов $\{f_j\}_1^\infty$ и сходящаяся к γ последовательность вещественных чисел $\{\gamma_j\}_1^\infty$ такие, что

$$\lim_{j\to\infty}(L(\gamma_j)f_j, f_j) = \lim_{j\to\infty}(L'(\gamma_j)f_j, f_j) = 0.$$

*) Аналогичное рассуждение для квадратичного пучка проведено в [4] (лемма 2.4).

Тогда, очевидно,

$$\lim_{j\to\infty}\big(L(\gamma)f_j,f_j\big)=\lim_{j\to\infty}\big(L'(\gamma)f_j,f_j\big)=0. \tag{1.6}$$

Так как коэффициенты многочленов $Q_j(\lambda)=\big(L(\lambda)f_j,f_j\big)$ ограничены, то (выделяя подпоследовательность и не меняя обозначений) можно считать, что последовательность $Q_j(\lambda)$ сходится (равномерно на любом компакте) к некоторому многочлену $Q(\lambda)$. При этом $Q(\lambda)\not\equiv 0$, так как коэффициент при λ^n положителен.

Из (1.6) следует, что $Q(\gamma)=Q'(\gamma)=0$, и по теореме Гурвица [8] в любой окрестности точки γ многочлен $Q_j(\lambda)$ при достаточно большом j имеет два корня, т. е.

$$\lim_{j\to\infty}[p_k(f_j)-p_{k+1}(f_j)]=0,$$

а это противоречит условию теоремы.

Так как по допущению $\gamma=\sup\Delta_{k+1}=\inf\Delta_k$, то существуют нормированные последовательности векторов $\{h_j\}_1^\infty$, $\{g_j\}_1^\infty$ и последовательности вещественных чисел $\{\alpha_j\}_1^\infty$, $\{\beta_j\}_1^\infty$ такие, что

$$\big(L(\alpha_j)h_j,h_j\big)=\big(L(\beta_j)g_j,g_j\big)=0,$$

$$\alpha_j=p_k(h_j),\ \beta_j=p_{k+1}(g_j)\ \ (j=1,2,\dots);\quad \lim\alpha_j=\lim\beta_j=\gamma. \tag{1.7}$$

Из (1.7) и доказанного выше утверждения следует, что при $j\geqq j_0$

$$\big|\big(L'(\alpha_j)h_j,h_j\big)\big|>\delta,\quad \big|\big(L'(\beta_j)g_j,g_j\big)\big|>\delta.$$

Точнее говоря, так как $\alpha_j=p_k(h_j)$, $\beta_j=p_{k+1}(g_j)$, то в силу (1.3)

$$(-1)^{k-1}\big(L'(\alpha_j)h_j,h_j\big)>\delta,\ (-1)^{k-1}\big(L'(\beta_j)g_j,g_j\big)<-\delta\qquad (j\geqq j_0). \tag{1.8}$$

Из соотношений (1.7) и (1.8) вытекает, что

$$\lim_{j\to\infty}\big(L(\gamma)h_j,h_j\big)=0,\quad (-1)^{k-1}\big(L'(\gamma)h_j,h_j\big)>\frac{\delta}{2}\qquad (j\geqq j_1), \tag{1.9}$$

$$\lim_{j\to\infty}\big(L(\gamma)g_j,g_j\big)=0,\quad (-1)^{k-1}\big(L'(\gamma)g_j,g_j\big)<-\frac{\delta}{2}\qquad (j\geqq j_1). \tag{1.10}$$

Выделяя из ограниченных последовательностей $\big(L'(\gamma)h_j,h_j\big)$ и $\big(L'(\gamma)g_j,g_j\big)$ сходящиеся подпоследовательности, получим

$$(-1)^{k-1}\lim_{m\to\infty}\big(L'(\gamma)h_{j_m},h_{j_m}\big)=t_1\left(\geqq\frac{\delta}{2}\right), \tag{1.11}$$

$$(-1)^{k-1}\lim_{m\to\infty}\big(L'(\gamma)g_{j_m},g_{j_m}\big)=t_2\left(\leqq-\frac{\delta}{2}\right). \tag{1.12}$$

Рассмотрим оператор

$$C = (-1)^{k-1}(L'(\gamma)+iL(\gamma)).$$

Пусть $\overline{W(C)}$ — замыкание числового образа оператора C. Из соотношений (1. 9) и (1. 11) следует, что $t_1 \in \overline{W(C)}$, а из соотношений (1. 10) и (1. 12) — что $t_2 \in \overline{W(C)}$. Так как $\overline{W(C)}$ — выпуклое множество, то $0 \in \overline{W(C)}$. Это означает, что найдется такая нормированная последовательнисть $\{\psi_j\}_1^\infty$, для которой $(C\psi_j, \psi_j) \to 0$, т. е. $(L(\gamma)\psi_j, \psi_j) \to 0$, $(L'(\gamma)\psi_j, \psi_j) \to 0$. Последние соотношения противоречат утверждению, установленному в начале доказательства теоремы. Теорема доказана.

§ 2. Факторизация гиперболического пучка

1. Ниже нам понадобится следующее предложение, вытекающее из результатов И. Ц. Гохберга и Ю. Лайтерера [9].

Теорема 3. *Пусть $A(\zeta)$ — голоморфная на окружности $|\zeta|=1$ оператор-функция со значениями в $\mathfrak{R}$. Если*

(2. 1) $$\operatorname{Re} A(\zeta) \gg 0 \qquad (|\zeta| = 1),$$

то $A(\zeta)$ допускает каноническую факторизацию

$$A(\zeta) = A_+(\zeta) A_-(\zeta),$$

где оператор-функция $A_+(\zeta)$ (соответственно $A_-(\zeta)$) голоморфна и обратима при $|\zeta| \leqq 1$ (соответственно $|\zeta| \geqq 1$), причем $A_-(\infty) = I$.

Следующие два вспомогательных предложения будут применяться для проверки выполнения условия (2. 1) в нашем случае.

Лемма 2. *Пусть заданы вещественные числа $\{c_j\}_1^n$, $\{p_j\}_1^n$ и положительное число r, причем*

$$c_n \leqq p_n \leqq \cdots \leqq c_{k+1} \leqq p_{k+1} < c_k - r < p_k < c_k + r < p_{k-1} \leqq c_{k-1} \leqq \cdots \leqq p_1 \leqq c_1 .$$

Положим

$$a(\lambda) = \frac{\prod_{j=1}^{n} (\lambda - p_j)}{\prod_{j=1}^{n} (\lambda - c_j)} .$$

Тогда для любого комплексного числа λ, лежащего на окружности $|\lambda - c_k| = r$, выполняется неравенство

(2. 2) $$\operatorname{Re} a(\lambda) > 0.$$

Доказательство. Достаточно установить, что *)

$$|\arg a(\lambda)| < \frac{\pi}{2}. \tag{2.3}$$

Положим

$$\theta_1(\lambda) = \sum_{j=k+1}^{n} [\arg(\lambda - p_j) - \arg(\lambda - c_j)], \quad \theta_2(\lambda) = \arg(\lambda - p_k) - \arg(\lambda - c_k),$$

$$\theta_3(\lambda) = \sum_{j=1}^{k-1} [\arg(\lambda - c_j) - \arg(\lambda - p_j)], \quad \theta(\lambda) = \theta_1(\lambda) + \theta_2(\lambda) - \theta_3(\lambda).$$

Очевидно, $\theta(\lambda) \equiv \arg a(\lambda) \pmod{2\pi}$.

Предположим вначале, что $\operatorname{Im} \lambda > 0$. Нетрудно заметить, что тогда

$$\arg(\lambda - p_{k+1}) \geqq \arg(\lambda - c_{k+1}) \geqq \dots \geqq \arg(\lambda - p_n) \geqq \arg(\lambda - c_n),$$

$$\arg(\lambda - c_1) \geqq \arg(\lambda - p_1) \geqq \dots \geqq \arg(\lambda - c_{k-1}) \geqq \arg(\lambda - p_{k-1}),$$

и поэтому

$$0 \leqq \theta_1(\lambda) \leqq \arg(\lambda - p_{k+1}), \quad 0 \leqq \theta_3(\lambda) \leqq \arg(\lambda - c_1) - \arg(\lambda - p_{k-1}).$$

Следовательно,

$$0 \leqq \theta_1(\lambda) < \arg(\lambda - c_k + r) < \frac{\pi}{2}, \tag{2.4}$$

$$0 \leqq \theta_3(\lambda) < \pi - \arg(\lambda - c_k - r) < \frac{\pi}{2}. \tag{2.5}$$

Рассмотрим случай, когда $p_k \geqq c_k$. В этом случае

$$0 \leqq \theta_2(\lambda) < \arg(\lambda - c_k - r) - \arg(\lambda - c_k). \tag{2.6}$$

Так как

$$\frac{(\lambda - c_k + r)(\lambda - c_k - r)}{\lambda - c_k} = \lambda - c_k - \frac{r^2}{\lambda - c_k} = \lambda - c_k - \overline{(\lambda - c_k)} = 2i \operatorname{Im} \lambda,$$

то

$$\arg \frac{(\lambda - c_k + r)(\lambda - c_k - r)}{\lambda - c_k} = \frac{\pi}{2},$$

и, следовательно,

$$\arg(\lambda - c_k + r) + \arg(\lambda - c_k - r) - \arg(\lambda - c_k) \equiv \frac{\pi}{2} \pmod{2\pi}.$$

Учитывая, что

$$0 < \arg(\lambda - c_k + r) < \frac{\pi}{2}, \quad 0 < \arg(\lambda - c_k - r) - \arg(\lambda - c_k) < \arg(\lambda - c_k - r) < \pi,$$

*) Мы выбираем в качестве области значений $\arg z$ промежуток $(-\pi, \pi]$.

получаем

(2. 7) $$\arg(\lambda-c_k+r)+\arg(\lambda-c_k-r)-\arg(\lambda-c_k)=\frac{\pi}{2}.$$

Поэтому из (2. 4) и (2. 6) вытекает, что

(2. 8) $$0\leqq\theta_1(\lambda)+\theta_2(\lambda)<\frac{\pi}{2}.$$

Из (2. 5) и (2. 8) непосредственно следует, что $|\theta(\lambda)|<\frac{\pi}{2}$, но тогда $\arg a(\lambda)=\theta(\lambda)$, и мы получаем неравенство (2. 3).

Рассмотрим теперь случай, когда $p_k<c_k$. Тогда

(2. 9) $$0>\theta_2(\lambda)>\arg(\lambda-c_k+r)-\arg(\lambda-c_k),$$

и из соотношений (2. 5), (2. 7) и (2. 9) вытекает, что

(2. 10)
$$0>\theta_2(\lambda)-\theta_3(\lambda)>\arg(\lambda-c_k+r)+\arg(\lambda-c_k-r)-\arg(\lambda-c_k)-\pi=-\frac{\pi}{2}.$$

Из (2. 4) и (2. 10) следует, что и в этом случае $|\theta(\lambda)|<\frac{\pi}{2}$, откуда снова вытекает (2. 3).

Так как $\arg a(\bar\lambda)=\arg\overline{a(\lambda)}=-\arg a(\lambda)$, то неравенство (2. 3) справедливо и при $\operatorname{Im}\lambda<0$. При $\operatorname{Im}\lambda=0$ неравенство (2. 3) очевидно, так как тогда $(\lambda-p_j)(\lambda-c_j)^{-1}>0$ $(j=1, 2, \ldots, n)$.

Лемма доказана.

Лемма 3. *Пусть заданы вещественные числа* $\{\alpha_j\}_1^n$, $\{\beta_j\}_1^n$, *c и положительное число r, причем*

$$\alpha_n\leqq\beta_n\leqq\cdots\leqq\alpha_{k+1}\leqq\beta_{k+1}<c-r<\alpha_k\leqq\beta_k<c+r<\alpha_{k-1}\leqq\beta_{k-1}\leqq\cdots\leqq\alpha_1\leqq\beta_1.$$

Положим

(2. 11) $$a(\lambda,p)=\frac{(\lambda-p_1)(\lambda-p_2)\cdots(\lambda-p_n)}{(\lambda-\beta_1)\cdots(\lambda-\beta_{k-1})(\lambda-\alpha_{k+1})\cdots(\lambda-\alpha_n)}$$

и через Γ обозначим окружность $|\lambda-c|=r$. Тогда существует такое число $\delta>0$, что $\operatorname{Re} a(\lambda,p)\geqq\delta$ для любых $\lambda\in\Gamma$ и $p_k\in[\alpha_k,\beta_k]$ $(k=1, 2, \ldots, n)$.

Доказательство. В силу леммы 2 $\operatorname{Re} a(\lambda,p)>0$, и остается заметить, что функция $\operatorname{Re} a(\lambda,p)$ непрерывна на $\Gamma\times[\alpha_1,\beta_1]\times\cdots\times[\alpha_n,\beta_n]$.

2. Будем говорить, что оператор $Z\,(\in\mathfrak{R})$ *симметризуется справа (слева)* самосопряженным оператором $S(\in\mathfrak{R})$, если оператор $ZS\,(SZ)$ самосопряженный. Если при этом $S\gg0$, то, очевидно, Z подобен самосопряженному оператору:

$$Z=S^{1/2}(S^{-1/2}ZSS^{-1/2})S^{-1/2}\qquad(Z=S^{-1/2}(S^{-1/2}SZS^{-1/2})S^{1/2}).$$

Основным результатом этой статьи является следующая

Теорема 4. *Пусть*

$$L(\lambda) = \lambda^n A_n + \lambda^{n-1} A_{n-1} + \cdots + \lambda A_1 + A_0$$

—гиперболический пучок). Если при некотором* k $(1 \leqq k \leqq n)$ *спектральная зона* Δ_k *пучка* $L(\lambda)$ *отделена от соседних зон,* то $L(\lambda)$ *допускает следующую факторизацию*

$$L(\lambda) = L_+(\lambda)(Z - \lambda I), \tag{2.12}$$

где $L_+(\lambda) = \sum_{j=0}^{n-1} \lambda^j B_j$ *обратим при всех* $\lambda \in \bar{\Delta}_k$, *спектр оператора* Z *содержится в* $\bar{\Delta}_k$ *и оператор* Z *подобен самосопряженному.*

Доказательство. Пусть $\alpha_j = \inf \Delta_j$, $\beta_j = \sup \Delta_j$ $(j = 1, 2, \ldots, n)$. Так как по условию $\beta_{k+1} < \alpha_k$ и $\beta_k < \alpha_{k-1}$, то можно выбрать вещественное число c и положительное число r так, чтобы $\beta_{k+1} < c - r < \alpha_k$, $\beta_k < c + r < \alpha_{k-1}$. Положим $\Gamma = \{\lambda : |\lambda - c| = r\}$ и

$$A(\lambda) = \frac{L(\lambda)}{(\lambda - \beta_1) \cdots (\lambda - \beta_{k-1})(\lambda - c)(\lambda - \alpha_{k+1}) \cdots (\lambda - \alpha_n)}.$$

Так как $(L(\lambda)f, f) = (A_n f, f) \prod_{j=1}^{n} (\lambda - p_j(f))$, то, обозначая $p_j(f)$ через p_j, получим, что

$$(A(\lambda)f, f) = (A_n f, f)\, a(\lambda, p),$$

где функция $a(\lambda, p)$ определена равенством (2. 11). Поэтому из леммы 3 вытекает, что

$$\operatorname{Re} A(\lambda) \geqq \delta_1 I \qquad (\lambda \in \Gamma), \tag{2.13}$$

где $\delta_1 = \delta \inf_{\|f\| = 1} (A_n f, f)$.

Таким образом, для рациональной оператор-функции $A(\lambda)$ выполнены условия теоремы 3 (относительно окружности Γ), и поэтому

$$A(\lambda) = A_+(\lambda) A_-(\lambda), \tag{2.14}$$

где множители $A_\pm(\lambda)$ обладают свойствами, указанными в теореме 3. Перепишем равенство (2. 14) в виде

$$A_+^{-1}(\lambda) A(\lambda) = A_-(\lambda) \qquad (\lambda \in \Gamma) \tag{2.15}$$

---

*) Используя последние результаты Г. Лангера ([16], теорема 1), нетрудно убедиться, что теорема 4 (без изменений в доказательстве) сохраняет силу если из определения гиперболичности исключить требование простоты корней $(L(\lambda)f, f)$. Аналогичные замечания имеют место для нижеследующих теорем 5 и 6.

Правая часть этого равенства голоморфна при $|\lambda-c| \geqq r$, а левая часть голоморфна при $|\lambda-c| \leqq r$, за исключением точки c, где она имеет простой полюс. Поэтому обе части равенства (2. 15) представляют единую голоморфную оператор-функцию, единственной особенностью которой во всей расширенной плоскости является простой полюс в точке c. Так как $A_-(\infty)=I$, то отсюда следует, что

$$A_-(\lambda) = I+\frac{X}{\lambda-c} \qquad (X\in\mathfrak{R}). \tag{2.16}$$

Переписывая (2. 14) в виде

$$A(\lambda)A_-^{-1}(\lambda) = A_+(\lambda).$$

видим, что обе части последнего равенства определяют единую голоморфную оператор-функцию, имеющую в расширенной плоскости особенности лишь в точках $\beta_1, \ldots, \beta_{k-1}, \alpha_{k+1}, \ldots, \alpha_n$, причем все эти точки являются простыми полюсами. Следовательно,

$$A_+(\lambda) = \Pi^{-1}(\lambda) \sum_{j=0}^{n-1} \lambda^j C_j, \tag{2.17}$$

где $\Pi(\lambda) = (\lambda-\beta_1)\cdots(\lambda-\beta_{k-1})(\lambda-\alpha_{k+1})\cdots(\lambda-\alpha_n)$ и $C_j\in\mathfrak{R}$.

Из (2. 14), (2. 16) и (2. 17) вытекает, что

$$L(\lambda) = \sum_{j=0}^{n-1} \lambda^j C_j(\lambda I-cI+X),$$

причем первый множитель обратим при $|\lambda-c| \leqq r$, а второй — при $|\lambda-c| \geqq r$. Полагая $B_j = -C_j$ ($j = 0, 1, \ldots, n-1$), $Z = cI-X$, получим равенство (2. 12)*)

Для окончания доказательства теоремы осталось установить, что оператор Z подобен самосопряженному. Положим

$$M(\lambda) = (\lambda-c)A(\lambda) = \Pi^{-1}(\lambda)L(\lambda)$$

и рассмотрим ограниченные (в силу предложения 1°) операторы

$$G = \frac{1}{2\pi i}\int_\Gamma M^{-1}(\lambda)\,d\lambda, \quad H = \frac{1}{2\pi i}\int_\Gamma A^{-1}(\lambda)\,d\lambda.$$

Очевидно,

$$G^* = -\frac{1}{2\pi i}\int_\Gamma [M^{-1}(\lambda)]^*\,d\bar\lambda = -\frac{1}{2\pi i}\int_\Gamma M^{-1}(\bar\lambda)\,d\bar\lambda = \frac{1}{2\pi i}\int_\Gamma M^{-1}(\lambda)\,d\lambda = G.$$

---

*) Отметим, что факторизация (2.12) единственна. В самом деле, если $L_+(\lambda)(Z-\lambda I) = \tilde L_+(\lambda)(\tilde Z-\lambda I)$, то из равенства $\tilde L_+^{-1}(\lambda)\,L_+(\lambda) = (\tilde Z-\lambda I)(Z-\lambda I)^{-1}$ вытекает, что $(\tilde Z-\lambda I)(Z-\lambda I)^{-1}$ — голоморфная во всей расширенной плоскости оператор-функция. Следовательно, $(\tilde Z-\lambda I)(Z-\lambda I)^{-1} \equiv I$, т. е. $\tilde Z=Z$, а значит и $\tilde L_+(\lambda) = L_+(\lambda)$.

Точно так же доказывается, что $H^* = H$.

Покажем, что оператор Z симметризуется справа оператором G. В самом деле,

$$ZG = \frac{1}{2\pi i}\int_\Gamma (Z-\lambda I)\, M^{-1}(\lambda)\, d\lambda + \frac{1}{2\pi i}\int_\Gamma (\lambda - c)\, M^{-1}(\lambda)\, d\lambda + cG =$$

$$= H + cG + \frac{1}{2\pi i}\int_\Gamma \Pi(\lambda)(Z-\lambda I)\, L^{-1}(\lambda)\, d\lambda.$$

В силу равенства (2. 12) подынтегральное выражение в последнем интеграле равно $\Pi(\lambda)L_+^{-1}(\lambda)$, и так как эта оператор-функция голоморфна внутри Γ, то интеграл равен нулю. Таким образом $ZG = H_1$ (где $H_1 = H + cG$), а это и означает, что оператор Z симметризуется справа оператором G. Для завершения доказательства осталось установить, что $G \gg 0$.

Рассмотрим квадратичную форму оператора G:

$$(Gf, f) = \frac{1}{2\pi i}\int_\Gamma (M^{-1}(\lambda) f, f)\, d\lambda = \frac{1}{2\pi i}\int_\Gamma (A^{-1}(\lambda) f, f)(\lambda - c)^{-1}\, d\lambda.$$

Производя замену $\lambda = c + re^{i\theta}$ и учитывая, что $(Gf, f) = \mathrm{Re}\,(Gf, f)$, получим

$$(Gf, f) = \frac{1}{2\pi}\int_0^{2\pi} (A^{-1}(c + re^{i\theta}) f, f)\, d\theta = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{Re}\,(A^{-1}(c + re^{i\theta}) f, f)\, d\theta =$$

$$= \frac{1}{2\pi}\int_0^{2\pi} \mathrm{Re}\,(f, A^{-1}(c + re^{i\theta}) f)\, d\theta = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{Re}\,(A(c + re^{i\theta})\, g(\theta), g(\theta))\, d\theta,$$

где $g(\theta) = A^{-1}(c + re^{i\theta}) f$. Так как $\|g(\theta)\| \geqq \varrho \|f\|$, где

$$\varrho^{-1} = \max_{\lambda \in \Gamma} \|A(\lambda)\|,$$

то в силу неравенства (2. 13) получим

$$(Gf, f) \geqq \frac{\delta_1}{2\pi}\int_0^{2\pi} \|g(\theta)\|^2\, d\theta \geqq \delta_1 \varrho^2 \|f\|^2.$$

Теорема доказана.

Следствие 1. *В условиях теоремы 4* $\sigma(Z) = \sigma(L) \cap \bar{\Delta}_k$, *собственные числа* $L(\lambda)$ *на отрезке* $\bar{\Delta}_k$ *совпадают с собственными числами оператора* Z, *и этим числам отвечают одни и те же собственные векторы.*

Это утверждение непосредственно следует из равенства (2. 12) и обратимости оператора $L_+(\lambda)$ $(\lambda \in \bar{\Delta}_k)$.

Следствие 2. *Если выполнены условия теоремы 4 и при некотором $\gamma \in \bar{\Delta}_k$ оператор $L(\gamma) \in \mathfrak{S}_\infty$, то $\sigma(L) \cap \bar{\Delta}_k$ состоит из γ и последовательности собственных чисел конечной кратности, сходящейся к γ. Если, кроме того, $\mathfrak{H}$ сепарабельно, то последовательность соответствующих собственных векторов пучка $L(\lambda)$ образует безусловный базис пространства $\mathfrak{H}$.*

В самом деле, из равенства (2. 12) и обратимости оператора $L_+(\lambda)$ вытекает, что $Z-\gamma I \in \mathfrak{S}_\infty$, и поэтому первое утверждение вытекает из следствия 1. Так как оператор $Z-\gamma I$ подобен самосопряженному, то его собственные векторы образуют безусловный базис в $\mathfrak{H}$, а в силу следствия 1 эти собственные векторы совпадают с соответствующими собственными векторами пучка $L(\lambda)$.

Замечание 1. Из факторизационного равенства

$$L(\lambda) = \sum_{j=0}^{n-1} \lambda^j B_j (Z-\lambda I)$$

следует, что

$$A_0 = B_0 Z, \quad A_j = B_j Z - B_{j-1} \quad (j = 1, 2, \ldots, n-1), \quad A_n = -B_{n-1},$$

и поэтому

$$\sum_{j=0}^{n} A_j Z^j = B_0 Z + \sum_{j=1}^{n-1} (B_j Z - B_{j-1}) Z^j - B_{n-1} Z^n = 0.$$

Таким образом, оператор Z является корнем уравнения

$$A_n Z^n + A_{n-1} Z^{n-1} + \cdots + A_1 Z + A_0 = 0. \tag{2.18}$$

Заметим, что и обратно, если Z — корень уравнения (2. 18), то

$$L(\lambda) = \sum_{k=0}^{n} A_k (\lambda^k I - Z^k) = L_+(\lambda)(Z-\lambda I),$$

где

$$L_+(\lambda) = -\sum_{k=1}^{n} A_k \sum_{j=0}^{k-1} \lambda^j Z^{k-j-1}.$$

Однако из (2. 18) еще не следует обратимость $L_+(\lambda)$ при $\lambda \in \sigma(Z)$, и для установления этого требуется провести дополнительное исследование (см., например, [2]).

Замечание 2. Нетрудно убедиться, что если Z является корнем уравнения (2. 18), то он симметризуется слева самосопряженным оператором

$$S = \sum_{k=0}^{n-1} \sum_{j=0}^{k} (Z^*)^j A_{k+1} Z^{k-j}. \tag{2.19}$$

Поэтому другой путь установления подобия оператора Z самосопряженному в теореме 4 состоит в доказательстве равномерной дефинитности оператора S. Этот метод и использовался вначале авторами, однако его применение натолкнулось на технические трудности, которые удалось преодолеть лишь в предположении, что $L(\gamma)\in\mathfrak{S}_\infty$ и $(-1)^{k-1}L'(\gamma)\gg 0$ для некоторого $\alpha\in\bar{\Delta}_k$*). Приведем схему этого доказательства.

С помощью сдвига $\lambda\to\lambda+\gamma$ доказательство сводится к случаю, когда $A_0\in\mathfrak{S}_\infty$, $(-1)^{k-1}A_1\gg 0$ и $0\in\bar{\Delta}_k$. Из неравенства (1. 2) выводится без труда, что

$$\|L^{-1}(\lambda)\|\leqq\frac{C_1}{|\operatorname{Im}\lambda|}\qquad(|\lambda-c|\leqq r),$$

и поэтому из (2. 12) вытекает, что

$$\|(Z-\lambda I)^{-1}\|\leqq\frac{C_2}{|\operatorname{Im}\lambda|}.$$

Так как оператор $Z\,(=L_+^{-1}(0)A_0)$ вполне непрерывен, то отсюда следует, что система собственных векторов каждого из операторов Z и Z^* полна в $\mathfrak{H}$ (см., например, [10], предложение 4. 5°). Если $Z\varphi_j=\lambda_j\varphi_j$, то, как нетрудно проверить, $(S\varphi_j,\varphi_j)=(L'(\lambda_j)\varphi_j,\varphi_j)$ и $(S\varphi_j,\varphi_k)=0$ $(\lambda_j\neq\lambda_k)$, и поэтому

$$\left(S\left(\sum_{j=1}^m a_j\varphi_j\right),\ \sum_{j=1}^m a_j\varphi_j\right)=\sum_{j=1}^m|a_j|^2\,(L'(\lambda_j)\varphi_j,\varphi_j).$$

Так как $\lambda_j=p_k(\varphi_j)$, то согласно (1. 3)

$$(-1)^{k-1}(S\psi,\psi)>0 \tag{2.20}$$

для любого ψ $(\neq 0)$, являющегося линейной комбинацией собственных векторов Z. В силу полноты этих векторов $(-1)^{k-1}S\geqq 0$.

Из равенства (2. 19) следует, что $S=A_1+T$, где $T\in\mathfrak{S}_\infty$, и поэтому $\operatorname{im}S$ замкнуто и $\dim\ker S=\dim\mathfrak{H}/\operatorname{im}S<\infty$. Следовательно, для доказательства соотношения $(-1)^{k-1}S\gg 0$ достаточно показать, что $\overline{\operatorname{im}S}=\mathfrak{H}$, а для этого (в силу полноты собственных векторов оператора Z^*) достаточно установить, что любое собственное подпространство $\mathfrak{L}_\lambda(Z^*)$ оператора Z^* входит в $\operatorname{im}S$. Для простоты будем далее предполагать, что $\lambda\neq 0$, т. е. что $\ker Z^*=\{0\}$. Общий случай сводится к этому, так как имеет место разложение пространства

$$\mathfrak{H}=\ker Z\dotplus\overline{\operatorname{im}Z}$$

(см. [11], стр. 637), а для оператора $\tilde{Z}=Z|\overline{\operatorname{im}Z}$ указанное предположение выполняется.

*) Отметим, что равномерная положительность S в условиях теоремы 4 (и более общих) была установлена В. И. Ломоносовым [15] еще до того, как авторами было получено приведенное выше доказательство равномерной положительности правого симметризатора G.

Так как $SZ=Z^*S$, то $S(\mathfrak{L}_\lambda(Z))\subset\mathfrak{L}_\lambda(Z^*)$. Поскольку $\dim\mathfrak{L}_\lambda(Z)=\dim\mathfrak{L}_\lambda(Z^*)<$ $<\infty$ и из равенства $S\varphi=0$ $(\varphi\in\mathfrak{L}_\lambda(Z))$ следует, что $\varphi=0$ (см. (2. 20)), то $S(\mathfrak{L}_\lambda(Z))=$ $=\mathfrak{L}_\lambda(Z^*)$, что и завершает доказательство.

3. Здесь мы приведем пример, показывающий что утверждение теоремы 4 перестает быть верным без условия отделенности спектральной зоны Δ_k. Точнее говоря, строится квадратичный гиперболический пучок $L(\lambda)$, для которого не существует факторизации (2. 12) с оператором Z, подобным самосопряженному.

Отметим, что в силу результата Г. Лангера [3] всякий квадратичный гиперболический пучок $L(\lambda)=\lambda^2I+\lambda B+C$ допускает факторизацию $L(\lambda)=$ $=(Z_2^*-\lambda I)(Z_1-\lambda I)$, где $\sigma(Z_j)\subset\bar{\Delta}_j$ $(j=1,2)$*). Естественно предположить, что этот результат допускает обобщение на случай $n>2$, т. е. что любой гиперболический пучок $L(\lambda)$ допускает факторизацию (2. 12), где $\sigma(Z)\subset\bar{\Delta}_k$ и $L_+(\lambda)$ обратим для внутренних точек λ зоны Δ_k.

Перейдем к построению указанного выше примера. Пусть $\mathfrak{H}$ — сепарабельное гильбертово пространство. Представим $\mathfrak{H}$ в виде ортогональной суммы двумерных подпространств $\mathfrak{H}_j$ и рассмотрим операторы B_j и C_j, заданные в некотором ортонормированном базисе пространства $\mathfrak{H}_j$ матрицами

$$B_j=\begin{pmatrix} b_j' & b_j \\ b_j & b_j'' \end{pmatrix}, \quad C_j=\begin{pmatrix} c_j' & 0 \\ 0 & c_j'' \end{pmatrix},$$

где числа b_j', b_j'' вещественны, а b_j, c_j', c_j'' — положительны.

Обозначим через B и C операторы, являющиеся ортогональными суммами операторов B_j и C_j соответственно, и положим $L(\lambda)=\lambda^2I+\lambda B-C$. При условии

$$\lim b_j=\lim b_j'=\lim b_j''=\lim c_j'=\lim c_j''=0, \tag{2.21}$$

B и C являются вполне непрерывными самосопряженными операторами, причем $C>0$. Из последнего неравенства вытекает, что квадратный трехчлен $(L(\lambda)f,f)$ имеет при любом $f\neq0$ различные вещественные корни

$$p_{1,2}(f)=\frac{-(Bf,f)\pm\sqrt{(Bf,f)^2+4(Cf,f)(f,f)}}{2}.$$

Очевидно, $p_2(f)<0<p_1(f)$, а так как $B, C\in\mathfrak{S}_\infty$, то

$$\sup p_2(f)=\inf p_1(f)=0,$$

т. е. 0 является общей точкой $\bar{\Delta}_1$ и $\bar{\Delta}_2$.

---

*) М. Г. Крейн и Г. Лангер [2] получили ранее этот результат при дополнительных условиях $C>0$ и $C\in\mathfrak{S}_\infty$, установив, кроме того, что в этом случае операторы Z_1 и Z_2 подобны самосопряженным. Приводимый пример показывает, что без дополнительных ограничений последнее утверждение не имеет места.

Пучок $L(\lambda)$ распадается в ортогональную сумму двумерных пучков $L_j(\lambda) = \lambda^2 I_j + \lambda B_j - C_j$. Каждый из этих пучков имеет два положительных собственных числа $\lambda_{j1}, \lambda_{j2}$ (а также два отрицательных собственных числа). Эти собственные числа являются корнями уравнения

(2.22) $$(\det L_j(\lambda) =)(\lambda^2 + \lambda b_j' - c_j')(\lambda^2 + \lambda b_j'' - c_j'') - \lambda^2 b_j^2 = 0.$$

Будем предполагать, что

(2.23) $$\lambda_{j1} \neq \lambda_{j2}.$$

Собственные векторы пучка $L_j(\lambda)$, отвечающие собственным числам λ_{j1} и λ_{j2}, обозначим φ_{j1} и φ_{j2}. Очевидно, можно положить

$$\varphi_{j1} = (\lambda_{j1} b_j, c_j' - \lambda_{j1} b_j' - \lambda_{j1}^2), \quad \varphi_{j2} = (\lambda_{j2} b_j, c_j' - \lambda_{j2} b_j' - \lambda_{j2}^2).$$

Если выполнены условия

(2.24) $$\lim \frac{c_j' - \lambda_{j1} b_j' - \lambda_{j1}^2}{\lambda_{j1} b_j} = \lim \frac{c_j' - \lambda_{j2} b_j' - \lambda_{j2}^2}{\lambda_{j2} b_j} = 0,$$

то, как легко видеть, угол между векторами φ_{j1} и φ_{j2} стремится к нулю. Отсюда вытекает, что последовательность векторов, полученная объединением в каком-либо порядке последовательностей $\{\varphi_{j1}\}_1^\infty$ и $\{\varphi_{j2}\}_1^\infty$, не является базисом $\mathfrak{H}$. Но тогда для спектральной зоны Δ_1 пучка $L(\lambda)$ не имеет место утверждение теоремы 4, так как в противном случае в силу условия $L(0) = -C \in \mathfrak{S}_\infty$ собственные векторы пучка $L(\lambda)$, отвечающие его положительным собственным числам, образовывали бы безусловный базис $\mathfrak{H}$ (см. следствие 2). Таким образом, построение примера свелось к выбору последовательностей вещественных чисел b_j', b_j'' и положительных чисел b_j, c_j', c_j'', λ_{j1}, λ_{j2} так, чтобы выполнялись условия (2.21), (2.23), (2.24) и чтобы числа λ_{j1}, λ_{j2} были корнями уравнения (2.22).

Потребуем, чтобы выполнялись равенства

(2.25) $$\lambda_{j1}^2 + \lambda_{j1} b_j' - c_j' - j^{-1} \lambda_{j1} b_j = 0,$$

(2.26) $$\lambda_{j2}^2 + \lambda_{j2} b_j' - c_j' + j^{-1} \lambda_{j2} b_j = 0,$$

что будет гарантировать справедливость соотношений (2.24). Если, кроме того, будут выполняться равенства

(2.27) $$\lambda_{j1}^2 + \lambda_{j1} + b_j'' - c_j'' \lambda_{j1} b_j = 0,$$

(2.28) $$\lambda_{j2}^2 + \lambda_{j2} b_j'' - c_j'' + j \lambda_{j2} b_j = 0,$$

то из (2.25) и (2.27) будет следовать, что λ_{j1} является корнем уравнения (2.22), а из (2.26) и (2.28) — что λ_{j2} является корнем этого уравнения.

Будем считать b_j, b'_j и c'_j заданными и выразим λ_{j1} из (2. 25), а λ_{j2} из (2. 26):

$$\lambda_{j1} = \tfrac{1}{2}\left(\sqrt{(b'_j - j^{-1}b_j)^2 + 4c'_j} - b'_j + j^{-1}b_j\right), \tag{2.29}$$

$$\lambda_{j2} = \tfrac{1}{2}\left(\sqrt{(b'_j + j^{-1}b_j)^2 + 4c'_j} - b'_j - j^{-1}b_j\right) \tag{2.30}$$

(очевидно, $\lambda_{j1} > 0$, $\lambda_{j2} > 0$). Далее, из (2. 27) и (2. 28) находим b''_j и c''_j:

$$b''_j = -(\lambda_{j1} + \lambda_{j2}) + \frac{jb_j(\lambda_{j1} + \lambda_{j2})}{\lambda_{j1} - \lambda_{j2}}, \tag{2.31}$$

$$c''_j = \lambda_{j1}\lambda_{j2}\left(\frac{2jb_j}{\lambda_{j1} - \lambda_{j2}} - 1\right). \tag{2.32}$$

Положим

$$b_j = j^{-2}, \quad b'_j = j^{-4}, \quad c'_j = j^{-7}. \tag{2.33}$$

Тогда из (2. 29), (2. 30) и (2. 33) вытекают равенства

$$\lim j^3 \lambda_{j1} = 1, \quad \lim j^4 \lambda_{j2} = 1. \tag{2.34}$$

Из (2. 31), (2. 33) и (2. 34) следует, что $b''_j \to 0$. В силу (2. 33) и (2. 34)

$$\lim \frac{2jb_j}{\lambda_{j1} - \lambda_{j2}} = +\infty. \tag{3.35}$$

Выберем j_0 настолько большим, чтобы при $j \geqq j_0$ выполнялись неравенства (2. 23) и неравенство $c''_j > 0$, и будем далее в качестве основного пространства $\mathfrak{H}$ рассматривать ортогональную сумму подпространств $\mathfrak{H}_j$ при $j \geqq j_0$.

Так как

$$c''_j < \lambda_{j1}\lambda_{j2}\frac{2jb_j}{\lambda_{j1} - \lambda_{j2}},$$

то $\lim c''_j = 0$. Таким образом, все требуемые условия выполнены, что и завершает построение примера.

4. Здесь мы рассмотрим некоторые обобщения полученных выше результатов на случай, когда старший коэффициент пучка не является равномерно положительным.

Всюду в этом пункте предполагается, что рассматриваемый пучок

$$L(\lambda) = \sum_{j=0}^{n} \lambda^j A_j \qquad (A_j^* = A_j \in \mathfrak{R},\ j = 0, 1, \ldots, n;\ A_n \neq 0) \tag{2.36}$$

удовлетворяет следующим условиям: 1) при $(A_n f, f) \neq 0$ многочлен $(L(\lambda)f, f)$ имеет n различных вещественных корней; 2) при $f \neq 0$ и $(A_n f, f) = 0$ выполняется неравенство $(A_{n-1}f, f) \neq 0$ и $(L(\lambda)f, f)$ имеет $n-1$ различных вещественных корней (в этом случае n-ый корень можно считать бесконечным); 3) из соотношений $\lim\limits_{j\to\infty} (A_k f_j, f_j) = 0$ $(k = 0, 1, \ldots, n)$ вытекает, что $\|f_j\| \to 0$.

Нетрудно убедиться, что условие 3) существенно. Если оно не выполнено, то спектр пучка $L(\lambda)$ может заполнить всю плоскость, как показывает следующий пример. Пусть a_0, a_1, a_2 — вещественные числа такие, что $a_1^2 > 4a_0a_2$; $T \in \mathfrak{S}_\infty$, $T > 0$ и $L(\lambda) = (a_0 + a_1\lambda + a_2\lambda^2)T$. Легко видеть, что $(L(\lambda)f, f)$ имеет при любом $f \neq 0$ различные вещественные корни (которые не зависят от f), и в то же время спектр $L(\lambda)$ есть вся плоскость.

Обозначим через $\Delta(L)$ множество вещественных чисел, состоящее из всех корней многочленов $(L(\lambda)f, f)$ при любых $f \neq 0$.

Лемма 4. *Если выполнены условия 1)—3), то $\sigma(L) \subset \overline{\Delta(L)}$. Если, кроме того, $\lambda_0 \notin \overline{\Delta(L)}$ и вещественно, то оператор $L(\lambda_0)$ равномерно дефинитный.*

Доказательство. Так как $\sigma(A) \subset \overline{W(A)}$ для любого оператора A, то достаточно показать, что из условия $\lambda_0 \notin \overline{\Delta(L)}$ вытекает: $0 \notin \overline{W(L(\lambda_0))}$. Допустим, что это не так. Тогда найдется последовательность векторов $\{f_j\}_1^\infty$ такая, что $\|f_j\| = 1$ и $\lim (L(\lambda_0)f_j, f_j) = 0$. Последовательность многочленов $(L(\lambda)f_j, f_j)$ содержит подпоследовательность, сходящуюся к некоторому многочлену $P(\lambda)$. Очевидно, $P(\lambda_0) = 0$ и $P(\lambda) \not\equiv 0$ (в силу условия 3)). По теореме Гурвица в любой окрестности точки λ_0 должен содержаться хотя бы один корень всякого многочлена с достаточно большим номером из указанной подпоследовательности, что противоречит условию $\lambda_0 \notin \overline{\Delta(L)}$.

Если же $\bar\lambda_0 = \lambda_0$, то $[L(\lambda_0)]^* = L(\lambda_0)$, и так как $0 \notin \overline{W(L(\lambda_0))}$, то оператор $L(\lambda_0)$ является равномерно дефинитным. Лемма доказана.

Покажем теперь, как с помощью леммы 4 можно свести пучок (2. 36) к некоторому гиперболическому пучку. При этом мы будем предполагать, что существует хотя бы одно вещественное число $\lambda_0 \notin \overline{\Delta(L)}$. Тогда в силу леммы 4 оператор $L(\lambda_0)$ является равномерно дефинитным. Полагая $\lambda = \lambda_0 + \mu^{-1}$, получим

(2. 37)
$$L(\lambda) = \sum_{k=0}^{n} \frac{L^{(k)}(\lambda_0)}{k!}\mu^{-k} = \frac{M(\mu)}{\mu^n},$$

где

$$M(\mu) = \sum_{k=0}^{n} \frac{1}{k!} L^{(k)}(\lambda_0)\mu^{n-k}.$$

Очевидно, один из пучков $M(\mu)$ и $-M(\mu)$ является гиперболическим.

Если Δ_j' $(j = 1, 2, \ldots, n)$ — спектральные зоны пучка $M(\mu)$, то спектральными зонами*) пучка $L(\lambda)$ назовем множества $\Delta_j = f(\Delta_j')$ $(j = 1, 2, \ldots, n)$, где $f(\mu) = \lambda_0 + \mu^{-1}$. Очевидно, $\bigcup_{j=1}^{n} \Delta_j = \Delta(L)$.

---

*) Можно было бы дать непосредственное определение спектральных зон для пучка (2.36), однако мы предпочитаем приведенное формальное определение, использующее редукцию к гиперболическому пучку.

Из приведенного определения следует, что у пучка (2. 36) могут быть одна или две неограниченные спектральные зоны.

Будем говорить, что две спектральные зоны пучка $L(\lambda)$ отделены, если отделены соответствующие спектральные зоны пучка $M(\mu)$.

Теорема 5. *Пусть $L(\lambda)$ удовлетворяет условиям 1)—3) и $\overline{\Delta(L)}$ не совпадает с вещественной осью. Если Δ — ограниченная (соответственно неограниченная) спектральная зона $L(\lambda)$, отделенная от других зон, то $L(\lambda)$ допускает факторизацию $L(\lambda) = L_+(\lambda)(Z-\lambda I)$ (соответственно $L(\lambda) = L_+(\lambda)(I-(\lambda-\lambda_0)Z))$,** где первый множитель обратим при всех $\lambda \in \bar{\Delta}$, а второй — при всех $\lambda \notin \bar{\Delta}$, причем Z подобен самосопряженному оператору.*

Эта теорема непосредственно выводится из теоремы 4 с помощью равенства (2. 37).

§ 3. Некоторые частные случаи

1. Здесь мы будем рассматривать гиперболический пучок (1. 1) с неотрицательными коэффициентами. Заметим, что с помощью сдвига $\lambda = \mu + a$ $(a>0)$ к такому виду можно свести любой гиперболический пучок. Начнем с алгебраической леммы, в доказательстве которой используются методы статьи [12].

Лемма 5. *Пусть многочлен*

$$l(\lambda) = a_n\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0 \qquad (a_n > 0;\ a_j \geqq 0,\ j = 0, 1, \ldots, n-1;\ n \geqq 2)$$

имеет различные вещественные корни $\{p_j\}_1^n$ $(p_1 > p_2 > \cdots > p_n)$. Если при некотором k $(1 \leqq k \leqq n-1)$

(3. 1) $$a_k^2 \geqq 4a_{k-1}a_{k+1},$$

то

(3. 2) $$p_k - p_{k+1} \geqq \frac{1}{2l'(|p_{k+1}|)}\left[\left(\frac{a_k}{2a_{k+1}}\right)^{k-2} a_{k-2} + \left(\frac{a_k}{2a_{k+1}}\right)^{k+2} a_{k+2}\right].^{*)}$$

Доказательство. Так как $p_j < p_1 \leqq 0$ $(j > 1)$, то $a_j > 0$ $(j > 0)$. Из вещественности всех корней $l(\lambda)$ следуют неравенства (см., например, [13], стр. 22)

(3. 3) $$\frac{a_{j-1}}{a_j} \leqq \frac{j}{j+1}\frac{a_j}{a_{j+1}} \qquad (j = 1, 2, \ldots, n-1).$$

---

*) Здесь λ_0 — какое-нибудь вещественное число, не принадлежащее $\overline{\Delta(L)}$.

**) Мы считаем $a_j = 0$ при $j < 0$ и $j > n$.

Положим $z_k = -a_k/2a_{k+1}$ и

$$l_1(\lambda) = \sum_{j=2}^{k} a_{k-j}\lambda^{k-j}, \quad l_2(\lambda) = a_{k-1}\lambda^{k-1} + a_k\lambda^k + a_{k+1}\lambda^{k+1}, \quad l_3(\lambda) = \sum_{j=2}^{n-k} a_{k+j}\lambda^{k+j}.$$

В силу неравенства (3. 1)

$$(3.4)\qquad (-1)^k l_2(z_k) = (-1)^{2k-1}\left(\frac{a_k}{2a_{k+1}}\right)^{k-1}\frac{4a_{k-1}a_{k+1}-a_k^2}{4a_{k+1}} \geqq 0.$$

Из неравенств (3. 1) и (3. 3) следует, что при $1 \leqq j \leqq k-2$

$$\frac{a_{j-1}|z_k|^{j-1}}{a_j|z_k|^j} = \frac{a_{j-1}}{a_j}\frac{2a_{k+1}}{a_k} \leqq \frac{1}{2}\frac{a_{j-1}a_k}{a_j a_{k-1}} < \frac{1}{2},$$

и поэтому $(-1)^k l_1(z_k)$ представляет собой сумму убывающих по абсолютной величине и знакочередующихся слагаемых. Следовательно,

$$(3.5)\qquad (-1)^k l_1(z_k) \geqq \tfrac{1}{2}(-1)^k a_{k-2} z_k^{k-2} = \tfrac{1}{2}\left(\frac{a_k}{2a_{k+1}}\right)^{k-2} a_{k-2}.$$

Из неравенств (3. 3) вытекает, что при $k+2 \leqq j < n$

$$\frac{a_{j+1}|z_k|^{j+1}}{a_j|z_k|^j} = \frac{a_{j+1}}{a_j}\frac{a_k}{2a_{k+1}} < \frac{1}{2},$$

и поэтому

$$(3.6)\qquad (-1)^k l_3(z_k) \geqq \tfrac{1}{2}(-1)^k a_{k+2} z_k^{k+2} = \tfrac{1}{2}\left(\frac{a_k}{2a_{k+1}}\right)^{k+2} a_{k+2}.$$

Из неравенств (3. 4)—(3. 6) получаем

$$(3.7)\qquad (-1)^k l(z_k) \geqq \tfrac{1}{2}\left[\left(\frac{a_k}{2a_{k+1}}\right)^{k-2} a_{k-2} + \left(\frac{a_k}{2a_{k+1}}\right)^{k+2} a_{k+2}\right].$$

Если мы установим, что $p_{k+1} < z_k \leqq p_k$, то тогда $p_k - p_{k+1} \geqq z_k - p_{k+1}$. По теореме Лагранжа

$$z_k - p_{k+1} = \frac{l(z_k) - l(p_{k+1})}{l'(\xi)} = \frac{l(z_k)}{l'(\xi)} \qquad (p_{k+1} < \xi < z_k),$$

и поэтому

$$(3.8)\qquad p_k - p_{k+1} \geqq \left|\frac{l(z_k)}{l'(\xi)}\right| \geqq \frac{|l(z_k)|}{l'(|p_{k+1}|)}.$$

Неравенство (3. 2) вытекает тогда из (3. 7) и (3. 8). Таким образом, для доказательства леммы осталось установить, что всякий многочлен $l(\lambda)$, удовлетворяющий условиям леммы, имеет на отрезке $[z_k, 0]$ ровно k корней.

Докажем это утверждение индукцией по k. Пусть вначале $k=1$. Так как $l(0) \geqq 0$ и $l(z_1) \leqq 0$ (см. (3. 7)), то на отрезке $[z_1, 0]$ есть хоть один корень $l(\lambda)$.

С другой стороны, $l'(\lambda) \geqq 3a_3\lambda^2 + \cdots + na_n\lambda^{n-1}$ при $z_1 \leqq \lambda \leqq 0$, и так как в силу (3. 3)

$$\frac{(j+1)a_{j+1}|\lambda|^j}{ja_j|\lambda|^{j-1}} \leqq \frac{(j+1)a_{j+1}}{ja_j}\frac{a_1}{2a_2} \leqq \frac{a_ja_1}{2a_{j-1}a_2} < \tfrac{1}{2} \qquad (|\lambda| \leqq |z_1|, \quad 3 \leqq j < n),$$

то $l'(\lambda) \geqq \frac{3}{2}a_3\lambda^2 \geqq 0$ $(z_1 \leqq \lambda \leqq 0)$. Поэтому $l(\lambda)$ имеет на отрезке $[z_1, 0]$ ровно один корень.

Предположим теперь, что утверждение верно для $k-1$ (и для многочленов любой степени $\geqq k$) и установим его справедливость для индекса k $(\geqq 2)$.

Рассмотрим $l'(\lambda) = a_1 + 2a_2\lambda + \cdots + na_n\lambda^{n-1}$. Из (3. 1) вытекает, что

$$(ka_k)^2 > 4(k-1)a_{k-1}(k+1)a_{k+1},$$

и в силу индуктивного предположения многочлен $l'(\lambda)$ имеет на отрезке $[z'_{k-1}, 0]$ (где $z'_{k-1} = -ka_k[2(k+1)a_{k+1}]^{-1}$) ровно $k-1$ корней. Очевидно, $z_k < z'_{k-1}$. Покажем, что $l'(\lambda) \neq 0$ на интервале (z_k, z'_{k-1}). Для этого представим $l'(\lambda)$ в виде $l'(\lambda) = l'_1(\lambda) + l'_2(\lambda) + l'_3(\lambda)$ и оценим каждое слагаемое.

Положим

$$r(\lambda) = (k-1)a_{k-1} + ka_k\lambda + (k+1)a_{k+1}\lambda^2.$$

Так как $r(z'_{k-1}) < 0$ и $r(z_k) \leqq 0$, то $r(\lambda) < 0$ на интервале (z_k, z'_{k-1}) и, следовательно,

(3.9) $$(-1)^{k-1}l'_2(\lambda) > 0 \qquad (z_k < \lambda < z'_{k-1}).$$

В силу (3. 1) и (3. 3) при $|\lambda| \geqq |z'_{k-1}|$ и $1 \leqq j \leqq k-3$ выполняются неравенства

$$\frac{ja_j|\lambda|^{j-1}}{(j+1)a_{j+1}|\lambda|^j} \leqq \frac{ja_j}{(j+1)a_{j+1}}\frac{2(k+1)a_{k+1}}{ka_k} < \frac{a_ja_k}{2a_{j+1}a_{k-1}} < \tfrac{1}{2},$$

и поэтому

(3.10) $$(-1)^{k-1}l'_1(\lambda) \geqq \tfrac{1}{2}(k-2)a_{k-2}|\lambda|^{k-3} \qquad (\lambda \leqq z'_{k-1})$$

Наконец, из (3. 3) вытекает, что при $|\lambda| \leqq |z_k|$ и $k+2 \leqq j < n$

$$\frac{(j+1)a_{j+1}|\lambda|^j}{ja_j|\lambda|^{j-1}} \leqq \frac{(j+1)a_{j+1}}{ja_j}\frac{a_k}{2a_{k+1}} \leqq \frac{a_ja_k}{2a_{j-1}a_{k+1}} < \tfrac{1}{2}.$$

Следовательно,

(3.11) $$(-1)^{k-1}l'_3(\lambda) \geqq \tfrac{1}{2}(k+2)a_{k+2}|\lambda|^{k+1} \qquad (0 \geqq \lambda \geqq z_k).$$

Из (3. 9)—(3. 11) получаем, что $(-1)^{k-1}l'(\lambda) > 0$ $(z_k < \lambda < z'_{k-1})$. Таким образом, $l'(\lambda)$ имеет ровно $k-1$ корней в полуинтервале $(z_k, 0]$. Так как корни $l(\lambda)$ и $l'(\lambda)$ перемежаются, то $l(\lambda)$ имеет на отрезке $[z_k, 0]$ не более k и не менее $k-2$ корней. Однако, учитывая, что $l(0) \geqq 0$, $l'(0) > 0$ и $(-1)^k\, l(z_k) \geqq 0$, заключаем, что $l(\lambda)$ имеет на отрезке, $[z_k, 0]$ ровно k корней.

Лемма доказана.

Нам понадобится следующая лемма, являющаяся следствием общего предложения Ю. Л. Шмульяна ([14], лемма 1. 1), которое относится к операторам в банаховом пространстве.

Лемма 6. *Пусть A, B, C — ограниченные неотрицательные операторы, $B \gg 0$ и $C \neq 0$. Если $(Af,f)^2 \geqq (Bf,f)(Cf,f)$ для любого $f \in \mathfrak{H}$, то $A \gg 0$.*

Отметим, что при условии $C > 0$ приведенный результат был установлен ранее М. Г. Крейном и Г. Лангером [2].

Теорема 6. *Пусть $L(\lambda) = \sum_{j=0}^{n} \lambda^j A_j$ $(n > 3)$ — гиперболический пучок, $A_j \geqq 0$ $(j = 0, 1, \ldots, n-1)$ и $A_0 \neq 0$. Если при некотором k $(1 \leqq k \leqq n)$ и при любом $f \in \mathfrak{H}$*

(3. 12) $(A_k f,f)^2 \geqq 4(A_{k-1}f,f)(A_{k+1}f,f), \quad (A_{k-1}f,f)^2 \geqq 4(A_k f,f)(A_{k-2}f,f);$ *)

то спектральная зона Δ_k отделена от соседних зон.

Доказательство. Так как $L(\lambda)$ — гиперболический пучок, то имеют место неравенства (3. 3), т. е.

(3. 13) $$(A_j f,f)^2 \geqq \frac{j+1}{j}(A_{j-1}f,f)(A_{j+1}f,f) \qquad (j = 1, 2, \ldots, n-1; \quad f \in \mathfrak{H}).$$

Отметим также, что $(A_j f,f) > 0$ $(j > 0, f \neq 0)$, ибо $p_j(f) < 0$ при $j > 1$. Поэтому применяя последовательно (начиная с $j = n-1$) лемму 6, получаем из (3. 13), что $A_j \gg 0$ $(j > 0)$.

Из леммы 5 вытекает, что

$$p_k(f) - p_{k+1}(f) \geqq \frac{1}{2\sum_{j=1}^{n} j\|A_j\||\alpha_{k+1}^{j-1}|}\left[\left(\frac{m_k}{2\|A_{k+1}\|}\right)^{k-2} m_{k-2} + \left(\frac{m_k}{2\|A_{k+1}\|}\right)^{k+2} m_{k+2}\right],$$

где $\alpha_{k+1} = \inf \Delta_{k+1}$, $m_j = \inf_{\|f\|=1} (A_j f,f)$. В силу сказанного выше $m_k > 0$, и так как $n > 3$, то положительно также хоть одно из чисел m_{k-2}, m_{k+2}. Поэтому существует число $\delta_1 > 0$ такое, что при любом $f \neq 0$

$$p_k(f) - p_{k+1}(f) > \delta_1.$$

Аналогично устанавливается, что при некотором $\delta_2 > 0$ и любом $f \neq 0$

$$p_{k-1}(f) - p_k(f) > \delta_2.$$

*) Мы полагаем $A_{-1} = A_{n+1} = 0$, так что при $k = 1$ или $k = n$ остается одно из этих неравенств.

Теперь для окончания доказательства теоремы достаточно сослаться на теорему 2.

Замечание 3. Здесь мы обсудим случаи $n=2, 3$, которые не охватываются приведенной формулировкой теоремы 6. Отметим прежде всего, что в случае $n=3$ и $k=1$ теорема 6 сохраняет силу (без всяких изменений в доказательстве). Из доказательства теоремы 6 видно также, что она остается справедливой в случае $n=3$ и $k=2$ при дополнительном ограничении $A_0 \gg 0$. С другой стороны, просматривая доказательство леммы 5, нетрудно убедиться, что для справедливости утверждения теоремы 6 в случае $n=3$ и $k=2$ достаточно потребовать, чтобы при любом $f \in \mathfrak{H}$

$$(A_1 f, f)^2 \geqq 4(A_0 f, f)(A_2 f, f), \quad (A_2 f, f) \geqq (4+\varepsilon)(A_1 f, f)(A_3 f, f),$$

где ε — некоторое положительное число*).

Если же $n=2$, то зоны Δ_1 и Δ_2 отделены тогда и только тогда, когда существует число $\varepsilon > 0$ такое, что

$$(A_1 f, f)^2 \geqq (4+\varepsilon)(A_0 f, f)(A_2 f, f) \qquad (f \in \mathfrak{H}).$$

Это утверждение непосредственно вытекает из равенства

$$p_1(f) - p_2(f) = \frac{\sqrt{(A_1 f, f)^2 - 4(A_0 f, f)(A_2 f, f)}}{(A_2 f, f)},$$

теоремы 2 и равномерной положительности A_1.

Замечание 4. Если $A_n \gg 0$, $A_j \geqq 0$ $(j=0, 1, \ldots, n-1)$ и $A_0 \neq 0$, то при $n > 3$ из неравенств

$$(A_k f, f)^2 \geqq 4(A_{k-1} f, f)(A_{k+1} f, f) \qquad (k = 1, 2, \ldots, n-1; \; f \in \mathfrak{H})$$

вытекает, что $L(\lambda) = \sum_{j=0}^{n} \lambda^j A_j$ — гиперболический пучок и что все его спектральные зоны отделены друг от друга.

Первое утверждение этого замечания выводится без труда из леммы 6 и теоремы А. Ю. Левина [12], а тогда второе утверждение непосредственно следует из теоремы 6.

2. Рассмотрим квадратичный пучок вида

$$L(\lambda) = A - \lambda I + \lambda^2 B,$$

*) Приведем пример, показывающий, что в указанном случае утверждение теоремы 6 без дополнительных ограничений уже не верно. Пусть $L(\lambda) = 8\lambda^3 I + 16\lambda^2 I + 8\lambda I + A_0$, где $A_0 > 0$, $A_0 \in \mathfrak{S}_\infty$ и $\|A_0\| \leqq 1$. При $k=2$ условия (3.12), очевидно, выполнены, однако, если $\|f_j\| = 1$ и $(A_0 f_j, f_j) \to 0$, то $p_2(f_j) - p_3(f_j) \to 0$, и, следовательно, $\bar{\Delta}_2 \cap \bar{\Delta}_3$ непусто.

где A и B — ограниченные самосопряженные операторы. Если выполнено условие

$$(Af,f)(Bf,f) < \tfrac{1}{4}(f,f)^2 \qquad (f \neq 0), \tag{3.14}$$

то при $(Bf,f) \neq 0$ трехчлен $(L(\lambda)f,f)$ имеет два различных вещественных корня

$$p_{1,2}(f) = \frac{(f,f) \pm \sqrt{(f,f)^2 - 4(Af,f)(Bf,f)}}{2(Bf,f)}. \tag{3.15}$$

и, как легко видеть, пучок $L(\lambda)$ удовлетворяет условиям 1)—3) п. 4 § 2.

Обозначим через Δ_1 (соответственно Δ_2) множество всех корней $p_1(f)$ (соответственно $p_2(f)$). Для применения теоремы 5 надо показать, что $\overline{\Delta(L)}$ не совпадает с вещественной осью. Покажем, что $\bar{\Delta}_1 \cap \bar{\Delta}_2 = \emptyset$. При этом потребуем, чтобы выполнялся следующий усиленный вариант неравенства (3. 14):

$$(Af,f)(Bf,f) \leqq (\tfrac{1}{4} - \delta)(f,f)^2 \qquad (f \in \mathfrak{H}) \tag{3.16}$$

при некотором $\delta > 0$.

Так как в силу (3. 15)

$$|p_1(f) - p_2(f)| = \frac{\sqrt{(f,f)^2 - 4(Af,f)(Bf,f)}}{|(Bf,f)|}, \tag{3.17}$$

то из (3. 16) вытекает, что

$$|p_1(f) - p_2(f)| \geqq \frac{2\sqrt{\delta}}{\|B\|}.$$

Если теперь допустить, что существует вещественное число $\gamma \in \bar{\Delta}_1 \cap \bar{\Delta}_2$, то, повторяя рассуждения из доказательства теоремы 2, придем к противоречию. При этом соотношение $Q(\lambda) \not\equiv 0$ гарантируется тем, что коэффициент при λ в пучке $L(\lambda)$ равен $-I$. Кроме того, надо воспользоваться следующими соотношениями, вытекающими из (3. 15):

$$(L'(p_1(f))f,f) > 0 \quad ((Bf,f) \neq 0), \; (L'(p_2(f))f,f) < 0 \qquad (f \neq 0).$$

Таким образом, при условии (3. 16) $\bar{\Delta}_1 \cap \bar{\Delta}_2 = \emptyset$. Теперь нетрудно убедиться, что Δ_1 и Δ_2 — спектральные зоны пучка $L(\lambda)$ в смысле определения п. 4 § 2. Легко видеть, что зона Δ_2 ограничена, а Δ_1 ограничена тогда и только тогда, когда оператор B равномерно дефинитен. Так как соотношение $\bar{\Delta}_1 \cap \bar{\Delta}_2 = \emptyset$ означает, что зоны Δ_1 и Δ_2 отделены, то, в силу теоремы 5, из приведенных рассуждений вытекает следующая

Теорема 7. *Если при некотором $\delta > 0$ выполнено неравенство* (3. 16), *то имеет место равенство*

$$A - \lambda I + \lambda^2 B = (I - BZ - \lambda B)(Z - \lambda I),$$

где линейный пучок $I - BZ - \lambda B$ обратим при всех $\lambda \in \bar{\Delta}_2$, спектр оператора Z содержится в $\bar{\Delta}_2$ и Z подобен самосопряженному оператору.

Мы не приводим здесь формулировку теоремы о факторизации относительно зоны Δ_1, которая также вытекает из теоремы 5.

Замечание 5. Очевидно условие (3.16) будет выполнено, если $A \geqq 0$ и $B \leqq 0$ (или $A \leqq 0$ и $B \geqq 0$). Оно также выполнено, если $\|A\| \|B\| < \frac{1}{4}$.

Замечание 6. Если оба оператора A и B вполне непрерывны, то теорема 7 сохраняет силу при замене условия (3.16) условием (3.14).

Действительно, достаточно показать, что

$$\inf_{\|f\|=1} |p_1(f) - p_2(f)| > 0,$$

а для этого, в силу равенства (3.17), достаточно установить, что $\inf\limits_{\|f\|=1} F(f) > 0$, где

$$F(f) = 1 - 4(Af, f)(Bf, f).$$

Допустим, что это не так. Тогда существует такая нормированная последовательность $\{f_n\}_1^\infty$, что $F(f_n) \to 0$. Без ограничения общности можно считать, что последовательность $\{f_n\}_1^\infty$ слабо сходится к некоторому вектору g, и, следовательно,

$$\lim (Af_n, f_n) = (Ag, g), \quad \lim (Bf_n, f_n) = (Bg, g).$$

Так как $F(f_n) \to 0$, то $4(Ag, g)(Bg, g) = 1$, и из условия (3.14) вытекает, что $\|g\| > 1$. Последнее неравенство невозможно, так как g является слабым пределом нормированной последовательности $\{f_n\}_1^\infty$.

3. В этом пункте приводятся две теоремы, показывающие, что методы настоящей статьи допускают применение к некоторым классам пучков, квадратичные формы которых могут иметь и невещественные корни.

Теорема 8. *Пусть*

$$L(\lambda) = \sum_{j=0}^{n} \lambda^j A_j,$$

где A_j ($j = 0, 1, \ldots, n$) — ограниченные самосопряженные операторы. Если существуют вещественное число c и положительное число r такие, что для всех точек окружности $\Gamma = \{\lambda : |\lambda - c| = r\}$ выполняется условие

(3.18) $$\operatorname{Re} \frac{L(\lambda)}{\lambda - c} \gg 0,$$

то $L(\lambda)$ допускает факторизацию

(3.19) $$L(\lambda) = L_+(\lambda)(Z - \lambda I),$$

где $L_+(\lambda) = \sum\limits_{j=0}^{n-1} \lambda^j B_j$ обратим при $|\lambda - c| \leqq r$, спектр Z лежит в круге $|\lambda - c| < r$ и Z подобен самосопряженному оператору.

Доказательство. Из условия (3. 18), в силу теоремы 3, следует что имеет место каноническая факторизация

$$\frac{L(\lambda)}{\lambda-c}=A_+(\lambda)A_-(\lambda).$$

Как и в доказательстве теоремы 4, убеждаемся, что

$$A_+(\lambda)=\sum_{j=0}^{n-1}\lambda^j C_j,\quad A_-(\lambda)=I+\frac{X}{\lambda-c}\qquad (C_j, X\in\mathfrak{R}).$$

Полагая $B_j=-C_j$ $(j=0, 1, \ldots, n-1)$, $Z=cI-X$, получим равенство (3. 19). Для окончания доказательства осталось установить подобие оператора Z самосопряженному оператору. Это осуществляется так же, как и в доказательстве теоремы 4, с тем лишь отличием, что оператор G, симметризующий Z, следует определить равенством

$$G=\frac{1}{2\pi i}\int_\Gamma L^{-1}(\lambda)\,d\lambda$$

Теорема 9. *Пусть A — ограниченный самосопряженный оператор и $r>\|A\|$. Тогда для любого операторного многочлена $B(\lambda)$ с ограниченными самосопряженными коэффициентами, удовлетворяющего условию*

(3. 20) $$\|B(\lambda)\|<r-\|A\|\qquad (|\lambda|=r),$$

пучок $L(\lambda)=A-\lambda I+B(\lambda)$ допускает факторизацию

$$L(\lambda)=L_+(\lambda)(Z-\lambda I),$$

где $L_+(\lambda)$ — операторный многочлен, обратимый в круге $|\lambda|\leqq r$, $\sigma(Z)\subset(-r, r)$ и Z подобен самосопряженному оператору.

В самом деле,

$$\operatorname{Re}\frac{(L(\lambda)f,f)}{\lambda}\leqq-1+r^{-1}\big(\|A\|+\max_{|\lambda|=r}\|B(\lambda)\|\big)\qquad (\|f\|=1).$$

Следовательно, при условии (3. 20) Re $[L(\lambda)/\lambda]\ll 0$, и утверждение теоремы вытекает из теоремы 8.

Отметим в заключение, что для теорем 5 и 7—9 имеют место естественные аналоги следствий 1 и 2.

Цитированная литература

[1] М. В. Келдыш, О собственных значениях и собственных функциях некоторых классов несамосопряженных уравнений, *ДАН СССР*, **77** (1951), 11—14.

[2] М. Г. Крейн и Г. К. Лангер, О некоторых математических принципах линейной теории демпфированных колебаний континуумов, *Труды международного симпозиума по применению теории функций комплексного переменного в механике сплошной среды*, т. 2 (Москва, 1965), 283—322.

[3] Г. К. Лангер, Об инвариантных подпространствах линейных операторов, действующих в пространстве с индефинитной метрикой, *ДАН СССР*, **169** (1966), 12—15.

[4] H. Langer, Über stark gedämpfte Scharen im Hilbertraum, *J. Math. and Mech.*, **17** (1968), 685—706.

[5] H. Langer, Über eine Klasse nichtlinearer Eigenwertprobleme, *Acta Sci. Math.* (в печати).

[6] R. J. Duffin, A minimax theory for overdamped networks, *J. Rat. Mech. and Anal.*, **4** (1955), 221—233.

[7] R. Kühne, Minimaxprinzipe für stark gedämpfte Scharen, *Acta Sci. Math.*, **29** (1968), 39—68.

[8] А. И. Маркушевич, *Теория аналитических функций*. 1 (Москва, 1967).

[9] И. Ц. Гохберг и Ю. Лайтерер, О канонической факторизации непрерывных оператор-функций относительно окружности, *Функц. анализ и его приложения*, **6:1** (1972), 73—74.

[10] А. С. Маркус, Некоторые признаки полноты системы корневых векторов линейного оператора в банаховом пространстве, *Матем. сб.*, **70** (112) (1966), 526—561.

[11] Н. Данфорди и Дж. Т. Шварц, *Линейные операторы. Общая теория* (Москва, 1962).

[12] А. Ю. Левин, Элементарный признак вещественности корней целой функции с положительными коэффициентами, *Проблемы матем. анализа сложных систем*, вып. 2 (Воронеж, 1968), 72—77.

[13] Э. Беккенбах и Р. Беллман, *Неравенства* (Москва, 1965).

[14] Ю. Л. Шмульян, Дробно-линейные преобразования верхней операторной полуплоскости, *Известия ВУЗов, Математика*, № **1** (80) (1969), 97—105.

[15] В. И. Ломоносов, О равномерной поломжительности симметризатора, *Матем. исслед. Кишинев* (в печати).

[16] H. Langer, Über eine Klasse polynomiaber Scharen selbstadjungierter Operatoren im Hilbertraum, *J. Funct. Anal.*, **12:1** (1973), 13—29.

(Поступило 3/VII/1972)

Operator inequalities and related dilations

By W. MLAK in Krakow (Poland)

Dedicated to Professor Béla Szőkefalvi-Nagy on his 60th birthday

We deal in the present paper with inequalities $T(a)^*ZT(a)\leqq Z$ where $T(\cdot)$ is a semi-group of operators in a sense to be precised below and Z is a fixed positive operator. We show that to such inequalities there correspond a uniquely determined positive definite function. Now the dilation theory enters which makes it possible to give a more or less precise intrinsic characterization of several properties of involved operators Z and $T(a)$. The inequalities in question have been studied by direct methods in [2] and [8] for $T(\cdot)$ being a semi-group of powers of a fixed operator.

In all what follows we consider the complex Hilbert spaces with usual notation for inner products and norms. If S is such a space then $L(S)$ stands for the algebra of all linear bounded operators in S and I_S denotes the identity operator in S. To begin with we formulate the following lemma:

Lemma. *Let H be a Hilbert space. Suppose we are given a set A totally ordered by the relation "$\leqq$". Let $Z\in L(H)$ be a positive operator. Assume that the function $T(\cdot,\cdot)\colon A\times A\to L(H)$ satisfies the following conditions:*

(1) $T(a, a)=I_H$ *for* $a\in A$.

(2) $T(a, b)T(b, c)=T(a, c)$ *if* $c\leqq b\leqq a$.

(3) $T(a, b)$ *form a commutative family.*

Then, if

(4) $T(a, b)^* ZT(a, b)\leqq Z$ *for* $b\leqq a$

then the function

$$T(a,b)=\begin{cases} ZT(a,b) & \text{if } b\leqq a,\\ T(b,a)^*Z & \text{if } a\leqq b\end{cases}$$

is positive definite, i.e.,

$$\sum_{i,k} (T(a_i, a_k) f_i, f_k) \geqq 0$$

for every finite choice $a, \ldots, a_n \in A$, $f, \ldots, f_n \in H$.

The proof of the lemma may be performed exactly in the same way as that of Th. 2 of [5] by using Halperin's factoring method. It is also possible to apply directly Th. 2 [5] when using the semi-inner product $\langle f, g\rangle = (Zf, g)$ $(f, g \in H)$ (see comments after Theorem 2 below and [4]).

Suppose G is an additive subgroup of reals and let $G_+ = \{a \in G | a \geqq 0\}$. The semi-group $T(\cdot)$ on G_+ is a function $T(\cdot): G_+ \to L(H)$ such that $T(0) = I_H$ and $T(a+b) = T(a)T(b)$ for $a, b \in G_+$. Applying Lemma to the function $T(a, b) = T(a-b)$ $(a \geqq b \geqq 0,\ a, b \in G)$ we infer that if for $Z \in L(H)$, $Z \geqq O$

$$T(a)^* Z T(a) \leqq Z \quad \text{for} \quad a \in G_+ \tag{5}$$

then the function

$$T(a) = \begin{cases} ZT(a) & \text{if} \quad a \in G_+, \\ T(-a)^* Z & \text{if} \quad (-a) \in G_+ \end{cases}$$

is positive definite on G. By a suitable dilation theorem ([1], [7]) we get therefore a generalization of the celebrated theorem of Sz.-Nagy on unitary dilations of contractions:

Theorem 1. *Suppose the semi-group* $T(\cdot)$ *satisfies* (5). *Then there is a Hilbert space* K *and a unitary representation* $S(\cdot): G \to L(K)$ *and an operator* $R: H \to K$ *such that*

$$ZT(a) = R^* S(a) R \quad \text{for} \quad a \in G_+. \tag{6}$$

The space K, *the operator* R *and the unitary group are determined uniquely up to equivalence by the minimality condition* $K = \bigvee_{a \in G} S(a)RH$.

If the minimality condition holds true then $S(\cdot)$ is called the minimal Z-dilation of $T(\cdot)$ and (6) the canonical representation for $T(\cdot)$.

Assume now that (5) holds true and let $S(\cdot)$ be the minimal Z-dilation of $T(\cdot)$. We define

$$M_- = \bigvee_{a \in G_+} S(-a) RH, \quad S_+(a) = S(-a) | M_- \quad (a \in G_+). \tag{7}$$

If $f, g \in H$ then for $a \in G_+$, $(-b) \in G_+$ we have

$$\big(RT(a)f, S(b)Rg\big) = \big(ZT(a-b)f, g\big) = \big(S_+(a)^* Rf, S(b)Rg\big).$$

Since the vectors $S(b)Rg$ $\big((-b) \in G_+, g \in H\big)$ span M_-, we conclude that the following theorem holds true:

Theorem 2. *Suppose that the semi-group $T(\cdot)$ satisfies* (5). *Let $S(\cdot)$ be the minimal Z-dilation of $T(\cdot)$ and let M_- and $S_+(\cdot)$ be defined by* (7). *Then $R_-\colon H\to M_-$ defined by $R_-f=Rf$ for $f\in H$ satisfies the following conditions:*

$$R_-T(a) = S_+(a)^*R_- \quad \text{for} \quad a\in G_+. \tag{8}$$

$$Z = R_-^*R_-. \tag{9}$$

The above theorem includes as particular cases the Prop. 5. 1 of [8] p. 210 and Th. 5 of [2]. Notice that we do not require $T(\cdot)$ to be contractive.

The study of minimal Z-dilations may be reduced within certain limits to the study of ordinary dilations i.e. that ones for which $Z=I_H$. This is shown by arguments developed below, which, when suitably rearranged may stand for a direct proof of Theorem 1 without any appeal to Lemma. Suppose just that (5) holds true and let $S(\cdot)$ be the minimal Z-dilation of $T(\cdot)$.

Define $Q=\sqrt{Z}$, $H_1=\overline{R(Z)}=\overline{R(Q)}$. The relation $\tilde{T}(a)Qf=QT(a)f$ $(f\in H)$ determines a well defined semi-group $\tilde{T}(\cdot)$ of contractions in $L(H_1)$. It follows — see [4] — that $\tilde{T}(\cdot)$ has an ordinary minimal unitary dilation $U(a)$. Consequently $(U(a)Qf, Qg)=(\tilde{T}(a)Qf, Qg)=(ZT(a)f, g)=(S(a)Rf, Rg)$ for $a\in G_+$, $f, g\in H$; which implies that $U(\cdot)$ and $S(\cdot)$ are unitarily equivalent.

Suppose now that the operators $Z_1, Z_2\in L(H)$ are positive and

$$T(a)^*Z_iT(a) \leqq Z_i \quad \text{for} \quad i=1,2, \quad a\in G_+ \tag{10}$$

and the difference $\Delta Z = Z_2-Z_1 \geqq O$ also satisfies the inequality

$$T(a)^*\Delta Z T(a) \leqq \Delta Z \quad \text{for} \quad a\in G_+. \tag{11}$$

Let $Z_iT(a)=R_i^*S_i(a)R_i$ $(i=1, 2)$ be the canonical expression and K_i the minimal dilation space corresponding to Z_i. Following the arguments developed in [1], Lemma 4. 1 we conclude first from (11) that

$$\left\|\sum_{i|1}^{n} S_1(a_i)R_1f_i\right\|^2 \leqq \left\|\sum_{i|1}^{n} S_2(a_i)R_2f_i\right\|^2$$

for $a_i\in G$, $f_i\in H$. It follows that there is unique contraction $T\colon K_2\to K_1$ such that $TS_2(a)R_2f=S_1(a)R_1f$ for $a\in G$ and $f\in H$. Since the things are going about minimal dilations, the last equality yields that $TS_2(a)=S_1(a)T$ for all $a\in G$. We have just proved the following theorem:

Theorem 3. *Suppose that Z_1 and Z_2 satisfy* (10) *and* (11). *Then there exists a unique contraction $T\colon K_2\to K_1$ such that $TR_2=R_1$ and $TS_2(a)=S_1(a)T$ for all $a\in G$.*

Next we describe briefly some properties of polynomially bounded operators. We say that the operator $B \in L(H)$ is polynomially bounded if

$$\left\| \sum_{k|0}^{n} a_k B^k \right\| \leqq M \sup_{|z|=1} \left| \sum_{k|0}^{n} a_k z^k \right|$$

for every polynomial $\sum_{k|0}^{n} a_k z^k$ and with some finite M. If B is polynomially bounded then there are (so called elementary) measures $p(f, g)$ $(f, g \in H)$ on the unit circle C such that $\|p(f, g)\| \leqq M \|f\| \|g\|$ and

$$(12) \qquad (B^n f, g) = \int_C z^n dp(f, g) \qquad (n = 0, 1, 2 \ldots)$$

for all $f, g \in H$. This is an easy consequence of results of [6] that then $H = H_a + H_s$, $B = B_a + B_s$ (both sums direct), $B_a \in L(H_a)$, $B_s \in L(H_s)$ and B_a, B_s are polynomially bounded and such that

$$(B_a^n f, g) = \int z^n dp^a(f, g) \qquad (f, g \in H_a; \quad n = 0, 1, \ldots),$$

$$(B_s^n f, g) = \int z^n dp^s(f, g) \qquad (f, g \in H_s; \quad n = 0, 1, \ldots),$$

where the elementary measures p^a and p^s satisfy the conditions:

$$(13) \qquad p^a(f, g) \ll m \quad \text{for} \quad f, g \in H_a,$$

$$(14) \qquad p^s(f, g) \perp m \quad \text{for} \quad f, g \in H_s.$$

$$(15) \qquad p(f, g) = p(f_a, g_a) + p(f_s, g_s),$$

m stands here for the normalized Lebesgue measure on C and f_a, g_a and f_s, g_s stand for projections of f, g on H_a and H_s respectively. One can show that B_s is similar to a unitary operator with singular spectrum. If B is a contraction then the above decompositions are orthogonal and B_s is unitary and singular. If $B = B_a$ $(B = B_s)$ then we say that B is m-continuous (m-singular respectively). The decomposition $B = B_a + B_s$ is called the Lebesgue decomposition of B.

Suppose that $Z \geqq O$ and $T \in L(H)$ satisfy the inequality

$$(16) \qquad T^* Z T \leqq Z.$$

Then for $H_1 = \overline{R(Z)}$, $Q = \sqrt{Z}$ the formula $\tilde{T} Q f = Q T f$ $(f \in H)$ defines a contraction $\tilde{T} \in L(H_1)$. Let $H_1 = H_1^a \oplus H_1^s$, $\tilde{T} = \tilde{T}_a \oplus \tilde{T}_s$ be the corresponding Lebesgue decomposition of $\tilde{T}$. $\tilde{T}_s$ is unitary and singular. We now define Z_a, $Z_s \in L(H)$ by the formula

$$Z_a f = \hat{Q} P_a Q f, \quad Z_s f = \hat{Q} P_s Q f \qquad (f \in H),$$

where P_a and P_s are projections within H_1 on H_1^a and H_1^s respectively and $\tilde{Q}$ equals the restriction of Q to H_1. Since $P_a+P_s=I_{H_1}$ then

$$((Z_a+Z_s)f,f)=(Q(P_a+P_s)Qf,f)=(Zf,f)$$

for $f\in H$ i.e. $Z=Z_a+Z_s$. On the other hand $P_aQTf=\tilde{T}_aP_aQf$ for $f\in H$ and $\|\tilde{T}_a\|\leqq 1$ which implies that $\|P_aQTf\|^2\leqq\|P_aQf\|^2$ i.e. $T^*Z_aT\leqq Z_a$. By similar token, since T_s is unitary we get that $T^*Z_sT=Z_s$. Let $ZT^n=R^*V^nR$, $Z_aT^n=R_a^*V_{(a)}^nR_a$, $Z_sT^n=$ $=R_s^*V_{(s)}^nR_s$ $(n\geqq 0)$ be the canonical expressions for positive definite functions related to Z, Z_a, Z_s according to Theorem 1, V, $V_{(a)}$ and $V_{(s)}$ being the corresponding unitary operators. Let F be the semi-spectral measure of $\tilde{T}$ and $F=F^a+F^s$ its Lebesgue decomposition relative to m. Then for $f\in H$, $n\geqq 0$,

$$(Z_aT^n,f,f)=(QP_aQT^nf,f)=(\tilde{T}_a^nQf,Qf)=\int z^n d(F^aQf,Qf)=\int z^n d(E_{(a)}R_af,R_af)$$

and

$$(Z_sT^nf,f)=\int z^n d(F^sQf,Qf)=\int z^n d(E_{(s)}R_sf,R_sf)$$

where $E_{(a)}$ and $E_{(s)}$ stand for spectral measure of $V_{(a)}$ and $V_{(s)}$ respectively. Since the disc algebra is a Dirichlet one on C we infer that

$$(F^aQf,Qf)=(E_{(a)}R_af,R_af)\ll m,\quad (F^sQf,Qf)=(E_{(s)}R_sf,R_sf)\perp m.$$

Consequently $V_{(a)}$ has a Lebesgue spectrum and $V_{(s)}$ is singular. On the other hand $V=V_a\oplus V_s$ (Lebesgue decomposition relative to m) and

$$\int z^n d(E_aRf,Rf)+\int z^n d(E_sRf,Rf)=(ZT^nf,f)=((Z_a+Z_s)T^nf,f)=$$

$$=\int z^n d(E_{(a)}R_af,R_af)+\int z^n d(E_{(s)}R_sf,R_sf)$$

where E is the spectral measure of V, and $E=E_a\oplus E_s$ its Lebesgue decomposition. We conclude that for $f,g\in H$

$$(E_aRf,Rg)=(E_{(a)}R_af,R_af),\quad (E_sRf,Rg)=(E_{(s)}R_sf,R_sf)$$

which implies that for $n\geqq 0$

$$R_a^*V_{(a)}^nR_a=R^*V_a^nR=(P^aR)^*V_a^n(P^aR)$$

$$R_s^*V_{(s)}^nR_s=R^*V_s^nR=(P^sR)^*V_s^n(P^sR)$$

where $P^a=E_a(C)$, $P^s=E_s(C)$.

Summing up we get the following theorem:

Theorem 4. *Suppose T and Z satisfy* (16). *Then Z has a unique decomposition $Z=Z_a+Z_s$, $Z_a\geqq 0$, $Z_s\geqq 0$ where $T^*Z_aT\leqq Z_a$ $T^*Z_sT=Z_s$. The minimal Z_a (resp. Z_s) dilation of T is the m-continuous (resp. m-singular) part of the Z-dilation of T. Consequently, the minimal Z-dilation of T is an orthogonal sum of Z_a and Z_s dilations of T.*

Assume now that T which satisfies (16) is polynomially bounded. Let $p(f, g) = p^a(f, g) + p^s(f, g)$ be the Lebesgue decomposition (relative to m) of the elementary measure $p(f, g)$ of T. Using the previous notation for V, E_a, E_s we get for $n \geqq 0$, $f, g \in H$

$$(ZT^n f, g) = \int z^n d(E_a Rf, Rg) + \int z^n d(E_s Rf, Rg) = \\ = \int z^n dp^a(f, Zg) + \int z^n dp^s(f, Zg).$$

It follows now from the M. and F. Riesz theorem [3], Chapt. 4 that

$$\int z^n d(E_a Rf, Rg) = \int z^n dp^a(f, Zg), \quad (E_s Rf, Rg) \equiv p^s(f, Zg).$$

We conclude that for polynomially bounded T the following corollaries hold true:

Corollary 1. *If T is m-continuous then $Z = Z_a$ for every Z satisfying* (16), *i.e., every Z-dilation of T is m-continuous.*

Corollary 2. *If $Z = Z_a$ for T satisfying* (16) *then the range $R(Z)$ is included in the m-continuous part H_a of H of the Lebesgue decomposition related to T.*

Cor. 2 generalizes Cor. 5. 5 of [2]. Indeed, if $T^{*n} Z T^n \to O$ strongly then V is a bilateral shift with a complete wandering subspace equal to $\overline{(RT - VR)H}$. Consequently $V = V_a$. Notice that we infer Cor. 2 without using lifting of commutants.

References

[1] W. B. Arveson, Subalgebras of C*-algebras, *Acta Math.*, **123** (1969), 141—224.

[2] R. G. Douglas, On the operator equation $S^*XT = X$ and related topics, *Acta Sci. Math.*, **30** (1969), 19—32.

[3] K. Hoffman, *Banach spaces of analytic functions* (Englewood Cliffs, N. Y., 1962).

[4] W. Mlak, Unitary dilations in case of ordered groups, *Ann. Pol. Math.*, **17** (1965), 321—328.

[5] ——— Positive definite contraction valued functions, *Bull. Acad. Polon. Sci. (sér. sci. math., astr. et phys.)*, **15** (1967), 509—512.

[6] ——— Decompositions and extensions of operator valued representations of function algebras, *Acta Sci. Math.*, **30** (1969), 181—193.

[7] B. Sz.-Nagy, Extensions of linear transformations in Hilbert spaces which extend beyond this space, Appendix to F. Riesz and B. Sz.-Nagy: *Functional analysis* (New York. 1960).

[8] B. Sz.-Nagy and C. Foiaş, *Harmonic analysis of operators on Hilbert space* (Budapest—Amsterdam—London, 1970).

INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
SECTION KRAKÓW

(Received March 10, 1972)

Finslerräume von identischer Torsion

Von A. MOÓR in Sopron (Ungarn)

Herrn Professor B. Sz.-Nagy auf seinem 60. Geburtstag gewidmet

§ 1. Einleitung

Es seien $g_{ij}(x, \dot{x})$ und $g_{ij}^*(x, \dot{x})$ die metrischen Grundtensoren der n-dimensionalen Finslerräume F_n und F_n^*. Die Metrik soll im folgenden immer den in den Finslerräumen gewöhnlichen Bedingungen genügen (vgl. [1] I. § 1). Wir definieren die Finslerräume von identischer Torsion durch die folgende

Definition. Die Finslerräume F_n und F_n^* sind Finslerräume von identischer Torsion, falls

(1. 1) $$A_{ijk}^* = \frac{F^*}{F} A_{ijk}$$

gültig ist, wo

(1. 2) $$A_{ijk} \overset{\text{def}}{=} \frac{F}{2} \partial_{\dot{x}^k} g_{ij} \quad \text{bzw.} \quad A_{ijk}^* \overset{\text{def}}{=} \frac{F^*}{2} \partial_{\dot{x}^k} g_{ij}^*$$

den Torsionstensor und

$$F \overset{\text{def}}{=} \sqrt{g_{ij} \dot{x}^i \dot{x}^j} \quad \text{bzw.} \quad F^* \overset{\text{def}}{=} \sqrt{g_{ij}^* \dot{x}^i \dot{x}^j}$$

die Grundfunktion bedeuten.

Aus (1. 1) und (1. 2) folgt somit, daß in diesen Räumen

(1. 3) $$\partial_{\dot{x}^k} g_{ij}^* = \partial_{\dot{x}^k} g_{ij}$$

besteht und somit g_{ij}^* einer Relation von der Form:

(1. 4) $$g_{ij}^*(x, \dot{x}) = g_{ij}(x, \dot{x}) + \gamma_{ij}(x)$$

genügt, wo $\gamma_{ij}(x)$ einen Tensor bedeutet, der von der Richtung: $\dot{x}^i$ unabhängig ist. Offenbar folgt aus (1. 4) die Relation (1. 1), d.h. (1. 4) ist für die Finslerräume von identischer Torsion charakteristisch.

Im folgenden wollen wir statt des beliebigen, nur vom Orte: x^i abhängigen Tensors γ_{ij} einen Tensor von der Form:

$$\hat{\gamma}_{ij}(x, \dot{x}) = \tfrac{1}{2}(p_i q_j + p_j q_i)$$

voraussetzen, wo p_i und q_i im allgemeinen von $(x, \dot{x})$ abhängige kovariante Vektoren bedeuten. Die Forderung, daß der Tensor

$$\hat{g}_{ij} \equiv g_{ij} + \tfrac{1}{2}(p_i q_j + p_j q_i) \tag{1.5}$$

den Maßtensor eines Finslerschen Raumes $\hat{F}_n$ bestimme, führt auf die charakteristische Gleichung (2. 1), woraus im Falle

$$q_i(x, \dot{x}) = p_i(x, \dot{x}) \tag{1.6}$$

gefolgert werden kann, daß p_i nur vom Orte: x^i abhängig ist, d.h. die Finslerräume F_n und $\hat{F}_n$ sind in diesem Falle von identischer Torsion.

Im Paragraphen 3 wollen wir diejenigen Bedingungen bestimmen, die notwendig und hinreichend sind für die Übereinstimmung der geodätischen Linien von F_n und $\hat{F}_n$.

Letztens, im Paragraphen 4, untersuchen wir die $(n-1)$-dimensionalen Hyperflächen der Finslerräume F_n und $\hat{F}_n$.

§ 2. Charakteristische Relationen für die Finslerräume $\hat{F}_n$

Es seien F_n und $\hat{F}_n$ je ein n-dimensionaler metrischer Linienelementraum (vgl. [2] und [3]) mit den Maßtensoren $g_{ij}(x, \dot{x})$ und $\hat{g}_{ij}(x, \dot{x})$, die miteinander durch die Formel (1. 5) verbunden sind. $p_i(x, \dot{x})$ und $q_i(x, \dot{x})$ bedeuten in (1. 5) kovariante Vektoren. Wir beweisen den folgenden

Satz 1. *Notwendig und hinreichend dafür, daß der durch* (1. 5) *bestimmte Tensor $\hat{g}_{ij}$ eine Finslersche Metrik definiere, falls g_{ij} der Grundtensor eines Finslerraumes ist, ist die Relation:*

$$P_{\dot{x}^j} Q_{\dot{x}^i} + Q_{\dot{x}^j} P_{\dot{x}^i} + P_{\dot{x}^i \dot{x}^j} Q + Q_{\dot{x}^i \dot{x}^j} P = p_i q_j + p_j q_i, \tag{2.1}$$

wo

$$P \overset{\text{def}}{=} p_i(x, \dot{x}) \dot{x}^i, \quad Q \overset{\text{def}}{=} q_i(x, \dot{x}) \dot{x}^i \tag{2.1a}$$

bedeuten, und p_i und q_i in den $\dot{x}^k$ homogen von nullter Dimension sind.

Beweis. Erstens zeigen wir, daß (2. 1) gültig ist, falls g_{ij} und $\hat{g}_{ij}$ Finslersche Metriken bestimmen. Aus (1. 5) folgt nämlich nach einer Überschiebung mit $\dot{x}^i\dot{x}^j$:

$$\hat{F}^2 = F^2 + PQ, \tag{2.2}$$

wo

$$\hat{F}^2 \overset{\text{def}}{=} \hat{g}_{ij}\dot{x}^i\dot{x}^j, \quad F^2 \overset{\text{def}}{=} g_{ij}\dot{x}^i\dot{x}^j.$$

Da $\hat{F}$ und F die Grundfunktionen der entsprechenden Finslerräume $\hat{F}_n$ und F_n mit den Grundtensoren $\hat{g}_{ij}$ und g_{ij} sind, erhält man somit aus (2. 2) nach der Operation: $\frac{1}{2}\partial^2_{\dot{x}^i\dot{x}^j}$:

$$\hat{g}_{ij} = g_{ij} + \tfrac{1}{2}(P_{\dot{x}^i\dot{x}^j}Q + Q_{\dot{x}^i\dot{x}^j}P) + \tfrac{1}{2}(P_{\dot{x}^i}Q_{\dot{x}^j} + Q_{\dot{x}^i}P_{\dot{x}^j}), \tag{2.3}$$

woraus mittels (1. 5) die beweisende Relation (2. 1) unmittelbar folgt.

Nehmen wir jetzt an, daß die Relationen (2. 1) und (1. 5) gültig sind. Wir müssen zeigen, daß dann $\hat{g}_{ij}$ eine Relation von der Form:

$$\hat{g}_{ij} = \tfrac{1}{2}\partial^2_{\dot{x}^i\dot{x}^j}\hat{F}^2$$

genügt, d.h. $\hat{g}_{ij}$ ist der metrische Grundtensor des Finslerraumes $\hat{F}_n$ mit der Grundfunktion $\hat{F}$. Es wird sich zeigen, daß $\hat{F}(x, \dot{x})$ eben durch die Formel (2. 2) angegeben ist. Es folgt nach der Operation $\frac{1}{2}\partial^2_{\dot{x}^i\dot{x}^j}$ angewandt auf die Funktion $(F^2 + PQ) \equiv \hat{F}^2$:

$$\tfrac{1}{2}\partial^2_{\dot{x}^i\dot{x}^j}\hat{F}^2 = g_{ij} + \tfrac{1}{2}(P_{\dot{x}^i\dot{x}^j}Q + Q_{\dot{x}^i\dot{x}^j}P) + \tfrac{1}{2}(P_{\dot{x}^i}Q_{\dot{x}^j} + Q_{\dot{x}^i}P_{\dot{x}^j}).$$

Beachten wir jetzt die Relationen (2. 1) und dann (1. 5), so erhalten wir:

$$\tfrac{1}{2}\partial^2_{\dot{x}^i\dot{x}^j}\hat{F}^2 = \hat{g}_{ij}$$

w. z. b. w.

Mit Hilfe des fundamentalen Satzes 1, können mehrere Sätze über Finslerräume, die durch (1. 5) miteinander verbunden sind, bewiesen werden. Z. B.:

Satz 2. *Ist in der Relation* (1. 5) p_i *von der Form:*

$$p_i = P_{\dot{x}^i}, \quad P \overset{\text{def}}{=} p_j(x, \dot{x})\dot{x}^j \tag{2.4}$$

und ist p_i in den $\dot{x}^k$ homogen nullter Dimension, so ist auch

$$q_i = Q_{\dot{x}^i}, \quad Q \overset{\text{def}}{=} q_j(x, \dot{x})\dot{x}^j. \tag{2.5}$$

Beweis. Aus der Relation (1. 5) folgt nach dem Satz 1 die Relation (2. 1), ferner es muß auch $q_j(x, \dot{x})$ in den $\dot{x}^k$ homogen von nullter Dimension sein, falls g_{ij} und $\hat{g}_{ij}$ in (1. 5) Finslersche Metriken bestimmen. Aus (2. 1) folgt nach einer Überschiebung mit $\dot{x}^j$ auf Grund der Homogenität in den $\dot{x}^k$:

$$PQ_{\dot{x}^i} + QP_{\dot{x}^i} = Pq_i + Qp_i,$$

woraus wegen der Bedingung (2. 4) die Relation (2. 5) unmittelbar folgt, w. z. b. w.

Satz 3. *Gilt in der Relation* (1. 5):

$$p_i(x, \dot{x}) = q_i(x, \dot{x}), \tag{2. 6}$$

so ist p_i von $\dot{x}^k$ unabhängig, falls g_{ij} und $\hat{g}_{ij}$ Finslersche Metriken bestimmen.

Beweis. Aus (1. 5) folgt nach unserem Satz 1 die Relation (2. 1), die jetzt wegen der Bedingung (2. 6) die Form

$$P_{\dot{x}^i} P_{\dot{x}^j} + P_{\dot{x}^i \dot{x}^j} P = p_i p_j \tag{2. 7}$$

haben wird. Da g_{ij} und $\hat{g}_{ij}$ Finslersche Metriken bestimmen, sind sie in den $\dot{x}^k$ homogen von nullter Dimension, folglich ist auch $p_i(x, \dot{x})$ in den $\dot{x}^k$ homogen von nullter Dimension. Dann folgt aber aus unserer letzten Gleichung, nach einer Überschiebung mit $\dot{x}^j$ und in Hinsicht auf (2. 1a):

$$P_{\dot{x}^i \dot{x}^j} \dot{x}^j = 0, \quad p_i = P_{\dot{x}^i}.$$

Die Relation (2. 7) reduziert sich somit auf $P_{\dot{x}^i \dot{x}^j} = 0$, woraus $p_i = P_{\dot{x}^i} = p_i^*(x^1, x^2, \ldots, x^n)$ folgt, w. z. b. w.

Satz 4. *Ist in* (1. 5) *p_i nur vom Orte x^k abhängig, so ist q_i auch nur vom Orte x^k abhängig und von $\dot{x}^k$ unabhängig.*

Beweis. Nach der Annahme des Satzes folgt

$$p_i = p_i(x) = P_{\dot{x}^i}, \quad P \equiv p_k(x) \dot{x}^k.$$

Nach dem Satz 2 ist somit

$$q_i = Q_{\dot{x}^i}, \quad Q = q_k(x, \dot{x}) \dot{x}^k.$$

Aus (1. 5) folgt nun wegen $p_{i \dot{x}^k} = 0$ nach einer partiellen Ableitung nach $\dot{x}^k$:

$$\partial_{\dot{x}^k} \hat{g}_{ij} = \partial_{\dot{x}^k} g_{ij} + \tfrac{1}{2}(p_i Q_{\dot{x}^j \dot{x}^k} + p_j Q_{\dot{x}^i \dot{x}^k}). \tag{2. 8}$$

In den Finslerräumen gilt nun bekanntlich

$$(\partial_{\dot{x}^k} \hat{g}_{ij}) \dot{x}^i \equiv 0, \quad (\partial_{\dot{x}^k} g_{ij}) \dot{x}^i \equiv 0;$$

somit bekommt man aus (2. 8) nach einer Überschiebung mit $\dot{x}^i$ in Hinsicht auf die Homogenität nullter Dimension von $q_i(x, \dot{x})$ in den $\dot{x}^i$: d.h. $Q_{\dot{x}^i \dot{x}^j} \dot{x}^i = 0$,

$$p_i(x) \dot{x}^i Q_{\dot{x}^j \dot{x}^k} = 0,$$

woraus $Q_{\dot{x}^j \dot{x}^k} = 0$ folgt, und das beweist schon den Satz.

Die beiden letzten Sätze bestimmen also solche Räume F_n und $\hat{F}_n$, die nach unserer Definition im Paragraphen 1. von identischer Torsion sind. Der Zusammenhang der metrischen Grundtensoren ist durch die Relation (1. 5) festgelegt, die sich im Falle $p_i = q_i$ auf

$$\hat{g}_{ij}(x, \dot{x}) = g_{ij}(x, \dot{x}) + p_i(x) p_j(x)$$

reduziert. Selbstverständlich kann im allgemeinen Fall, d.h. beim Typ (1. 5) p_i bzw. q_i von den $\dot{x}^k$ abhängig sein. Z. B. in trivialer Weise ist (2. 1) erfüllt, falls p_i und q_i die Form:

$$p_i = f_i(x)\,h(\dot{x}), \quad q_i = f_i^*(x)\,h^{-1}(\dot{x})$$

haben. F_n und $\hat{F}_n$ sind in diesem Falle von identischer Torsion.

Die Relation (2. 1) war notwendig und hinreichend dafür, daß $\hat{F}_n$ mit dem Maßtensor $\hat{g}_{ij}$ ein Finslerraum sei, falls g_{ij} der Maßtensor eines solchen Raumes war. Wir geben statt (2. 1) noch eine andere hinreichende Bedingung dafür, daß $\hat{F}_n$ ein Finslerraum sei.

Satz 5. *Ist in der Formel* (1. 5) g_{ij} *der Grundtensor eines Finslerraumes, so ist:*

(2. 9) $$p_k Q_{\dot{x}^i} + q_k P_{\dot{x}^i} + p_{k\dot{x}^i} Q + q_{k\dot{x}^i} P = p_i q_k + p_k q_i$$

— *wo P und Q durch* (2. 1a) *festgelegt sind, ferner p_i und q_i in den $\dot{x}^k$ homogen von nullter Dimension sind — hinreichend dafür, daß $\hat{g}_{ik}$ eine Finslersche Metrik bestimme.*

Beweis. Wir überschieben die Bedingungsgleichung (2. 9) mit $\dot{x}^k$. Wegen der Homogenität in den $\dot{x}^j$ wird:

$$PQ_{\dot{x}^i} + QP_{\dot{x}^i} + (P_{\dot{x}^i} - p_i)\,Q + (Q_{\dot{x}^i} - q_i)\,P = p_i Q + q_i P,$$

die in der From $p_k Q + q_k P = PQ_{\dot{x}^k} + P_{\dot{x}^k} Q$ geschrieben werden kann, wo wir statt des Index „i" den Index „k" gesetzt haben. Differenzieren wir nun die letzte Gleichung partiell nach $\dot{x}^i$, beachten wir die Bedingungsgleichung (2. 9), so wird:

$$p_i q_k + p_k q_i = P_{\dot{x}^i\dot{x}^k} Q + Q_{\dot{x}^i\dot{x}^k} P + P_{\dot{x}^i} Q_{\dot{x}^k} + P_{\dot{x}^k} Q_{\dot{x}^i}.$$

Offenbar ist diese Gleichung mit (2. 1) identisch, somit folgt nach Satz 1 auch der Satz 5.

Wir gehen jetzt zur Bestimmung der kontravarianten Komponenten des Maßtensors $\hat{g}_{ik}$ über. Kürze halber bezeichnen wir im folgenden durch h_{ij} den Ausdruck:

$$h_{ij} \overset{\text{def}}{=} \tfrac{1}{2}(p_i q_j + p_j q_i).$$

Es sei p^{kj} die Abweichung des Tensors $\hat{g}^{kj}$ von g^{kj}. Es ist somit:

(2. 10) $$\hat{g}^{kj} = g^{kj} + p^{kj}.$$

Offenbar ist p^{kj} in ihren Indizes symmetrisch, ferner es gilt nach der gewöhnlichen Definition des kontravarianten Maßtensors:

$$(g_{ij} + h_{ij})(g^{kj} + p^{kj}) = \delta_i^j = \begin{cases} 1, & \text{wenn} \quad i = j, \\ 0, & \text{wenn} \quad i \neq j. \end{cases}$$

Beachten wir jetzt, daß auch $g_{ij}g^{kj}=\delta_i^k$ ist, so bekommen wir für p^{kj} die Gleichung:

$$(g_{ij}+h_{ij})p^{kj} = -h_i^k \equiv -h_{ij}g^{kj}.$$

Angenommen, daß Det $|g_{ij}+h_{ij}| \neq 0$, wird

$$(g_{ij}+h_{ij})Q^{im} = \delta_j^m$$

für Q^{im} eindeutig lösbar. Mit Hilfe von Q^{im} wird dann offensichtlich $p^{km}=-h_i^k Q^{im}$, womit nach (2. 10) der Tensor $\hat{g}^{kj}$ vollständig bestimmt ist.

§ 3. Übereinstimmende geodätische Linien

In diesem Paragraphen wollen wir solche Bedingungen bestimmen, die notwendig und hinreichend sind dafür, daß eine geodätische Linie von F_n auch bezüglich der Metrik von $\hat{F}_n$ eine geodätische Linie sei.

Die geodätischen Linien genügen der Differentialgleichung

$$\text{(3. 1)} \qquad \frac{\ddot{x}^i+2G^i(x,\dot{x})}{\dot{x}^i} = \frac{\ddot{x}^k+2G^k(x,\dot{x})}{\dot{x}^k} \qquad (i,k=1,2,\dots,n),$$

wo

$$G^h \overset{\text{def}}{=} \tfrac{1}{4} g^{hj}(\partial_k g_{ij}+\partial_i g_{jk}-\partial_j g_{ik})\dot{x}^i\dot{x}^k$$

bedeutet (vgl. [1], IV. (8. 2), III. (1. 26) und II. (2. 3)). Offenbar ist (3. 1) eine parameterinvariante Form der geodätischen Linien. Bezeichnen $\hat{G}^i$ $(i=1,2,\dots,n)$ die die geodätischen Linien bestimmenden Funktionen des Finslerraumes $\hat{F}_n$, ferner sind diese Linien auch für F_n geodätisch, so hängen $\hat{G}^h$ und G^h durch die Formeln

$$\text{(3. 2)} \qquad \hat{G}^h = G^h + \dot{x}^h \psi(x,\dot{x})$$

zusammen, wo $\psi(x,\dot{x})$ einen Skalar bedeutet der in den $\dot{x}^k$ homogen von erster Dimension ist.

Wir berechnen jetzt $\hat{G}^i$ falls (1. 5) gilt. Nach (1. 5) und (2. 10) ist:

$$\text{(3. 3)} \qquad \hat{G}^h \equiv \hat{g}^{hj}\hat{G}_j = (g^{hj}+p^{hj})(G_j+[ijk]\dot{x}^i\dot{x}^k),$$

wo

$$G_j \overset{\text{def}}{=} \tfrac{1}{2}(\partial_k g_{ij}-\tfrac{1}{2}\partial_j g_{ik})\dot{x}^i\dot{x}^k, \quad \hat{G}_j \overset{\text{def}}{=} \tfrac{1}{2}(\partial_k \hat{g}_{ij}-\tfrac{1}{2}\partial_j \hat{g}_{ik})\dot{x}^i\dot{x}^k$$

$$[ijk] \overset{\text{def}}{=} \tfrac{1}{4}(q_j\partial_k p_i+p_i\partial_k q_j+q_i\partial_k p_j+p_j\partial_k q_i-q_k\partial_j p_i-p_i\partial_j q_k)$$

bedeuten. Schreiben wir nun (3. 3) in der Form:

$$\text{(3. 4)} \qquad \hat{G}^h = G^h+p^{hj}G_j+(g^{hj}+p^{hj})[ijk]\dot{x}^i\dot{x}^k,$$

so haben wir eine fundamentale Formel erhalten, die den Zusammenhang von $\hat{G}^h$

und G^h bestimmt, falls die metrischen Grundtensoren von F_n und $\hat{F}_n$ miteinander durch (1. 5) verbunden sind.

Substituieren wir nun $\hat{G}^h$ aus (3. 2) in (3. 4), so wird nach einer Überschiebung mit $g_{hs}\dot{x}^s$:

$$\psi = \frac{1}{F^2}\{p^{hj}g_{hs}\dot{x}^s G_j + (\dot{x}^j + p^{hj}g_{hs}\dot{x}^s)[ijk]\dot{x}^i\dot{x}^k\}. \tag{3. 5}$$

Setzen wir das in (3. 2), ferner beachten wir (3. 4), so erhält man die Differentialgleichung:

$$p^{hj}G_j + (g^{hj}+p^{hj})[ijk]\dot{x}^i\dot{x}^k = \frac{\dot{x}^h}{F^2}\{p^{tj}g_{ts}\dot{x}^s G_j + (\dot{x}^j + p^{tj}g_{ts}\dot{x}^s)[ijk]\dot{x}^i\dot{x}^k\}, \tag{3. 6}$$

die offenbar notwendig ist, falls $x^i(t)$ sowohl für F_n wie auch für $\hat{F}_n$ geodätische Linie sein soll.

Die Relation (3. 6) ist aber auch hinreichend dafür, daß wenn $x^i(t)$ eine geodätische Linie von F_n ist, dann auch eine solche von $\hat{F}_n$ sei. Nach (3. 4) und (3. 6) folgt nämlich, daß $\hat{G}^h$ die Form (3. 2) hat, wo ψ durch (3. 5) angegeben ist. Genügt also die Kurve $x^i(t)$ der Gleichung (3. 1), so genügt sie auch der analogen Gleichung mit $\hat{G}^i$.

Unsere Resultate können wir im folgenden Satz zusammenfassen:

Satz 6. *Die Gleichung* (3. 6) *ist notwendig und hinreichend dafür, daß die geodätische Linie von F_n auch geodätische Linie von $\hat{F}_n$ sei.*

§ 4. Hyperflächen bezüglich der Metriken g_{ij} und $\hat{g}_{ij}$

Eine Hyperfläche F_{n-1} eines Finslerraumes F_n ist eine $(n-1)$-dimensionale Mannigfaltigkeit der Linienelemente $(u^\alpha, \dot{u}^\alpha)$,[1] deren Raumkomponenten durch die Gleichungen

$$x^i = x^i(u^1, u^2, \ldots, u^{n-1})$$

$$\dot{x}^i = B^i_\alpha \dot{u}^\alpha, \quad B^i_\alpha \overset{\text{def}}{=} \frac{\partial x^i}{\partial u^\alpha}$$

festgelegt sind.

Der metrische Grundtensor von F_{n-1} ist

$$a_{\alpha\beta} = g_{ij}\big(x^k(u), B^k_\gamma \dot{u}^\gamma\big)B^i_\alpha B^j_\beta;$$

---

[1]) Griechische Indizes bedeuten im folgenden immer die Zahlen 1, 2, ..., $(n-1)$.

auf Grund von (1. 5) wird:

$$\hat{a}_{\alpha\beta} = \hat{g}_{ij} B^i_\alpha B^j_\beta = a_{\alpha\beta} + \tfrac{1}{2}(p_\alpha q_\beta + p_\beta q_\alpha),$$

wo $p_\alpha = p_i B^i_\alpha$ bzw. $q_\alpha = q_i B^i_\alpha$ die Projektion von p_i bzw. q_i bedeutet.

Die Normalvektoren bezüglich der Metriken g_{ij} und $\hat{g}_{ij}$ sind durch die Gleichungen

$$\left.\begin{aligned} g_{ij} N^i B^j_\alpha &= 0 \\ g_{ij} N^i N^j &= 1 \end{aligned}\right\} \quad \left.\begin{aligned} \hat{g}_{ij} \hat{N}^i B^j_\alpha &= 0 \\ \hat{g}_{ij} \hat{N}^i \hat{N}^j &= 1 \end{aligned}\right\}$$

fesgelegt.

Wir beweisen den folgenden

Satz 7. *Stehen p_i und q_i zu den Vektoren B^i_α ($\alpha = 1, 2, \ldots, m-1$) normal, so sind $\hat{N}^i$ und N^i bis auf einen skalaren Faktor identisch.*

Beweis. In einem Linienelement $(u^\alpha, \dot{u}^\alpha)$ von F_{n-1} gilt nach der Annahme:

$$p_i B^i_\alpha = 0, \quad q_i B^i_\alpha = 0, \tag{4.1}$$

woraus folgt, daß

$$\hat{g}_{ij} B^i_\alpha \hat{N}^j = \left(g_{ij} + \tfrac{1}{2}(p_i q_j + p_j q_i)\right) B^i_\alpha \hat{N}^j = g_{ij} B^i_\alpha \hat{N}^j \tag{4.2}$$

ist. Da N^i und $\hat{N}^i$ Normalvektoren bezüglich der entsprechenden Metriken sind, ist

$$g_{ij} B^i_\alpha N^j = 0, \quad \hat{g}_{ij} B^i_\alpha \hat{N}^j = 0. \tag{4.3}$$

Auf Grund von (4. 2) ist somit $g_{ij} B^i_\alpha \hat{N}^j = 0$, die mit (4. 3) zusammen beweist, daß

$$\hat{N}^j = \lambda N^j \tag{4.4}$$

ist, wo λ einen Skalar bedeutet. Unsere letzte Formel drückt aber schon den Satz aus.

Bemerkung. Die Behauptung des Satzes 7, genau gesagt: die Formel (4. 4) gilt auch dann, wenn statt (4. 1) die schwächere Bedingungen

$$(p_i q_j + p_j q_i) B^i_\alpha \hat{N}^j = 0 \qquad (\alpha = 1, 2, \ldots, n-1) \tag{4.5}$$

gelten. Offenbar ist auch jetzt (4. 2) gültig, woraus (4. 4) — wie vorher — abgeleitet werden kann. —

Nun wollen wir die geodätischen Linien der Hyperfläche F_{n-1} charakterisieren und das Analogon des im Paragraphen 3 behandelten Problems untersuchen.

Die Gleichung der geodätischen Linien einer Hyperfläche hat die Form:

$$\frac{d^2 u^\alpha}{ds^2} + 2G^\alpha\left(u, \frac{du}{ds}\right) = 0, \tag{4.6}$$

wo der Parameter s die Bogenlänge bedeutet und

$$(4.7)\qquad 2G^\alpha(u,\dot u) = B_i^\alpha(B^i_{\gamma\beta}+\Gamma^{*i}_{hk}B^h_\gamma B^k_\beta)\dot u^\beta \dot u^\gamma \equiv 2G^i B_i^\alpha + B_i^\alpha B^i_{\beta\gamma}\dot u^\beta \dot u^\gamma,\quad B^i_{\beta\gamma} \overset{\text{def}}{=} \frac{\partial^2 x^i}{\partial u^\beta \partial u^\gamma}$$

ist (vgl. [1] Kap. V. § 3.)[2]). In parameterinvarianter Form lautet bekanntlich (4. 6) wie folgt:

$$(4.8)\qquad \big(\ddot u^\alpha + 2G^\alpha(u,\dot u)\big)\dot u^\beta = \big(\ddot u^\beta + 2G^\beta(u,\dot u)\big)\dot u^\alpha.$$

Da für die Raumkoordinaten

$$\dot x^i = B^i_\beta \dot u^\beta,\quad \ddot x^i = B^i_{\beta\gamma}\dot u^\beta \dot u^\gamma + B^i_\beta \ddot u^\beta$$

besteht, wird man nach einer Überschiebung mit B_i^α, auf Grund der Relation:

$$B^i_\beta B^\alpha_i = \delta^\alpha_\beta,$$

die folgende Formel bekommen:

$$(4.9)\qquad \ddot u^\alpha = B^\alpha_i \ddot x^i - B^\alpha_i B^i_{\beta\gamma}\dot u^\beta \dot u^\gamma.$$

Setzen wir G^α und $\ddot u^\alpha$ aus (4. 7) und (4. 9) in die Gleichung (4. 8), so wird nach einer kleinen Umformung:

$$(4.10)\qquad (\ddot x^i + 2G^i)(B^\alpha_i \dot u^\beta - B^\beta_i \dot u^\alpha) = 0$$

eben die Gleichung der flächengeodätischen Linien sein.

Auf Grund von (3. 4) ist nun

$$G^h = \hat G^h - p^{hj}G_j - (g^{hj}+p^{hj})[ijk]\dot x^i \dot x^k.$$

Substituiert man das in die Gleichung (4. 10), so wird:

$$(4.11)\qquad \{\ddot x^i + 2\hat G^i - 2p^{ij}G_j - 2(g^{ij}+p^{ij})[tjk]\dot x^t \dot x^k\}(B^\alpha_i \dot u^\beta - B^\beta_i \dot u^\alpha) = 0.$$

Da die Gleichung der flächengeodätischen Linien bezüglich der Metrik $\hat g_{ij}$ analog zur Formel (4. 10) die Form:

$$(\ddot x^i + 2\hat G^i)(B^\alpha_i \dot u^\beta - B^\beta_i \dot u^\alpha) = 0$$

hat, folgt aus (4. 11) unmittelbar der

Satz 8. *Die notwendigen und hinreichenden Bedingungen dafür, daß die geodätische Linie $u^\alpha(t)$ einer Hyperfläche F_{n-1} auch bezüglich der Metrik $\hat g_{ij}$ geodätische Linie sei, sind die Relationen:*

$$(4.12)\qquad \{p^{ij}G_j + (g^{ij}+p^{ij})[tjk]\dot x^t \dot x^k\}(B^\alpha_i \dot u^\beta - B^\beta_i \dot u^\alpha) = 0.$$

---

[2]) In der Formel V. (3.14) von [1] ist selbstverständlich: $B^h_\beta \dot u^\beta = \dot x^h$, $\Gamma^{*i}_{hk}\dot x^h \dot x^k = 2\,G^i$.

Zum Schluß bemerken wir noch, daß die Gleichung (4. 12) erfüllt ist, falls die Relation

$$p^{ij}G_j+(g^{ij}+p^{ij})[tjk]\dot{x}^t\dot{x}^k=\dot{x}^i\psi(x,\dot{x})$$

besteht, wo $\psi(x,\dot{x})$ einen Skalar bedeutet, da nach unserer Annahme über $\dot{x}^i$, die Formel

$$B_i^\beta\dot{x}^i=\dot{u}^\beta$$

gültig ist.

Literatur

[1] HANNO RUND, *The differential geometry of Finsler spaces*, Springer-Verlag (Berlin—Göttingen—Heidelberg, 1959).

[2] A. MOÓR, Entwicklung einer Geometrie der allgemeinen metrischen Linienelementräume, *Acta Sci. Math.*, **17** (1956), 85—120.

[3] A. MOÓR, Eine Verallgemeinerung der metrischen Übertragung in allgemeinen metrischen Räumen, *Publ. Math. Debrecen*, **10** (1963), 145—150.

(Eingegangen am 9. Februar, 1972)

The order of magnitude of the Lebesgue functions and summability of function series

By FERENC MÓRICZ in Szeged.*)

Dedicated to Professor Béla Sz.-Nagy on his 60th birthday

1. Let X be a measurable space with a positive measure μ and let $F=\{f_k(x)\}$ $(k=0, 1, \ldots)$ be a sequence of L_μ-integrable functions on the set $E(\subset X)$ of positive measure. We shall consider the series

$$\sum_{k=0}^{\infty} c_k f_k(x) \tag{1}$$

with real coefficients satisfying

$$\sum_{k=0}^{\infty} c_k^2 < \infty \tag{2}$$

Let $T=(\alpha_{nk})$ $(n, k=0, 1, \ldots)$ be a doubly infinite matrix of real numbers determining a general summation process with the aid of the linear means

$$t_n(x) = \sum_{k=0}^{\infty} \alpha_{nk} c_k f_k(x).$$

We say that the series (1) is *T-summable* at the point $x(\in X)$ if the series defining $t_n(x)$ converges in the ordinary sense for all n (except perhaps finitely many of them) and the limit $\lim_{n\to\infty} t_n(x)$ exists at the point x in question.

Form the *Lebesgue functions* belonging to the sequence F of functions and to the summation process T as follows:

$$L_n(T, F; x) = \int_E |K_n(T, F; x, t)|\, d\mu(t),$$

where

$$K_n(T, F; x, t) = \sum_{k=0}^{\infty} \alpha_{nk} f_k(x) f_k(t).$$

*) This paper was written while the author stayed at the Steklov Mathematical Institute in Moscow.

To avoid the unnecessary complications concerning the existence (in a certain sense) of $t_n(x)$ and $L_n(T, F; x)$, we shall consider the following two particular cases of summation processes T:

(i) If the functions $f_k(x)$ are assumed to be only L_μ-integrable on E, we shall confine ourselves to matrices T that have only finitely many nonzero elements in each row, i.e., which are such that $\alpha_{nk}=0$ for $k>k_n$ $(n=0, 1, \ldots)$.

(ii) If F is an orthonormal system defined on a set E of finite measure, then we shall only consider matrices T satisfying the condition

$$\sum_{k=0}^{\infty} \alpha_{nk}^2 < \infty \qquad (n = 0, 1, \ldots).$$

In this case, from (2) and this condition it immediately follows that $\sum_{k=0}^{\infty} \alpha_{nk}^2 c_k^2 < \infty$, and so we have by the Riesz—Fischer theorem that $t_n(x)$ is L_μ^2-integrable on E for every n. Furthermore, by virtue of

$$\sum_{k=0}^{\infty} \alpha_{nk}^2 \int_E f_k^2(x)\, d\mu(x) = \sum_{k=0}^{\infty} \alpha_{nk}^2 < \infty \qquad (n = 0, 1, \ldots),$$

and by B. Levi's theorem we can conclude, that $\sum_{k=0}^{\infty} \alpha_{nk}^2 f_k^2(x) < \infty$ for almost every x in E, and consequently $K_n(T, F; x, t)$ is L_μ^2-integrable on E as a function of t for almost every x in E and for every n. This implies, in particular, the existence of $L_n(T, F; x)$ for almost every x in E and for every n.

2. The order of magnitude of the Lebesgue functions may, in many cases, be a decisive factor in convergence problems.

In particular, taking

$$\alpha_{nk} = 1 - \frac{k}{n+1} \qquad (k = 0, 1, \ldots, n), \quad \alpha_{nk} = 0$$

$$(k = n+1, n+2, \ldots) \quad (n = 0, 1, \ldots),$$

we obtain the classical $(C, 1)$-summation process. Now we have

$$L_n((C, 1), F; x) = \int_E \left| \sum_{k=0}^{n} \left(1 - \frac{k}{n+1}\right) f_k(x) f_k(t) \right| d\mu(t).$$

In this case G. Alexits and A. Sharma [1] have proved the following theorems:

A. *Let F be a sequence of L_μ-integrable functions on a measurable set E of finite measure and let $\{\mu_n\}$ be a non-decreasing sequence of positive numbers. If $\sum c_k^2 < \infty$ and the condition $L_n((C, 1), F; x) = O(\mu_n)$ is uniformly satisfied on E, then the sums*

$$\sigma_n(x) = \sum_{k=0}^{n} \left(1 - \frac{k}{n+1}\right) c_k f_k(x)$$

have the order of magnitude $O_x(\sqrt{\mu_n})$ on E almost everywhere.

B. *Let F be a sequence of L_μ-integrable functions on a measurable set E of finite measure satisfying the condition*

$$\int_E \left| \sum_{k=0}^{n} c_k d_k f_k(x) \right| d\mu(x) = O(1) \qquad (n = 0, 1, \ldots)$$

whenever $\sum c_k^2 d_k^2 < \infty$, and let $\{\mu_n\}$ be a non-decreasing sequence of positive numbers that is concave from below. Suppose that $L_n((C, 1), F; x) = O(\mu_n)$ for every $x \in E$ and

$$\sum_{k=0}^{\infty} c_k^2 \mu_k < \infty. \tag{3}$$

Then the series (1) *is* $(C, 1)$*-summable on E almost everywhere.*

They also remark that these results remain valid for any (C, α)-summation $(\alpha > 0)$ if we replace $L_n((C, 1), F; x)$ by the corresponding Lebesgue functions $L_n((C, \alpha), F; x)$.

We note that the above theorem for orthonormal systems is a well-known theorem of S. KACZMARZ [2]. G. SUNOUCHI [3] and L. LEINDLER [4] have extended Kaczmarz's theorem to the Riesz summation of orthogonal series. In this case

$$\alpha_{nk} = 1 - \frac{\lambda_k}{\lambda_{n+1}} \qquad (k = 0, 1, \ldots, n),$$

$$\alpha_{nk} = 0 \qquad (k = n+1, n+2, \ldots) \quad (n = 0, 1, \ldots),$$

and

$$L_n(R, F; x) = \int_E \left| \sum_{k=0}^{n} \left(1 - \frac{\lambda_k}{\lambda_{n+1}}\right) f_k(x) f_k(t) \right| d\mu(t),$$

where $\{\lambda_n\}$ is a strictly increasing sequence of numbers with $\lambda_0 = 0$ and $\lambda_n \to \infty$ as $n \to \infty$.

To our knowledge, no analogous theorem for other summation processes has yet been proved. The following problem can be quite naturally raised: If for a summation process T and for an orthonormal system F defined on E the condition

$$L_n(T, F; x) = O(\mu_n)$$

is uniformly satisfied on E, is then the series (1), under condition (3), summable with respect to the concerning process almost everywhere in E?

A. V. EFIMOV [5] in the case of $\mu_n \to \infty$ $(n \to \infty)$, and K. TANDORI and the present author in a joint paper [6] in general, have essentially showed that the answer to this question is in the negative.

The aim of the present paper is to give a positive answer to the above question for a relatively large class of summation processes.

3. In the sequel we shall consider summation processes T with the following property: the estimate

$$\left|\sum_{k=0}^{\infty} \alpha_{nk}\alpha_{mk} f_k(x) f_k(y)\right| \leqq \sum_{i=0}^{\min(m,n)} \beta_i |K_i(T, F; x, y)| \tag{4}$$

holds for every m and n, where the positive numbers $\beta_i = \beta_i(T, F; \min(m, n))$ satisfy the inequalities

$$\sum_{i=0}^{\min(m,n)} \beta_i = O(1) \qquad (m, n = 0, 1, \ldots).$$

We note that if F is an orthonormal system defined on E, then the estimate (4) can be written in a more natural form as follows:

$$\left|\int_E K_n(T, F; x, t) K_m(T, F; y, t)\, d\mu(t)\right| = \left|\sum_{k=0}^{\infty} \alpha_{nk}\alpha_{mk} f_k(x) f_k(y)\right| \leqq$$

$$\leqq \sum_{i=0}^{\min(m,n)} \beta_i |K_i(T, F; x, y)|,$$

where the β_i's have the properties mentioned above.

We show that if T is the Riesz summation process defined by the sequence $\{\lambda_n\}$, then condition (4) is satisfied. (See also [4]). Supposing $n<m$, we obtain with the aid of the Abel transform that

$$\frac{1}{\lambda_{n+1}\lambda_{m+1}} \sum_{k=0}^{n} (\lambda_{n+1}-\lambda_k)(\lambda_{m+1}-\lambda_k) f_k(x) f_k(y) =$$

$$= \frac{\lambda_{m+1}-\lambda_{n+1}}{\lambda_{n+1}\lambda_{m+1}} \sum_{k=0}^{n} (\lambda_{n+1}-\lambda_k) f_k(x) f_k(y) + \frac{1}{\lambda_{n+1}\lambda_{m+1}} \sum_{k=0}^{n} (\lambda_{n+1}-\lambda_k)^2 f_k(x) f_k(y) =$$

$$= \frac{\lambda_{m+1}-\lambda_{n+1}}{\lambda_{m+1}} K_n(R, F; x, y) +$$

$$+ \frac{1}{\lambda_{n+1}\lambda_{m+1}} \sum_{k=0}^{n} (2\lambda_{n+1}-\lambda_k-\lambda_{k+1})(\lambda_{k+1}-\lambda_k) \sum_{l=0}^{k} f_l(x) f_l(y).$$

Substituting here $\lambda_{k+1} K_k(R, F; x, y) - \lambda_k K_{k-1}(R, F; x, y)$ for $(\lambda_{k+1}-\lambda_k) \sum_{l=0}^{k} f_l(x) f_l(y)$

$(k=0, 1, \ldots, n)$, a repeated Abel transform gives that the right-hand side can be written as

$$\frac{\lambda_{m+1}-\lambda_{n+1}}{\lambda_{m+1}} K_n(R,F;x,y)+\frac{1}{\lambda_{n+1}\lambda_{m+1}} \sum_{k=0}^{n-1} (\lambda_{k+2}-\lambda_k)\,\lambda_{k+1} K_k(R,F;x,y)+$$

$$+\frac{\lambda_{n+1}-\lambda_n}{\lambda_{m+1}} K_n(R,F;x,y) = \frac{\lambda_{m+1}-\lambda_n}{\lambda_{m+1}} K_n(R,F;x,y)+$$

$$+\frac{1}{\lambda_{n+1}\lambda_{m+1}} \sum_{k=0}^{n-1} (\lambda_{k+2}-\lambda_k)\,\lambda_{k+1} K_k(R,F;x,y).$$

Therefore, the estimate (4) holds for every n and m with $\beta_k = (\lambda_{k+2}-\lambda_k)/\lambda_{n+1} \geqq \geqq (\lambda_{k+2}-\lambda_k)\lambda_{k+1}/\lambda_{n+1}\lambda_{m+1}$ $(k = 0, 1, \ldots, n-1)$ and $\beta_n = 1 \geqq (\lambda_{m+1}-\lambda_{n+1})/\lambda_{m+1}$, for which we have

$$\sum_{k=0}^{n} \beta_k = \frac{\lambda_{n+1}-\lambda_0}{\lambda_{n+1}}+1 \leqq 2 \qquad (n = 0, 1, \ldots).$$

4. After these preliminaries our first result can be formulated as follows:

Theorem 1. *Let $F=\{f_k(x)\}$ be a sequence of L_μ-integrable functions on a measurable set E of finite measure, let c_k be a sequence of coefficients satisfying* (2), *and assume that the summation process T satisfies condition* (4). *If $\{\mu_n\}$ is a non-decreasing sequence of positive numbers for which the relation*

$$L_n(T,F;x) = O(\mu_n) \tag{5}$$

uniformly holds on E, then the estimate

$$t_n(x) = O_x(\sqrt{\mu_n})$$

holds almost everywhere in E.

The proof is a modification of the well-known proof of A. Kolmogoroff—G. Seliverstoff [7] and A. Plessner [8] for the trigonometric system, and of S. Kaczmarz [2] for arbitrary orthonormal systems.

We shall use an idea of C. J. Preston [9] which consists in a special representation of $t_n(x)$. Introduce an arbitrary orthonormal system $\{g_k(y)\}$ defined on a measure space Y with positive measure ν; then

$$t_n(x) = \int_Y \sum_{k=0}^{\infty} c_k g_k(t)\cdot \sum_{k=0}^{\infty} \alpha_{nk} f_k(x) g_k(t)\, d\nu(t).$$

Let $n(x)$ be the smallest index $\leqq n$ such that

$$\frac{t_{n(x)}(x)}{\sqrt{\mu_{n(x)}}} = \max_{0\leqq k\leqq n} \frac{t_k(x)}{\sqrt{\mu_k}}$$

holds. By Schwarz's inequality we have

$$I_n = \left| \int_E \frac{t_{n(x)}(x)}{\sqrt{\mu_{n(x)}}} d\mu(x) \right| \leqq \left\{ \int_Y \left[\sum_{k=0}^{\infty} c_k g_k(t) \right]^2 d\nu(t) \times \right.$$

$$\left. \times \int_Y \left[\int_E \frac{1}{\sqrt{\mu_{n(x)}}} \sum_{k=0}^{\infty} \alpha_{n(x),k} f_k(x) g_k(t) d\mu(x) \right]^2 d\nu(t) \right\}^{\frac{1}{2}} \leqq$$

$$\leqq \left\{ \sum_{k=0}^{\infty} c_k^2 \right\}^{\frac{1}{2}} \left\{ \int_E \int_E \int_Y \frac{1}{\sqrt{\mu_{n(x)}}} \sum_{k=0}^{\infty} \alpha_{n(x),k} f_k(x) g_k(t) \times \right.$$

$$\left. \times \frac{1}{\sqrt{\mu_{n(y)}}} \sum_{k=0}^{\infty} \alpha_{n(y),k} f_k(y) g_k(t) d\mu(x) d\mu(y) d\nu(t) \right\}^{\frac{1}{2}} =$$

$$= O(1) \left\{ \int_E \int_E \frac{1}{\sqrt{\mu_{n(x)} \mu_{n(y)}}} \left| \sum_{k=0}^{\infty} \alpha_{n(x),k} \alpha_{n(y),k} f_k(x) f_k(y) \right| d\mu(x) d\mu(y) \right\}^{\frac{1}{2}}.$$

Using estimate (4) and the monotonity of $\{\mu_n\}$ we have

$$I_n = O(1) \left\{ \int_E \int_E \frac{1}{\sqrt{\mu_{n(x)} \mu_{n(y)}}} \sum_{i=0}^{\min(n(x), n(y))} \beta_i |K_i(T, F; x, y)| \, d\mu(x) d\mu(y) \right\}^{\frac{1}{2}} =$$

$$= O(1) \left\{ \int_E \int_E \frac{1}{\mu_{n(x)}} \sum_{i=0}^{n(x)} \beta_i |K_i(T, F; x, y)| \, d\mu(x) d\mu(y) + \right.$$

$$\left. + \int_E \int_E \frac{1}{\mu_{n(y)}} \sum_{i=0}^{n(y)} \beta_i |K_i(T, F; x, y)| \, d\mu(x) d\mu(y) \right\}.$$

The validity of relation (5) on E implies the estimate

$$I_n = O(1) \left\{ \int_E \frac{1}{\mu_{n(x)}} \sum_{i=0}^{n(x)} \beta_i L_i(T, F; x) d\mu(x) + \int_E \frac{1}{\mu_{n(y)}} \sum_{i=0}^{n(y)} \beta_i L_i(T, F; y) d\mu(y) \right\}^{\frac{1}{2}} =$$

$$= O(1) \left\{ \int_E \sum_{i=0}^{n(x)} \beta_i d\mu(x) + \int_E \sum_{i=0}^{n(y)} \beta_i d\mu(y) \right\}^{\frac{1}{2}} = O(1).$$

Since the sequence $\{t_{n(x)}(x)/\sqrt{\mu_{n(x)}}\}$ is increasing, it follows by B. Levi's theorem that

$$\frac{t_{n(x)}(x)}{\sqrt{\mu_{n(x)}}} < \infty$$

almost everywhere in E. The same is true for the sequence $\{-t_{n(x)}(x)/\sqrt{\mu_{n(x)}}\}$; hence

$$\frac{t_{n(x)}(x)}{\sqrt{\mu_{n(x)}}} = O_x(1)$$

almost everywhere, which implies our statement.

5. We need the following auxiliary result:

Lemma. *Let $\{\mu_n\}$ be a non-decreasing sequence of positive numbers. Let F be a sequence of L_μ-integrable functions on the set E and let $\{n_k\}$ be an increasing sequence of indices such that*

(6) $$\int_E |s_n(x)|\,d\mu(x) = O(1)$$

and

(7) $$s_{n_k}(x) = O_x(\sqrt{\mu_{n_k}})$$

holds almost everywhere in E for every sequence of coefficients satisfying (2), *where $s_n(x)$ is the nth partial sum of the series* (1). *Then condition* (3) *implies the convergence of the partial sums $s_{n_k}(x)$ almost everywhere in E as $k\to\infty$.*

This lemma is contained in the cited paper of G. Alexits and A. Sharma [1]. (See there Theorem 3.) We remark that (6) is trivially satisfied for orthonormal systems defined on a set E of finite measure.

In the sequel we suppose that the sequence F and the summation process T are such that there exists an increasing sequence n_k of indices for which the conditions

$$\text{(i)}\quad s_{n_k}(x) - t_{n_k}(x) = o_x(1) \quad\text{and}\quad \text{(ii)}\quad \max_{n_k<n\leqq n_{k+1}} |t_n(x) - t_{n_k}(x)| = o_x(1)$$

hold almost everywhere in E as $k\to\infty$ for every sequence of coefficients satisfying (2).

In particular, if F is an orthonormal system defined on a set E of finite measure and T is the Riesz summation process defined by $\{\lambda_n\}$, then the conditions (i) and (ii) are fulfilled by every sequence $\{n_k\}$ of indices for which

$$1 < q \leqq \frac{\lambda_{n_{k+1}}}{\lambda_{n_k}} \leqq r < \infty \quad (k = 1, 2, \ldots).\quad \text{(See A. Zygmund [10].)}$$

6. Now we are in a position to formulate our second result:

Theorem 2. *Suppose the sequence F of L_μ-integrable functions and the summation process T are such that there exists an increasing sequence $\{n_k\}$ of indices satisfying* (i) *and* (ii) *and such that condition* (6) *is also satisfied. If the inequality*

$$t_n(x) = O_x(\sqrt{\mu_n})$$

holds almost everywhere in E for every sequence of coefficients satisfying (2), *where* $\{\mu_n\}$ *is a non-decreasing sequence of positive numbers, then condition* (3) *implies the T-summability of the series* (1) *almost everywhere in E.*

In fact, by (i) we have that the inequality (7) holds almost everywhere in E for every sequence of coefficients satisfying (2). Applying our Lemma we get that condition (3) implies the convergence of the partial sums $\{s_{n_k}(x)\}$ almost everywhere in E. Using (i) and (ii) we obtain that, under (3), $\{t_n(x)\}$ converges almost everywhere in E, which entails our assertion.

7. Remarks. (i) It is clear that our Theorem 1 contains Theorem A as a special case. In particular, Theorem A remains valid for any Riesz summation process if we replace $L_n((C, 1), F; x)$ with the corresponding Lebesgue functions $L_n(R, F; x)$.

As for the orthonormal systems F, the results of G. SUNOUCHI [3] and L. LEINDLER [4] on the Riesz-summability of orthogonal series are also special cases of our Theorems 1 and 2.

(ii) We mention without proofs that Theorems 1 and 2 can be extended for other particular summation processes T such as, e.g., the de la Vallée Poussin summation, the Euler summation, etc.

References

[1] G. ALEXITS and A. SHARMA, The influence of Lebesgue functions on the convergence and summability of function series, *Acta Sci. Math.*, **33** (1972), 1—10.

[2] S. KACZMARZ, Sur la convergence et la sommabilité des développements orthogonaux, *Studia Math.*, **1** (1929), 87—121.

[3] G. SUNOUCHI, On the Riesz summability of Fourier series, *Tôhoku Math. J.*, second series, **11** (1959), 319—326.

[4] L. LEINDLER, Über die sehr starke Riesz-Summierbarkeit der Orthogonalreihen und Konvergenz lückenhafter Orthogonalreihen, *Acta Math. Acad. Sci. Hung.*, **13** (1962), 401—414.

[5] А. В. Ефимов, О несуммируемых линейными методами ортогональных рядах, *ДАН СССР*, **152** (1963), 31—34.

[6] F. MÓRICZ and K. TANDORI, On a problem of summability of orthogonal series, *Acta Sci. Math.*, **29** (1968), 331—350.

[7] A. KOLMOGOROFF—G. SELIVERSTOFF, Sur la convergence des séries de Fourier, *Rendiconti Acad. Lincei*, **3** (1926), 307—310.

[8] A. PLESSNER, Über Konvergenz von trigonometrischen Reihen, *J. reine angew. Math.*, **155** (1926), 15—25.

[9] C. J. PRESTON, On the convergence of multiplicatively orthogonal series, *Proc. Amer. Math. Soc.*, **28** (1971), 453—455.

[10] A. ZYGMUND, Sur la sommation des séries de fonctions orthogonales, *Bulletin Intern. Acad. Polonaise Sci. Lettres Cracovie*, série A, 1927, 293—308.

(Received August 28, 1972)

Über die Metrisierung der affinen Geometrie

Von ROLF NEVANLINNA in Helsinki (Finnland)

Prof. Béla Szőkefalvi-Nagy zum 60. Geburtstag gewidmet

§ 1. Affine Geometrie

Im folgenden beschränken wir uns auf die Frage der Einführung der Metrik in der *zweidimensionalen affinen Ebene.*

Die affine Geometrie kann auf Grund der folgenden Axiome aufgebaut werden.

1. 1. *Axiome der Inzidenz.* Gegeben seien zwei Mengen: der *Punkte* (P) und der *Geraden* (L). Jedem Punkt P ist eine Menge von mit P inzidenten Geraden (L_P) zugeordnet. Die Inzidenz von P und L bezeichnen wir: $P-L$. (oder $L-P$). Die Inzidenz soll folgenden Axiomen genügen:

I. 1. Zu jedem Punkt gibt es mindestens zwei inzidente Geraden und zu jeder Geraden mindestens zwei inzidente Punkte.

I. 2. Zu zwei Punkten gibt es genau eine inzidente Gerade.

I. 3. (Parallelenaxiom). Zu jedem Punkt P, der mit einer Geraden L_1 nichtinzident ist, gibt es genau eine Gerade L_2-P, so daß L_1 und L_2 keinen gemeinsamen inzidenten Punkt besitzen (L_1 und L_2 sind parallel: $L_1 \| L_2$).

I. 4. (Satz von Desargues). Wenn die Punkttripel A_1, B_1, C_1 und A_2, B_2, C_2 den Bedingungen $A_1A_2 \| B_1B_2 \| C_1C_2$, $A_1B_1 \| A_2B_2$, $B_1C_1 \| B_2C_2$ genügen, so gilt auch $A_1C_1 \| A_2C_2$.

1. 2. *Axiome der Anordnung.* Als zweite Grundrelation des Systems (P), (L) führt man die *Anordnung* ein: *jedem Punktepaar A, B ist eine wohlbestimmte Untermenge* (C) von der Menge (P) zugeordnet. Man sagt, C liegt zwischen A und B, und bezeichnet dies: ACB (oder BCA).

Die Anordnung wird durch nachstehende Axiome geregelt:

II. 1. Wenn ABC, so sind die Punkte A, B, C kollinear (es gibt eine mit ihnen inzidente Gerade).

II. 2. Von drei kollinearen Punkten liegt genau ein Punkt zwischen den zwei übrigen.

II. 3. Zu dem Punktepaar A, B gibt es mindestens drei weitere Punkte C, D, E, so daß ACB, ABD, EAB.

II. 4. (Axiom von Pasch.) Wenn die Gerade L mit den drei Punkten A, B, C nichtinzident ist, und ist $L-P$, APB, so gibt es einen Punkt $Q-L$, so daß AQC oder BQC gilt.

II. 5. (Stetigkeitsaxiom.) Wenn die Menge aller Punkte (P), $P-L$, in zwei punktfremde Klassen (A) und (B) eingeteilt sind, so daß, kein Punkt A zwischen zwei B-Punkten liegt und kein Punkt B zwischen zwei A-Punkten liegt, so gibt es entweder in (A) oder in (B) einen *äußersten* Punkt C, der weder zwischen zwei A-Punkten noch zwischen zwei B-Punkten liegt.

Das affine System (I, II) genügt den logischen Grundforderungen der Unabhängigkeit, der Widerspruchsfreiheit und der Vollständigkeit.

§ 2. Einführung der Metrik

2. 1. *Kongruenz paralleler Vektoren.* Das Parallelenaxiom und der Desargues'sche Satz ermöglichen die eindeutige parallele Verschiebung eines Vektors von der Anfangslage A_0B_0 in die Endlage A_nB_n, längs eines Polygonzuges $A_0A_1\ldots A_n$. Erklärt man zwei parallel verschobene Vektoren bzw. die entsprechenden nicht orientierten Strecken) als kongruent, so bilden sie eine Äquivalenzklasse.

2. 2. *Kongruenz nichtparalleler Strecken. Die Eichlinie.* Unter Beachtung des Permanenzprinzips sucht man den Kongruenzbegriff für beliebige (nichtparallele) Vektoren und Strecken zu erweitern. Zu diesem Zweck verschieben wir die Vektoren parallel so, daß sie einen beliebig festgesetzen Punkt O als Anfangspunkt erhalten. Man betrachte dann eine Punktmenge E_0, mit folgenden Eigenschaften:

1°. E_0 ist in Bezug auf O symmetrisch.

2°. E_0 ist in Bezug auf O sternförmig (jede Halbstrahl, der von 0 ausgeht, trifft E_0 in genau einem Punkt).

Die Wahl einer solchen *Eichlinie* („Kreislinie") E_0 setzt den Kongruenzbegriff der Strecken eindeutig fest: 1) Alle „Radien" OX ($X\in E_0$) werden als kongruent erklärt; 2) Die affinen Axiome definieren eine wohlbestimmte Menge von weiteren (punktfremden) Eichlinien E, die mit E_0 in Bezug auf O homotetisch sind. Die Radien einer Linie E werden als kongruent definiert.

Damit ist der Kongruenzbegriff für beliebige Vektoren (Strecken) als eine Äquivalenz erklärt.

Falls die Eichlinie E_0 *konvex* ist, so gelangt man zu der von Minkowski eingeführten Metrik, bei der die Maßzahlen der Strecken der *Dreiecksungleichung* genügen[1]).

---

[1]) Die Erweiterung der Minkowskischen Geometrie zu unendlichdimensionalen Räumen ergibt die Banach-Geometrie.

2. 3. *Winkelkongruenz.* Zwei Winkel W und W', deren Schenkel paarweise *parallel* sind, werden als kongruent definiert. Sucht man diese Erklärung auf beliebige Winkel zu erweitern, unter Beachtung des Permanenzprinzips, so liegt es nahe die Winkelkongruenz auf die Streckenkongruenz zurückzuführen.

Hierzu bemerke man, daß wenn die Winkel W und W' (mit den Scheitelpunkten O bzw. O') *parallel* sind, folgendes gilt:

Trägt man auf den Schenkeln L_1 und L_2 von W zwei beliebige Vektoren OA_1 und OA_2 ab, und auf den Schenkeln L_1' und L_2' von W' zwei zu jenen Vektoren kongruente Vektoren $O'A_1'(=OA_1)$ und $O'A_2'(=OA_2)$ ab, so gilt, daß die Vektoren A_1A_2 und $A_1'A_2'$ kongruent sind, und umgekehrt.

2. 4. *Erweiterung der Definition der Winkelkongruenz.* Diese letzte Eigenschaft läßt sich als erweiterte Erklärung der Kongruenz von zwei *beliebigen* Winkeln W und W' verwenden: Sei $O'A_1'=OA_1$ (wobei $=$ das Zeichen für die Kongruenzrelation ist) und $O'A_2'=OA_2$; falls dann auch die Strecken A_1A_2 und $A_1'A_2'$ kongruent sind, so heißen die Winkel W und W' kongruent.

2. 5. *Eindeutigkeit der Definition der Winkelkongruenz.* Im Falle paralleler Winkel ist diese definierende Eigenschaft von der Wahl der Punkte A_1 und A_2 unabhängig. Damit die aufgestellte Erklärung der Kongruenz nichtparalleler Winkel eindeutig (ebenfalls von der Wahl der Hilfspunkte A_1, A_2 unabhängig) ist, muß folgendes gefordert werden:

Postulat III: Es seien die vier Punkte O, A_1, A_2, A_3 so gegeben, daß O, A_2, A_3 kollinear sind. Sei ferner O', A_1', A_2', A_3' eine zweite Gruppe von vier Punkten (O', A_2', A_3' kollinear). Wenn dann $OA_1=O'A_1'$, $OA_2=O'A_2'$, $OA_3=O'A_3'$, $A_1A_2=A_1'A_2'$ gilt, so soll auch $A_1A_3=A_1'A_3'$ sein.

Bei beliebiger Wahl der symmetrischen, sternförmigen Eichlinie E ist dieses Postulat i. A. nicht erfüllt. In der Tat schränkt das Bestehen des für die Eindeutigkeit des Begriffs der Winkelkongruenz *notwendigen* Postulats III die Wahl der Linie E stark ein, wie im folgenden kurz gezeigt werden soll.

2. 6. *Konstruktion der Eichlinie mittels des Postulats III.* Wir zeigen zunächst, daß das Postulat III erlaubt, zu *drei* gegebenen Punkten A, B, C auf einer Eichlinie E einen neuen, vierten Punkt C_1 von E zu konstruieren. Wählt man OA und OB als aufspannende Einheitsvektoren eines affinen Koordinatensystems, so erhalten die drei gegebenen Punkte A, B, C die Koordinaten $(x=1, y=0)$, $(x=0, y=1)$ und $C=C_0(x_0, y_0)$. Mittels des Postulats III bestimmt man dann einen vierten Punkt $C_1(x_1, y_1)$ mit den Koordinaten

$$x_1 = \frac{x_0^2-1}{y_0}, \qquad y_1 = x_0,$$

der *ebenfalls auf der Eichlinie E liegt.*

Anderseits bestimmen die Punkte O, A, B, C_0 eine einzige quadratische Form

$$Q(x,y) = x^2 - 2cxy + y^2 \qquad \left(c = \frac{x_0^2 + y_0^2 - 1}{2x_0 y_0}\right),$$

die an den Punkten A, B, C_0 den Wert 1 annimmt. Die Koordinaten $(x_1 y_1)$ von C_1 lassen sich schreiben:

$$x_1 = 2cx_0 - y_0, \qquad y_1 = x_0, \tag{2.6}$$

und diese affine Transformation erhält die quadratische Form Q invariant.

2.7. *Euklidische Metrik.* Wir nehmen jetzt an, daß die vorgegebenen drei Punkte A, B, C der Eichlinie E eine in Bezug auf den Mittelpunkt O *konvexe* Figur bilden; d.h. daß die Punkte O und $C(=C_0)$ auf *verschiedenen* Seiten der Geraden AB liegen. Die quadratische Form Q ist dann *positiv definit* ($|c|<1$).

Durch Iteration der Transformation (2.6) bestimmt man eine Folge von Punkten $C_n(x_n, y_n)$, die alle sowohl auf der Eichlinie E als auf der Ellipse $Q=1$ liegen. Nimmt man dann noch an, daß die Eichlinie *beschränkt* ist, so ergibt das Iterationsverfahren:

1) *Die Eichlinie E ist stetig,* und daraus ferner:

2) *E fällt mit der Ellipse $Q=1$ zusammen.*

Damit ist gezeigt, daß die *Vorgabe von drei Punkten A, B, C einer Eichlinie E, die in Bezug auf den Mittelpunkt O eine konvexe Konfiguration bilden,* alle Eichlinien (mit O als Mittelpunkt) eindeutig festlegt, als eine Schar von (in Bezug auf O) homotetischen Ellipsen.

Die Streckenmetrik ist eindeutig bestimmt, und zwar als eine *euklidische.*

Das skalare Produkt wird durch Polarisation aus Q erhalten. Die euklidische Winkelmessung ergibt sich dann in üblicher Weise.

2.8. *Der indefinite und der semidefinite Fall.* Falls die Konfiguration A, B, C in Bezug auf O *konkav* ist, so führt die Anwendung des Postulats III (sofern man die Forderung der Beschränktheit der Eichlinie fallen läßt) zu dem Ergebnis, daß E mit den *Hyperbeln* $Q(x,y) = \pm 1$ ($|C|>1$) zusammenfällt. Die Metrik ist also *indefinit*; die Richtungen der Asymptoten verbleiben unmetriziert.

Sind schließlich die drei Eichpunkte A, B, C *kollinear*, so ist $Q(x,y)$ *semidefinit*, und die Anwendung des Postulats III zeigt, daß die Eichlinie E mit dem *Geradenpaar* $Q=\pm 1$ zusammenfällt.

Die Transformationen (2.6), die im konvexen Fall die euklidischen Drehungen darstellen, können im indefiniten Fall als die Lorentztransformationen, im semidefiniten Fall als die Galileitransformationen gedeutet werden.

Die vollständige Ausführung der Beweise der Ergebnisse von 2.6 bis 2.8 soll in einem anderen Zusammenhang gegeben werden.

(Eingegangen am 31. Oktober 1972)

On quasi-equivalence of matrices over H^∞

By ERIC A. NORDGREN in Durham (New Hampshire, U.S.A.)

Dedicated to Professzor Béla Szőkefalvi-Nagy on his 60th birthday

Introduction

The purpose of this paper is to introduce a relation of quasi-equivalence for matrices over H^∞ that generalizes the relation of equivalence for matrices over principal ideal domains (cf. [3], p. 79) and leads to an analogous theory. In [6] Sz.-Nagy and Foiaş began a study of a class C_0 of Hilbert space contractions that possess a minimal function analogous to the minimal polynomial of finite matrices. This study was continued in [7] where it was shown that the minimal function of a C_0 contraction T of finite defect bears the same relation to the characteristic operator function Θ_T of T that the minimal polynomial of a finite matrix A bears to the polynomial matrix $A-\lambda$. In this paper an equivalence theory is developed which will be used in a subsequent paper [4] to show that the invariant factors of Θ_T determine the Jordan model of T, which was introduced by Sz.-Nagy and Foiaş in [8]. Thus the analogy between such contractions and finite matrices is complete.

1. Preliminaries

We will be concerned with matrices over the Hardy class H^∞ of bounded analytic functions on the unit disc, and a few of the pertinent facts will be set forth here. See [2] or [9] for details. Since H^∞ is an integral domain, the usual terminology for factorization applies. In particular, for a, b in H^∞, a is said to divide b if there exists c in H^∞ such that $ac=b$, in which case we will write $a|b$. According to Fatou's theorem, every H^∞ function has a radial limit at almost every point of the unit circle, and if these radial limits have modulus one almost everywhere, then the function is called inner. Every H^∞ function $f\not\equiv 0$ can moreover be factored into an inner function f_i and a function f_e having only constant inner divisors. We will require

[1]) Research supported in part by a grant from the National Science Foundation.

that the first nonvanishing Taylor coefficient of f_i be positive, and in this case the factorization is unique. The inner part of f can be further factored into a Blaschke product b_f, determined by the zeros of f, and a singular inner function s_f, determined by a measure ν_f on the unit circle that is singular with respect to Lebesgue measure. An inner function g divides f if and only if every zero of g is a zero of f, counting multiplicity, and $\nu_g \leqq \nu_f$.

Every subset Φ of H^∞ has a greatest common divisor $\wedge \Phi$, i.e. an inner function that divides every member of Φ and is in turn divisible by every other inner function with this property. We will require that $\wedge \Phi = (\wedge \Phi)_i$, thus insuring uniqueness. For pairs f, g, we will write $f \wedge g$ simply in place of $\wedge \{f, g\}$. A subset Φ of H^∞ is *relatively prime* if $\wedge \Phi = 1$. If Φ is any subset of H^∞, then let $\Phi^k = \{f^k : f \in \Phi\}$.

Lemma 1. 1. *If Φ is relatively prime, then so is Φ^k for every positive integer k. If Ψ is also relatively prime and if*

$$\Phi\Psi = \{\varphi\psi : \varphi \in \Phi, \psi \in \Psi\},$$

then $\Phi\Psi$ is relatively prime.

Proof. The members of Φ have no common zero, and the same is true of Φ^k. Thus no nonconstant Blaschke product divides each member of Φ^k. If ν is any nonzero singular measure, then there is an f in Φ such that $\frac{1}{k}\nu$ is not dominated by ν_f. Consequently, if $g = f^k$, then g^k is in Φ^k and ν is not dominated by $k\nu_f = \nu_g$. Thus no nonconstant singular inner function divides every member of Φ^k, and hence Φ^k is relatively prime.

If g is inner and if $g | \varphi\psi$ for all φ in Φ and ψ in Ψ, then $g | \psi$ since Φ is relatively prime, and it follows that g is constant since Ψ is relatively prime.

A notion of length can be attached to elements of a principal ideal domain, and this idea can be used to show that any matrix over such a ring can be reduced to a diagonal one by a finite number of equivalence transformations. For H^∞ a different route to diagonalization is available because of the possibility of forming a Lebesgue decomposition of one inner function with respect to another. If f and g are inner functions, if every zero of f is a zero of g, and if $\nu_f \ll \nu_g$, then we will write $f \ll g$. On the other hand, if f and g have no common zeros and if $\nu_f \perp \nu_g$, then we will write $f \perp g$. It is easy to see that $f \perp g$ if and only if $f \wedge g = 1$. Suppose f and g are arbitrary inner functions. Then $f = f_a f_s$, where $f_a \ll g$ and $f_s \perp g$. For let $\nu_f = \nu_a + \nu_s$ be the Lebesgue decomposition of ν_f with respect to ν_g. Let each zero of f that is a zero of g be a zero of f_a, and let ν_a determine the singular inner factor of f_a. Let the remaining zeros of f be zeros of f_s, and let ν_s determine the singular part of f_s. Then the desired factorization of f results. The essential lemma for the diagonalization later is the following.

Lemma 1. 2. *If a and b are relatively prime H^∞ functions, and if ω and ψ are arbitrary inner functions, then there exists H^∞ functions x and y such that $y\wedge\omega = 1$ and $(ax+by)\wedge\psi = 1$.*

Proof.[2]) Let $\omega=\omega_1\omega_2\omega_3$, where $\omega_1\ll a_i$, $\omega_2\ll b_i$, and $\omega_3\perp a_i b_i$. Setting $a'=a\omega_3$ we have $a'_i\perp b_i$. Now factor ψ in the form $\psi=\psi_1\psi_2\psi_3$, where $\psi_1\ll a'_i$, $\psi_2\ll b_i$, and $\psi_3\perp a'_i b_i(=a_i b_i\omega_3)$; hence $\psi_3\perp\omega$. Set $x' = \psi_1\psi_2+\psi_3$, $y=\psi_3$, and $\delta = a'x'+by$. Clearly $y\wedge\omega = 1$. We shall also show that $\delta\wedge\psi = 1$.

Consider to this effect any inner divisor φ of $\delta\wedge\psi$. Since $\varphi|\psi$, we have $\varphi=\varphi_1\varphi_2\varphi_3$, where $\varphi_k = \varphi\wedge\psi_k$ $(k=1, 2, 3)$. Set $\hat\varphi_1 = \varphi_1\wedge a'_i$, $\hat\varphi_2 = \varphi_2\wedge b_i$, and observe that φ_1 is constant if $\hat\varphi_1$ is so, and similarly for φ_2 and $\hat\varphi_2$. Since $\hat\varphi_1|\delta$ and $\hat\varphi_1|a'$, we have $\hat\varphi_1|by$, and as $a'\perp b$ we deduce that $\hat\varphi_1|y$, i.e. $\hat\varphi_1|\psi_3$. But $\hat\varphi_1|\varphi_1|\psi_1$ and $\psi_1\perp\psi_3$ so $\hat\varphi_1$ — and therefore φ_1 — are constant. Similarly, from $\hat\varphi_2|\delta$ and $\hat\varphi_2|b$ we deduce that $\hat\varphi_2|a'x'$ and as $a'\perp b$ we conclude that $\hat\varphi_2|x'(=\psi_1\psi_2+\psi_3)$. But $\hat\varphi_2|\varphi_2|\psi_2$ and $\psi_2\perp\psi_3$ so $\hat\varphi_2$ — and therefore φ_2 — are constant. Thus $\varphi(=\varphi_1\varphi_2\varphi_3)|\psi_3$. On the other hand, we have $\varphi|\delta(=a'\psi_1\psi_2+a'\psi_3+b\psi_3)$, and hence $\varphi|a'\psi_1\psi_2$. As the factors of the last product are prime to ψ_3, φ is constant. This proves that $\delta\wedge\psi = 1$. To obtain x as required by the lemma, we only have to set $x=\omega_3 x'$.

2. Definition and elementary properties of quasi-equivalence

If A and B are $m\times n$ matrices over H^∞, then equivalence of A and B is defined by requiring the existence of units X and Y of orders m and n respectively such that $XA=BY$. Here a unit X of order m is an $m\times m$ matrix over H^∞ for which there exists another such matrix Z such that $XZ=ZX=I_m$, where I_m is the $m\times m$ identity matrix. Since only weak* closed ideals in H^∞ are principal [5], this is not the appropriate relation to study if one hopes to obtain a theory analogous to the classical one, as may be seen from the following example. Suppose a and b are relatively prime inner functions, and let $A=\operatorname{diag}(a, b)$, $B=\operatorname{diag}(ab, 1)$. A calculation shows that A and B are equivalent only if there exist x and y in H^∞ such that $ax+by = 1$, which implies that a and b have to satisfy the Carleson condition in addition to being relatively prime.

A *quasi-unit* $\mathbf{X}$ of order n is a collection of $n\times n$ matrices over H^∞ such that $\det\mathbf{X}$ is relatively prime, where $\det\mathbf{X}=\{\det X: X\in\mathbf{X}\}$. Clearly, if X is a unit, then $\{X\}$ is a quasi-unit, but the collection of all nonconstant inner functions a quasi-unit of order one which contains no unit. It is easy to see that products of quasi-units

[2]) I am indebted to the Referee for supplying the present version of the proof, which is much more lucid than the original.

are also quasi-units: if $\mathbf{X}$ and $\mathbf{Y}$ are quasi-units of the same order, then $\mathbf{XY}=\{XY: X\in\mathbf{X}$ and $Y\in\mathbf{Y}\}$ is also a quasi-unit since $\det\mathbf{XY} = \det\mathbf{X}\cdot\det\mathbf{Y}$, which is relatively prime by Lemma 1. 1.

If A and B are $m\times n$ matrices over H^∞, then A will be called *quasi-equivalent* to B if there exist quasi-units $\mathbf{X}$ and $\mathbf{Y}$ of orders m and n respectively such that $\mathbf{X}A=B\mathbf{Y}$. By the remarks of the preceding paragraph, equivalence implies quasi-equivalence, and quasi-equivalence is transitive.

In presenting some of our arguments the following definition will be found useful. If A and B are $m\times n$ matrices and δ is an H^∞ function, then A will be called *δ-equivalent* to B if there are square matrices X and Y of orders m and n respectively such that $XA=BY$ and $(\det X)_i$ and $(\det Y)_i$ are factors of δ. It is immediate that if A is δ-equivalent to B for all δ in a relatively prime family, then A is quasi-equivalent to B. Let A^t denote the transpose of A.

Lemma 2. 1.

a) *If A is δ-equivalent to B and B is ε-equivalent to C, then A is $\delta\varepsilon$-equivalent to C.*

b) *If A is δ-equivalent to B, then B is $\delta^{k(k-1)}$-equivalent to A, where k is the larger of the dimensions of A and B.*

c) *If A is δ-equivalent to B, then A^t is $\delta^{k(k-1)}$-equivalent to B^t, where k is as above.*

Proof.

a) If $XA=BY$ and $UB=CV$, then $UXA=CVY$, and the assertion follows from the multiplicative property of determinants.

b) If $XA=BY$, then multiplying this equation on the left by adj X, the classical adjoint of X, and on the right by adj Y leads to

$$(\det Y)(\operatorname{adj} X)B=A(\det X)(\operatorname{adj} Y).$$

If X is $m\times m$, then

$$\det(\det Y\cdot\operatorname{adj} X)=(\det Y)^m(\det X)^{m-1},$$

and this together with the corresponding relation for $\det X\cdot\operatorname{adj} Y$ implies the assertion.

c) This part follows from the defining relation for δ-equivalence by taking transposes and applying part b).

Invariant factors for matrices over H^∞ may be defined in the usual way. If A is an $m\times n$ matrix let $\mathscr{D}_0(A)=1$ and let $\mathscr{D}_k(A)$ be the greatest common divisor of all minors of order k of A, where k is no larger than $\min\{m, n\}$. The invariant factors are then defined by $\mathscr{E}_k(A)=\mathscr{D}_k(A)/\mathscr{D}_{k-1}(A)$ for $k\geqq 1$ such that $\mathscr{D}_k(A)\neq 0$.

Lemma 2. 2. *If A is δ-equivalent to B, then $\mathscr{D}_k(A)|\delta^k\mathscr{D}_k(B)$ and $\mathscr{D}_k(B)|\delta^k\mathscr{D}_k(A)$ for all k such that $\mathscr{D}_k(A)\neq 0$.*

Proof. Suppose $XA=BY$, $(\det X)_i|\delta$ and $(\det Y)_i|\delta$. From the fact that the minors of a product of matrices are linear combinations of the minors of corresponding order of either factor, it follows that $\mathscr{D}_k(A)|\mathscr{D}_k(XA)$ and also $\mathscr{D}_k(BY)|\mathscr{D}_k((\det Y)B)$, since $(\det Y)B=BY \operatorname{adj} Y$. By supposition, $\mathscr{D}_k(XA)=\mathscr{D}_k(BY)$, and hence

$$\mathscr{D}_k(A)\,|\mathscr{D}_k((\det Y)B),\quad \text{i.e.}\quad \mathscr{D}_k(A)|(\det Y)^k\mathscr{D}_k(B).$$

This implies $\mathscr{D}_k(A)|\delta^k\mathscr{D}_k(B)$, and the other relation may be obtained similarly.

Theorem 2. 1. *If two matrices over H^∞ are quasi-equivalent, then they have the same invariant factors.*

Proof. Suppose A and B are matrices over H^∞, and $\mathbf{X}$ and $\mathbf{Y}$ are quasi-units such that $\mathbf{X}A=B\mathbf{Y}$. If $XA=BY$, then as in the proof of Lemma 2. 2, $\mathscr{D}_k(A)|(\det Y)^k\mathscr{D}_k(B)$. Since $\mathbf{Y}$ is a quasi-unit, it follows from Lemma 1. 1 that $\mathscr{D}_k(A)|\mathscr{D}_k(B)$. The relation $\mathscr{D}_k(B)|\mathscr{D}_k(A)$ follows similarly, and hence $\mathscr{D}_k(A)=\mathscr{D}_k(B)$ which implies the assertion.

3. Diagonalization

Our principal goal is to prove the converse of Theorem 2. 1, and this will be accomplished by showing that every matrix is quasi-equivalent to a canonical one. A matrix E over H^∞ is in *normal form* (or simply *normal*) provided

$$E=\begin{bmatrix} D & 0\\ 0 & 0\end{bmatrix},$$

where D is a diagonal matrix of nonzero inner functions, each with a positive first nonvanishing Taylor coefficient, and each one except the first divisible by its predecessor. (Some of the blocks of zeros or even D may not be present.) As in the classical case, the diagonal entries of D are the invariant factors of E (see e.g. [3], p. 91).

Lemma 3. 1. *If $A=\begin{bmatrix} A_{11} & 0\\ A_{21} & A_{22}\end{bmatrix}$, then for each inner function ψ there is a matrix X such that $(\det X)\wedge\psi=1$ and AX is of the form $\begin{bmatrix} B_{11} & 0\\ B_{21} & B_{22}\end{bmatrix}$, where $B_{11}=A_{11}$, and except possibly for the first, all entries of the first row of B_{22} are zeros.*

Proof. Let $(a_1, a_2, \ldots, a_k)$ be the first row of A_{22}. It will suffice to produce a $k\times k$ matrix X whose determinant is relatively prime to ψ such that $(a_1, a_2, \ldots, a_k)X$ has at most a nonvanishing first entry. For the required matrix may then be produced by forming the direct sum of an appropriate identity matrix with X.

The matrix X is obtained in $k-1$ steps, each step changing one a_j to zero. If the a_j are all zero, then there is nothing to prove. Permuting columns if necessary, assume $a_1 \neq 0$ and let $\omega = a_1 \wedge a_2$. By Lemma 1. 2, x and y may be chosen so that if $\delta_1 = (a_1 x + a_2 y)/\omega$, then $\delta_1 \wedge \psi = 1$. Let

$$X_1 = \begin{bmatrix} x & -a_2/\omega \\ y & a_1/\omega \end{bmatrix} \oplus I_{k-2}.$$

Then $\det X_1 = \delta_1$, the second component of $(a_1, a_2, \dots, a_k)X_1$ is zero, and the entries beyond the second are unchanged.

After the second and third columns of the result are permuted this procedure may be repeated, and in $k-1$ steps a matrix is produced that has its only nonzero entry as the first. The matrix X is obtained as a product of permutation matrices and matrices of the form X_1, and the result follows from the fact that products of functions relatively prime to ψ are also relatively prime to ψ.

Lemma 3. 2. *If A is an $m \times n$ matrix over H^∞ and ψ is any inner function then there exist an $m \times m$ matrix X and an $n \times n$ matrix Y, each with determinant relatively prime to ψ, such that XA is upper triangular and AY is lower triangular.*

Proof. The upper triangular case follows from the lower by taking transposes, and the lower triangular case is proved by repeated use of Lemma 3. 1.

Theorem 3. 1. *Every finite matrix over H^∞ is quasi-equivalent to a unique normal matrix. In fact given any $m \times n$ matrix A over H^∞ and any inner function ψ, A is δ-equivalent to the normal matrix formed from the invariant factors of A for some δ relatively prime to ψ.*

Proof. Since a normal matrix is determined by its invariant factors, uniqueness is a consequence of Theorem 2. 1. The second statement implies the remaining part of the first, and by Lemma 2. 1. c, it suffices to consider the case $m \leqq n$, since the case $m > n$ follows by taking transposes.

Given an inner function ψ, if δ is relatively prime to an inner multiple of ψ, then it is also relatively prime to ψ. Hence there is no loss of generality in considering a ψ divisible by each nonzero $\mathcal{D}_k(A)$. We suppose further without loss of generality that $\mathcal{D}_1(A) = 1$. The major portion of the proof consists of verifying that there exists η relatively prime to ψ such that A is η-equivalent to a normal matrix E_1.

The proof is by induction on m, and the case $m=1$ (and arbitrary $n \geqq 1$) is an easy consequence of Lemma 3. 2. For it implies the existence of an $n \times n$ matrix Y having a determinant relatively prime to ψ such that AY is lower triangular, i.e. AY is a $1 \times n$ matrix with at most its first entry a nonzero. Thus $AY = XE_1$, where X is the 1×1 matrix whose single entry is the outer factor of a, and E_1 is the $1 \times n$

normal matrix whose first entry is the inner factor of a. Taking $\eta=(\det Y)^{n(n-1)}$, we see by Lemma 2. 1 that A is η-equivalent to E_1.

Suppose the assertion true, therefore, for $(m-1)\times\nu$ matrices with $\nu \geqq m-1$. By Lemmas 3. 2 and 2. 1 again, there exists a δ_1, relatively prime to ψ such that A is δ_1-equivalent to a lower triangular matrix A_1. The last $n-m$ columns of A_1, which consist only of zeros, do not essentially affect the subsequent calculations, and thus it will be assumed that A_1 is an $m\times m$ lower triangular matrix.

If A_1' consists of the last $m-1$ rows and columns of A_1, then the inductive hypothesis implies there exists a δ_2 relatively prime to ψ such that A_1' is δ_2-equivalent to a normal matrix E_1', the equivalence being effected by a pair of matrices X_1' and Y_1'. If A_1'' consists of the last $m-1$ rows of the first column of A_1, then let A_2 have the same first row as A_1 and $(X_1' A_1'' E_1')$ as its last $m-1$ rows. Let $X_1 = I_1\oplus X_1'$ and $Y_1 = I_1\oplus Y_1'$. Then $X_1 A_1=A_2 Y_1$, i.e. A_1 is δ_2-equivalent to A_2, and $\mathscr{D}_1(A_2)$ is the greatest common divisor of the entries in the first two columns of A_2.

By Lemma 3. 2, there exists X_2 with determinant δ_3 relatively prime to ψ such that if $A_3'=X_2 A_2$, then A_3' is upper triangular. The greatest common divisor ε of the elements in the first two columns of A_3' is a factor of $\delta_3\mathscr{D}_1(A_2)$, as may be seen by applying Lemma 2. 2 to the first two columns of A_2 and A_3'. But A is $\delta_1\delta_2$-equivalent to A_2 by Lemma 2. 1. a, and thus Lemma 2. 2 together with the initial supposition on $\mathscr{D}_1(A)$ yield $\varepsilon|\delta_1\delta_2\delta_3$. Hence if A_3 is obtained from A_3' by dividing the entries in the first two columns by ε, then A_3' is ε^2-equivalent to A_3 and $\varepsilon^2\wedge\psi = 1$. Further, if a and b are the first two entries of the first row of A_3 and if c is the second entry of the second row, then $\wedge\{a, b, c\}=1$.

It may be assumed that a or b is nonzero, for otherwise the interchange of the first two rows and columns yields an equivalent matrix satisfying this condition. Let $\omega = a\wedge b$, and choose x and y in H^∞ according to Lemma 1. 2 so that $y\wedge\omega = 1$ and if $\delta_4 = (ax+by(/\omega$, then $\delta_4\wedge\psi = 1$. If

$$X_2' = \begin{bmatrix} x & -b/\omega \\ y & a/\omega \end{bmatrix} \oplus \delta_4 I_{m-2},$$

then all entries of A_3X_2', except possibly for those in the second row, are divisible by δ_4, and hence if $Y_2=\operatorname{diag}(\delta_4, 1, \delta_4, \delta_4, \ldots, \delta_4)$, then $A_3X_2'=Y_2A_4$, where the first two entries ω and cy of the first column of A_4 are relatively prime. From the form of X_2' and Y_2 it is not hard to see that A_3 is δ_4-equivalent to A_4.

Since ω and cy are relatively prime, another application of Lemma 1. 2 yields H^∞ functions u and v such that $\delta_5 = u\omega+vcy$ is relatively prime to ψ. Let

$$X_3 = \begin{bmatrix} u & v \\ -cy & \omega \end{bmatrix} \oplus I_{m-2};$$

X_3 has determinant δ_5, and $X_3 A_4$ has δ_5 as the only nonzero entry of the first column.

If A_5 is obtained from $X_3 A_4$ be replacing this entry by 1, and if $Y_3 \doteq (\delta_5) \oplus I_{m-1}$, then $X_3 A_4 = A_5 Y_3$. By equivalence transformations, all entries but the first of row one of A_5 can be changed to zeros yielding an equivalent matrix A_6 that is a direct sum of (1) with an $(m-1)\times(m-1)$ matrix. A second application of the induction hypothesis then yields δ_6 relatively prime to ψ such that A_6 is δ_6-equivalent to a normal matrix E_1.

Combining the above steps, we see that if η is the product of the six δ_j's and ε^2, then η is relatively prime to ψ and A is η-equivalent to a normal matrix E_1. This completes the induction.

In general it can not be supposed that E_1 is the matrix E formed from the invariant factors of A. By Lemma 2. 2, however, there exist inner functions $\alpha_1, \alpha_2, \dots, \alpha_k$, where k is the largest of the indices for which $\mathscr{D}_j(A) \neq 0$, such that each α_j divides η^j and

$$\mathscr{D}_j(E_1) = \alpha_j \mathscr{D}_j(A).$$

Since each $\mathscr{D}_j(A)$ divides ψ, and since η is prime to ψ, it follows that α_j is prime to $\mathscr{D}_l(A)$ for all j and l. Thus with $\alpha_0 = 1$,

$$\mathscr{E}_j(E_1) = (\alpha_j/\alpha_{j-1})\mathscr{E}_j(A),$$

and each α_j/α_{j-1} is inner. Thus if

$$Y_4 = \operatorname{diag}(\alpha_1/\alpha_0, \alpha_2/\alpha_1, \dots, \alpha_k/\alpha_{k-1}) \oplus I_{n-k},$$

then $\det Y_4 = \alpha_k$, which divides η, and $E_1 = EY_4$. Hence if $\delta = \eta^2$, then $\delta \wedge \psi = 1$ and A is δ-equivalent to E. On the other hand,

$$\mathscr{E}_j(A)\alpha_j^2 \mathscr{E}_{j+1}(E_1)/\mathscr{E}_j(E_1) = \alpha_{j+1}\alpha_{j-1}\mathscr{E}_{j+1}(A) \quad (j = 1, \dots, k-1),$$

and this together with the fact that $\mathscr{E}_j(A)$ is relatively prime to $\alpha_{j+1}\alpha_{j-1}$ imply that $\mathscr{E}_j(A) | \mathscr{E}_{j+1}(A)$, i. e. E is normal. This completes the proof.

Corollary 3. 1. *An $m\times n$ matrix A is quasi-equivalent to an $m\times n$ matrix B over H^∞ if and only if A and B have the same invariant factors.*

Proof. Necessity was established in Theorem 2. 1. If A and B have the same invariant factors, then each one determines the same normal matrix E. By the theorem, A is quasi-equivalent to E which is quasi-equivalent to B, and this establishes the result.

Corollary 3. 2. *Quasi-equivalence is an equivalence relation.*

Corollary 3. 3. *If A and B are quasi-equivalent, then there exist matrices X and Y each of whose determinants is relatively prime to all the invariant factors of A and B, and such that $XA = BY$.*

Proof. Let E be the normal matrix that A and B are quasi-equivalent to, and let ψ be a multiple of each of the nonzero entries of E. Two applications of the theorem yield δ_1 and δ_2 relatively prime to ψ such that A is δ_1-equivalent to E and B is δ_2-equivalent to E. If k is the larger of m and n, then setting $\delta=\delta_1\delta_2{}^{k(k-1)}$, we see from Lemma 2. 1 that A is δ-equivalent to B. Thus there exist X and Y such that $(\det X)_i|\delta$, $(\det Y)_i|\delta$, (hence $\det X$ and $\det Y$ are relatively prime to ψ) and $XA=BY$.

Corollary 3. 4. *Suppose $XA=BY$. If the determinants of X and Y are relatively prime to the invariant factors of A and B, then A and B are quasi-equivalent.*

Proof. The first hypothesis implies that A and B are δ-equivalent for $\delta = (\det X\cdot\det Y)_i$, By Lemma 2. 2, $\mathscr{D}_k(A)|\delta^k\mathscr{D}_k(B)$, and by the second hypothesis, $\mathscr{D}_k(A)|\mathscr{D}_k(B)$. Similarly $\mathscr{D}_k(B)|\mathscr{D}_k(A)$, and hence A and B have the same invariant factors. Thus Corollary 3. 1 implies the result.

Corollary 3. 5. *Suppose $XA=BY$, where A and B are square matrices. If X and Y have the same determinant, and it is relatively prime to* $\det A$, *then A and B are quasi-equivalent.*

Proof. If $\det A=0$, then $\det X$ and $\det Y$ are outer and hence relatively to prime to the invariant factors of A and B. If $\det A\neq 0$, then the relation $XA=BY$ implies that $\det B=\det A$. Since $\det A$ is the product of the invariant factors of A up to an outer factor, it follows that $\det X$ and $\det Y$ are relatively prime to the invariant factors of A, and similarly to those of B. Hence in either case the result follows from Corollary 3. 4.

4. A reformulation

The definition of quasi-equivalence may be formulated in a slightly different way which is more general and leads to an open question. Let $\mathfrak{H}$ be a separable Hilbert space and suppose X is a bounded analytic function on the unit disc D whose values are operators on $\mathfrak{H}$ (see [1] or [7], Chap. V.). If X admits a scalar multiple, let Φ_X be the set of scalar multiples of X, and let $\varphi_X=\wedge\Phi_X$. Then in the case of X inner φ_X is the characteristic scalar inner function of X (cf. [1], p. 81). If $\mathbf{X}$ is a collection of analytic operator valued functions admitting scalar multiples such that $\{\varphi_X : X\in\mathbf{X}\}$ is relatively prime, then $\mathbf{X}$ is called a quasi-unit on $\mathfrak{H}$. In the case of $\mathfrak{H}$ finite dimensional, this definition agrees with the one given previously ([1], p. 81; [7]).

Let $\mathfrak{H}$ and $\mathfrak{K}$ be a pair of separable Hilbert spaces. If A and B are bounded analytic functions on D whose values are operators from $\mathfrak{H}$ to $\mathfrak{K}$, then A is quasi-equivalent to B in case there exist quasi-units $\mathbf{X}$ and $\mathbf{Y}$ on $\mathfrak{H}$ and $\mathfrak{K}$ respectively such that $\mathbf{X}A=B\mathbf{Y}$. Is every operator function or every operator function that admits a scalar multiple quasi-equivalent to a diagonal one?

References

[1] H. HELSON, *Lectures on invariant subspaces*, Academic Press (New York, 1964).
[2] K. HOFFMAN, *Banach spaces of analytic functions*, Prentice Hall (Englewood Cliffs, N. J., 1962).
[3] J. JACOBSON, *Lectures in abstract algebra*, Vol. II, Van Nostrand (Princeton, N. J., 1953).
[4] B. MOORE, III and E. A. NORDGREN, On quasi-equivalence and quasi-similarity, *Acta Sci. Math.*, **34** (1973), 283—288.
[5] T. P. SRINIVASAN, Simply invariant subspaces and generalized analytic functions, *Proc. Amer. Math. Soc.*, **16** (1965), 813—818.
[6] B. SZ.-NAGY and C. FOIAŞ, Sur les contractions de l'espace de Hilbert. VII. Triangulations canoniques. Fonction minimum, *Acta Sci. Math.*, **25** (1964), 12—37.
[7] B. SZ.-NAGY and C. FOIAŞ, Sur les contractions de l'espace de Hilbert. VIII Fonctions caractéristiques. Modèles fonctionnels, *Acta Sci. Math.*, **25** (1964), 38—71.
[8] B. SZ.-NAGY and C. FOIAŞ, Modèle de Jordan pour une classe d'opérateurs de l'espace de Hilbert, *Acta Sci. Math.*, **31** (1970), 93—117.
[9] B. SZ.-NAGY and C. FOIAŞ, *Harmonic analysis of operators on Hilbert space*, Akadémiai Kiadó (Budapest, 1970).

(Received August 12, 1971)

On quasi-equivalence and quasi-similarity

By BERRIEN MOORE, III and ERIC A. NORDGREN in Durham (N. H., U.S.A.)[1]

Dedicated to Professor Béla Szőkefalvi-Nagy on his 60th birthday

Introduction

The object of this note is to show how the theory of Jordan models for C_0 contractions of finite defect, which was developed by Sz.-Nagy and Foiaş [10—17], may be approached from the relation of quasi-equivalence for H^∞ matrices [6]. We thereby give a different approach to the Jordan model theory and in so doing establish a conjecture of Sz.-Nagy and Foiaş [15] on the relation between the inner functions in the Jordan model of an operator T and its characteristic operator function Θ_T. Our results show that the analogy with the finite dimensional situation is complete.

1. Preliminaries

Let $\mathfrak{E}$ be a separable complex Hilbert space and m normalized Lebesgue measure on the unit circle C of the complex plane. Then $L^2(\mathfrak{E})$ is the Hilbert space of all weakly measurable functions from C to $\mathfrak{E}$ having square integrable norm, and $H^2(\mathfrak{E})$ is the corresponding Hardy subspace. If χ is the identity function on C, then the bilateral shift operator U on $L^2(\mathfrak{E})$ is given by

$$Uf=\chi f \quad (f\in L^2(\mathfrak{E})),$$

where the operation is that of pointwise multiplication. The unilateral shift U_+ on $H^2(\mathfrak{E})$ is simply the restriction of U to $H^2(\mathfrak{E})$. The above is discussed in detail by Helson [4] and Sz.-Nagy and Foiaş [18].

The algebra of weakly measurable, essentially bounded functions from C to the algebra $\mathscr{B}(\mathfrak{E})$ of bounded operators on $\mathfrak{E}$ is $L^\infty(\mathscr{B}(\mathfrak{E}))$. A function Θ in $L^\infty(\mathscr{B}(\mathfrak{E}))$ is said to be analytic if

$$\int (\Theta(z)f, g) z^n \, dm(z)=0 \quad (f, g\in\mathfrak{E}, n=1, 2, \ldots),$$

and $H^\infty(\mathscr{B}(\mathfrak{E}))$ is the algebra of analytic functions in $L^\infty(\mathscr{B}(\mathfrak{E}))$.

---

[1]) This research was supported in part by a grant from the National Science Foundation

If Θ is in $H^\infty(\mathscr{B}(\mathfrak{E}))$, then there are two operators that are naturally associated with it. Each sends a function u to Θu, where

$$\Theta u(z) = \Theta(z)u(z) \quad (z \in C),$$

but the domain of one is $L^2(\mathfrak{E})$ whereas that of the other is $H^2(\mathfrak{E})$. The former is the associated analytic Laurent operator and the latter the associated analytic Toeplitz operator. We will be somewhat imprecise and use the same notation for each of these operators as for the function that induces it, relying on the context to make it clear which is intended. If the operator Θ on $H^2(\mathfrak{E})$ is a partial isometry, then the analytic function inducing it will be called inner. This usage does not quite conform to that of SZ.-NAGY and FOIAŞ [18, p. 190], who use the term inner for an analytic function whose values are operators from one Hilbert space $\mathfrak{E}_1$ to another $\mathfrak{E}_2$ such that the induced analytic Toeplitz operator from $H^2(\mathfrak{E}_1)$ to $H^2(\mathfrak{E}_2)$ is isometric. The difference is inessential in that with either definition the typical invariant subspace of U_+ is the range of an analytic Toeplitz operator induced by some inner function ([1], [3], [5], [7] and [8]).

For each inner Θ let $S(\Theta)$ be the compression of U_+ to the orthogonal complement $\mathfrak{H}(\Theta)$ of $\Theta H^2(\mathfrak{E})$. It is now well known that every contraction T on a separable Hilbert space such that $\{T^{*n}\}$ converges to zero strongly is unitarily equivalent to some $S(\Theta)$ [18].

Our present interest centers on the case where $\mathfrak{E}$ is finite dimensional and Θ is an isometric operator on $H^2(\mathfrak{E})$. If Θ is purely contractive [18, p. 188], and if $\mathfrak{E}$ has dimension N, then both defect indices of $S(\Theta)$ are N, and $S(\Theta)$ belongs to class $C_0(N)$. In fact, up to unitary equivalence the most general contraction of class $C_0(N)$ arises this way [11]. In general, we will not require Θ to be purely contractive, and thus the resulting $S(\Theta)$ is of class $C_0(M)$ for some positive integer $M \leqq N$.

An operator X from a Hilbert space $\mathfrak{H}_1$ to another $\mathfrak{H}_2$ is called a quasi-affinity in case it is one to one and has dense range. An operator T_1 on $\mathfrak{H}_1$ is a quasi-affine transform of an operator T_2 on $\mathfrak{H}_2$ in case there exists a quasi-affinity X from $\mathfrak{H}_1$ to $\mathfrak{H}_2$ such that

$$XT_1 = T_2 X,$$

in which case we write $T_1 \prec T_2$. If $T_1 \prec T_2$ and $T_2 \prec T_1$, then T_1 and T_2 are called quasi-similar. Again, for a more detailed discussion of these ideas refer to the text by SZ.-NAGY and FOIAŞ [18].

Two functions Θ_1 and Θ_2 in $H^\infty(\mathscr{B}(\mathfrak{E}))$ are said to be *equivalent* in case there exist two invertible functions Δ and Λ in $H^\infty(\mathscr{B}(\mathfrak{E}))$ such that

$$\Delta\Theta_1 = \Theta_2\Lambda.$$

A function Δ in $H^\infty(\mathscr{B}(\mathfrak{E}))$ is said to have a scalar H^∞ function δ as a scalar multiple if there exists a function Ω in $H^\infty(\mathscr{B}(\mathfrak{E}))$ such that

$$\Omega\Delta = \Delta\Omega = \delta I_{\mathfrak{E}},$$

where $I_{\mathfrak{E}}$ is the identity operator on $\mathfrak{E}$. Finally, the following concepts introduced in [6] are fundamental to our present considerations. A *quasi-unit* $Д$ is a subset of $H^\infty(\mathscr{B}(\mathfrak{E}))$ with the property that the collection of all scalar multiples of the functions in $Д$ is nonempty and relatively prime, i.e. has no nonconstant common inner factor. If Θ_1 and Θ_2 belong to $H^\infty(\mathscr{B}(\mathfrak{E}))$ and if there exist quasi-units $Д$ and $Л$ such that

$$Д\Theta_1 = \Theta_2 Л,$$

then Θ_1 and Θ_2 are called *quasi-equivalent*.

In this paper we will study the relation between quasi-equivalence and quasi-similarity, but we pause to note that in one direction at least the relation between equivalence and similarity parallels the finite dimensional situation.

Theorem 1. *Let Θ_1 and Θ_2 be inner functions in $H^\infty(\mathscr{B}(\mathfrak{E}))$. If Θ_1 and Θ_2 are equivalent, then $S(\Theta_1)$ and $S(\Theta_2)$ are similar.*

Proof. Let Δ and Λ be invertible functions in $H^\infty(\mathscr{B}(\mathfrak{E}))$ such that

(1) $$\Delta\Theta_1 = \Theta_2\Lambda.$$

It follows from (1) that $\Delta\Theta_1 H^2(\mathfrak{E}) \subset \Theta_2 H^2(\mathfrak{E})$, and hence

(2) $$P_2\Delta P_1 = P_2\Delta,$$

where P_j is the orthogonal projection of $H^2(\mathfrak{E})$ onto $\mathfrak{H}(\Theta_j)$ for $j=1, 2$. Define an operator X by

(3) $$X = P_2\Delta|\mathfrak{H}(\Theta_1).$$

For every $f \in \mathfrak{H}(\Theta_1)$ we have by (2) and (3) that

$$XS(\Theta_1)f = P_2\Delta P_1 U_+ f = P_2\Delta U_+ f = P_2 U_+ \Delta f = P_2 U_+ P_2 \Delta f = S(\Theta_2)Xf,$$

where the next to last equality follows from the invariance of $\Theta_2 H^2(\mathfrak{E})$ under U_+. It remains only to show that X is invertible. We have by hypothesis that Δ is invertible; therefore, if v is an element of $\mathfrak{H}(\Theta_2)$, then there exists a unique element u in $H^2(\mathfrak{E})$ such that

$$v = \Delta u = P_2\Delta u.$$

Employing (2), we have

$$v = P_2\Delta P_1 u = XP_1 u;$$

thus X is onto. If $Xu=0$ for some u in $\mathfrak{H}(\Theta_1)$, then $\Delta u = \Theta_2 v$ for some v in $H^2(\mathfrak{E})$, and since Λ is invertible there exists some v' in $H^2(\mathfrak{E})$ such that $\Lambda v'=v$. Hence by (1),

$$\Delta(u-\Theta_1 v') = \Delta u - \Theta_2 \Lambda v' = 0.$$

But Δ is invertible and u is orthogonal to $\Theta_1 v'$; therefore, $u=0$, and X is one-to-one. Thus X is invertible.

It would be interesting to know to what extent Theorem 1 has a converse; it is known [2] that defect indices are not similarity invariants.

2. Main theorems

The following result is an analog of Theorem 1 for quasi-equivalence.

Theorem 2. *Let $\mathfrak{E}$ be finite dimensional and suppose that Θ_1 and Θ_2 are inner functions in $H^\infty(\mathscr{B}(\mathfrak{E}))$ that induce isometries on $H^2(\mathfrak{E})$. If Θ_1 and Θ_2 are quasi-equivalent, then $S(\Theta_1)$ and $S(\Theta_2)$ are quasi-similar.*

Proof. Since quasi-equivalence is an equivalence relation [6, Cor. 3. 2], it will suffice to prove $S(\Theta_1)\prec S(\Theta_2)$. The facts that E is finite dimensional and that Θ_1 and Θ_2 induce isometries on $H^2(E)$ imply Θ_1 and Θ_2 are each unitary valued a.e. and consequently their determinants are nonzero a.e. By Corollary 3. 3 of [6], there exist functions Δ and Λ in $H^\infty(\mathscr{B}(\mathfrak{E}))$ whose determinants are relatively prime to those of Θ_2 and Θ_1, and such that

$$\Delta\Theta_1 = \Theta_2 \Lambda.$$

Define X as in (3). The same reasoning used previously implies

$$XS(\Theta_1) = S(\Theta_2)X;$$

therefore, we need only show that X is a quasi-affinity.

Suppose v in $\mathfrak{H}(\Theta_2)$ is orthogonal to the range of X, and let u be any vector in $H^2(\mathfrak{E})$. There exist u' in $\mathfrak{H}(\Theta_1)$ and u'' in $H^2(\mathfrak{E})$ such that

$$u = u' + \Theta_1 u''.$$

By supposition,

$$(v, \Delta u') = (P_2 v\, \Delta u') = (v, P_2\, \Delta u') = v, Xu') = 0,$$

and since $\Delta\Theta_1 = \Theta_2 \Lambda$, we have

$$(v, \Delta\Theta_1 u'') = (v, \Theta_2 \Lambda u'') = 0.$$

Thus v is orthogonal to $\Delta H^2(\mathfrak{E})$, which includes $(\det \Delta) H^2(\mathfrak{E})$. But v is also orthogonal to $\Theta_2 H^2(\mathfrak{E})$, which includes $(\det \Theta_2) H^2(\mathfrak{E})$. Since $\det \Delta$ and $\det \Theta_2$ are relatively prime, it follows that $\Delta H^2(\mathfrak{E})$ and $\Theta_2 H^2(\mathfrak{E})$ span $H^2(\mathfrak{E})$ (see [1]); thus $v=0$. Consequently X has dense range.

Now consider an arbitrary $u \in \mathfrak{H}(\Theta_1)$ such that $Xu=0$; i.e. for which $\Delta u \in \Theta_2 H^2(\mathfrak{E})$. As mentioned earlier, each of Θ_1 and Θ_2 is unitary valued a.e., and therefore, the operators Θ_1 and Θ_2 on $L^2(\mathfrak{E})$ are unitary. Using this fact and setting $f=\Theta_1^* u$, it follows from (1) that $\Theta_2 \Lambda f = \Delta \Theta_1 f = \Delta u \in \Theta_2 H^2(\mathfrak{E})$, and hence $\Lambda f \in H^2(\mathfrak{E})$. On the other hand, $\Theta_1 f = \Theta_1 \Theta_1^* u = u$. Thus, Λf and $\Theta_1 f$ are both in $H^2(\mathfrak{E})$. From this it follows that

$$(\det \Lambda) f \in H^2(\mathfrak{E}) \quad \text{and} \quad (\det \Theta_1) f \in H^2(\mathfrak{E}).$$

Since $\det \Lambda$ and $\det \Theta_1$ are relatively prime, we infer by using a lemma of SZ.-NAGY [9, p. 74], that $f \in H^2(\mathfrak{E})$. This implies $u \in \Theta_1 H^2(\mathfrak{E})$. As $u \in \mathfrak{H}(\Theta_1)$, we necessarily have $u=0$. Hence, X is one to one, and the proof of the theorem is complete.

It is shown in [6], Theorem 3. 1 that every $N \times N$ matrix over H^∞ is quasi-equivalent to a diagonal one with the invariant factors on the main diagonal. Thus if Θ is the characteristic operator function of an operator T of class $C_0(N)$, then Θ is quasi-equivalent to a normal matrix Θ', i.e. Θ' is diagonal and the diagonal entries of Θ' are the invariant factors of Θ. From the theorem then we have that T is quasi-similar to $S(\Theta')$. Operators of the form $S(\Theta')$ are called Jordan operators by SZ.-NAGY and FOIAŞ, and they were the first to show that every $C_0(N)$ contraction T is quasi-similar to a Jordan operator and that the minimal inner function of T is the first invariant factor of Θ_T [11, 15]. We have thus obtained their results via a different route; moreover, we have shown that the inner functions that appear in the Jordan model of an operator are related to the characteristic operator function in the manner they conjectured. To summarize, we have established the following:

Theorem 3. *If T is an operator of class $C_0(N)$ for some integer N, then T is quasi-similar to a Jordan operator determined by the invariant factors of the characteristic operator function of T.*

Finally, Theorem 2 has a converse; that the statement is more general is illusory.

Theorem 4. *Let $\mathfrak{E}$ be finite dimensional and suppose that Θ_1 and Θ_2 are inner functions that induce isometries on $H^2(\mathfrak{E})$. If $S(\Theta_1)$ is a quasi-affine transform of $S(\Theta_2)$, then Θ_1 and Θ_2 are quasi-equivalent.*

Proof. Let Θ_1' and Θ_2' be the normal matrices that are quasi-equivalent to Θ_1 and Θ_2 respectively. Then by Theorem 2 and the hypothesis,

$$S(\Theta_1') \prec S(\Theta_1) \prec S(\Theta_2) \prec S(\Theta_2').$$

It was shown by SZ.-NAGY and FOIAŞ [14] that if one Jordan operator is a quasi-affine transform of another, then they are both determined by the same nonconstant inner functions. Consequently, $S(\Theta_1')=S(\Theta_2')$, and hence $\Theta_1'=\Theta_2'$. It follows by transitivity that Θ_1 and Θ_2 are quasi-equivalent.

We should like to express our gratitude to Professor Sz.-Nagy for suggesting a substantial improvement on the proof of Theorem 2.

References

[1] A. Beurling, On two problems concerning linear transformations in Hilbert space, *Acta Math.*, **81** (1949), 239—255.

[2] C. Foiaş and J. Williams, Some remarks on the Volterra operator. *Proc. Amer. Math. Soc.*, **31** (1972), 177—184.

[3] P. R. Halmos, Shifts on Hilbert spaces, *J. reine angew. Math.*, **208** (1961), 102—112.

[4] H. Helson, *Lectures on invariant subspaces*, Academic Press (New York, 1964).

[5] P. D. Lax, Translation invariant spaces, *Acta Math.*, **101** (1959), 163—178.

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

[7] M. Rosenblum, Vectorial Toeplitz operators and the Fejér—Riesz theorem, *J. Math. Anal. Appl.*, **23** (1968), 139—147.

[8] J. Rovnyak, Ideals of square summable power series, *Proc. Amer. Math. Soc.*, **13** (1962), 360—365.

[9] B. Sz.-Nagy, Hilbertraum-Operatoren der Klasse C_0, *Abstract spaces and approximation*, Proc. M.R.I. Oberwolfach, Birkhäuser (Basel, 1968), 72—81.

[10] B. Sz.-Nagy and C. Foiaş, Sur les contractions de l'espace de Hilbert. VII. Triangulations canoniques. Fonction minimum, *Acta Sci. Math.*, **25** (1964), 12—37.

[11] B. Sz.-Nagy and C. Foiaş, Sur les contractions de l'espace de Hilbert. VIII. Fonctions caractéristiques. Modèles fonctionnels, *Acta Sci. Math.*, **25** (1964), 38—71.

[12] B. Sz.-Nagy and C. Foiaş, Sur les contractions de l'espace de Hilbert. XI. Transformations unicellulaires, *Acta Sci. Math.*, **26** (1965), 301—324 and **27** (1966), 265.

[13] B. Sz.-Nagy and C. Foiaş, Vecteurs cycliques et quasi-affinités, *Studia Math.*, **31** (1968), 35—42.

[14] B. Sz.-Nagy and C. Foiaş, Opérateurs sans multiplicité, *Acta Sci. Math.*, **30** (1969), 1—18.

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

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

[17] B. Sz.-Nagy and C. Foiaş, Local characterizations of operators of class C_0, *J. Functional Analysis*, **8** (1971), 76—81.

[18] B. Sz.-Nagy and C. Foiaş, *Harmonic analysis of operators on Hilbert spaces*, Akadémiai Kiadó (Budapest, 1970).

(Received November 14, 1971)

A Riesz decomposition theorem in W^*-algebras

By C. PELIGRAD and L. ZSIDÓ in Bucharest (Romania)

To Prof. B. Sz.-Nagy on his 60th birthday

An extension in to W^*-algebras of the Riesz theory ([7]) of compact operators was initiated and developped by M. BREUER in [2]. Here the ideal of compact operators is replaced by the norm closed two-sided ideal generated by all finite projections. A problem raised by M. Breuer was solved by V. I. OVTSHINNIKOV in the case of von Neumann algebras on a separable Hilbert spaces ([6]). Our purpose is to solve the Breuer problem in the general case.

For notions and knowledges about W^*-algebras we send the reader to [5].

The authors are indebted to Dr. S. TELEMAN who called their attention to the papers of Breuer and Ovtshinnikov.

In the first section we prove some technical results. Lemma 3 in the second section solves the conjecture of Breuer and, together with Breuer's results, it allows us to formulate an extension of the Riesz decomposition theorem for compact operators.

1. The two-sided ideal associated to a semi-finite trace

Let M be a W^*-algebra and M^+ its positive part. A mapping

$$\varphi: M^+ \to R^+ \cup \{+\infty\}$$

is called a *normal trace* on M if:

1. $\varphi(x+y) = \varphi(x)+\varphi(y), \quad x, y \in M^+,$
2. $\varphi(\lambda x) = \lambda\varphi(x), \quad x \in M^+, \ \lambda \in R^+,$
3. $\varphi(u^* x u) = \varphi(x), \quad x \in M^+, \ u \in M$ unitary.
4. $\varphi(x) = \sup \varphi(x_\iota)$ for every increasingly filtered family (x_ι) in M^+ with $x = \sup x_\iota$.

We denote by $\mathfrak{M}_\varphi$ the linear span of $\{x: x \in M^+, \varphi(x) < +\infty\}$. $\mathfrak{M}_\varphi$ is a two sided ideal, called the two-sided ideal associated to φ. φ can be extended to a positive linear form on $\mathfrak{M}_\varphi$. If $a\in\mathfrak{M}_\varphi$, then $x\mapsto\varphi(ax)=\varphi(xa)$ is a w-continuous linear form on M, where the w-topology is the weak topology defined by the predual of M. φ is said to be *semi-finite* if:

5. $\mathfrak{M}_\varphi$ is w-dense in M.

Finally, a normal trace φ on M is called *faithful* if

6. $x\in M^+,\quad \varphi(x)=0\Rightarrow x=0$.

If there exists a faithful semi-finite normal trace on M, then M is called semi-finite.

A projection $e\in M$ is called finite if every projection $f\in M$ $f\leqq e$, $f\sim e$, equals e. The norm-closed linear span $\mathcal{I}$ of all finite projections $e\in M$ is a two-sided ideal. Every projection $e\in\mathcal{I}$ is finite.

Let M be a W^*-algebra. For every set $S\subset M$ we denote by $\langle S\rangle$ the set of all elements $\sum_{\iota\in I} x_\iota z_\iota$, where $x_\iota\in S$ and (z_ι) is a family of orthogonal central projections with $\sum_{\iota\in I} z_\iota=1$. If M is semi-finite, φ is a faithful normal semi-finite trace on M and e is a projection in $\mathfrak{M}_\varphi$, then e is finite. So, denoting by $\overline{\mathfrak{M}}_\varphi$ the norm-closure of $\mathfrak{M}_\varphi$, we have $\langle\overline{\mathfrak{M}}_\varphi\rangle\subset\mathcal{I}$.

Lemma 1. *Let M be a W^*-algebra, φ a normal semi-finite trace on M and $e\in M$ a finite projection. Then $e\in\langle\mathfrak{M}_\varphi\rangle$.*

Proof. Let $(z_\iota)_{\iota\in I}$ be a maximal family of orthogonal central projections such that $ez_\iota\in\mathfrak{M}_\varphi$ for all $\iota\in I$. Suppose that $z_0 = 1-\sum_{\iota\in I} z_\iota\neq 0$.

Since $\mathfrak{M}_\varphi$ is w-dense in M, there exists a non-zero projection $f\in\mathfrak{M}_\varphi z_0$. By the comparison theorem there exists a central projection $z_1\leqq z_0$, such that

$$fz_1\prec ez_1,\qquad f(z_0-z_1)\succ e(z_0-z_1).$$

The maximality of $(z_\iota)_{\iota\in I}$ implies $z_0-z_1=0$. So

$$f=fz_0\prec ez_0$$

that is, for some projection $f_1\in M$,

$$f\sim f_1\leqq ez_0.$$

Using again the comparison theorem, there exists a central projection $z_2\leqq z_0$, such that

$$fz_2\prec(ez_0-f_1)z_2,\qquad f(z_0-z_2)\succ(ez_0-f_1)(z_0-z_2).$$

Again, by the maximality of $(z_\iota)_{\iota\in I}$ we have $z_0-z_2=0$. Hence

$$f=fz_0\prec ez_0-f_1$$

that is, for some projection $f_2\in M$

$$f\sim f_2\leqq ez_0-f_1.$$

By induction, we obtain a sequence (f_n) of orthogonal projections such that, for every n, $f_n \sim f$ and $f_n \leqq ez_0$. This contradicts the finiteness of e.

Consequently $z_0=0$ that is $e \in \langle \mathfrak{M}_\varphi \rangle$.

Lemma 2. *Let $\mathfrak{N}$ be a norm-closed two-sided ideal in a W^*-algebra M. Then $\langle \mathfrak{N} \rangle$ is a norm-closed two-sided ideal.*

Proof. Obviously, $\langle \mathfrak{N} \rangle$ is a two-sided ideal.

Let (x_n) be a sequence in $\langle \mathfrak{N} \rangle$ convergent in the norm-topology to $x \in M$. For every n there exists a family $(z_n^\iota)_{\iota \in I_n}$ of orthogonal central projections, such that

$$x_n z_n^\iota \in \mathfrak{N}, \iota \in I_n, \quad \text{and} \quad \sum_{\iota \in I_n} z_n^\iota = 1.$$

Denote by Ω the maximal ideal space of the center of M. Then Ω is a hyperstonean space (see [3]). For every n, $\Omega_n = \bigcup_{\iota \in I_n} (z_n^\iota)^{-1}(1)$ is an open dense set, and $\bigcap_{n=1}^{\infty} \Omega_n$ contains an open dense set Ω_0. Now it is easy to see, that there exists a family $(K_\iota)_{\iota \in I}$ of mutually disjoint open and closed subsets of Ω_0 such that $\bigcup_{\iota \in I} K_\iota$ is dense in Ω. Denote by z_ι the characteristic function of K_ι. Then $(z_\iota)_{\iota \in I}$ is a family of orthogonal central projections with $\sum_{\iota \in I} z_\iota = 1$. For every $\iota_0 \in I$ and for every n, $((z_n^\iota)^{-1}(1))_{\iota \in I}$ is an open convering of $(z_{\iota_0})^{-1}(1) = K_{\iota_0}$. Hence there exist $\iota_1, \dots, \iota_{k_n} \in I$ such that

$$(z_{\iota_0})^{-1}(1) \subset \bigcup_{j=1}^{k_n} (z_n^{\iota_j})^{-1}(1), \text{ that is } z_{\iota_0} \leqq \sum_{j=1}^{k_n} z_n^{\iota_j}.$$

Consequently, for every $\iota \in I_n$ and for every n we have $x_n z_\iota \in \mathfrak{N}$. $\mathfrak{N}$ being norm-closed, we obtain $xz_\iota \in \mathfrak{N}$, $\iota \in I$. So $x \in \langle \mathfrak{N} \rangle$.

In [1], a norm-closed two-sided ideal $\mathfrak{M}$ in M such that $\langle \mathfrak{M} \rangle = \mathfrak{M}$ is called continuous. Examples of continuous ideals are the above defined ideal $\mathscr{I}$ and the strong radical. Lemma 2 states that the continuous hull of $\mathfrak{N}$ is $\langle \mathfrak{N} \rangle$.

Theorem 1. *Let M be a W^*-algebra and φ a normal semi-finite trace on M. Then $\mathscr{I} \subset \langle \overline{\mathfrak{M}}_\varphi \rangle$.*

Proof. Let $\mathscr{I}_0$ be the two-sided ideal generated by all finite projections $e \in M$. By Lemma 1, $\mathscr{I}_0 \subset \langle \mathfrak{M}_\varphi \rangle \subset \langle \overline{\mathfrak{M}}_\varphi \rangle$. On the other hand, by Lemma 2 $\langle \overline{\mathfrak{M}}_\varphi \rangle$ is norm-closed. Hence the norm-closure $\mathscr{I}$ of $\mathscr{I}_0$ is included in $\langle \overline{\mathfrak{M}}_\varphi \rangle$.

Using Theorem 1 and the remark preceding Lemma 1, we obtain:

Corollary. *If M is a semi-finite W^*-algebra and φ is a faithful normal semi-finite trace on M, then $\mathscr{I} = \langle \overline{\mathfrak{M}}_\varphi \rangle$.*

We need also the following result:

Theorem 2. *Let M be a semi-finite W^*-algebra, φ a faithful normal semi-finite trace on M, and N a W^*-subalgebra of M such that $N\cap\mathfrak{M}_\varphi$ is w-dense in N. Then there exists a positive linear map $P:M\to N$ such that*

1) *$P(1)$ is the unit of N.*
2) *$P(y_1xy_2)=y_1P(x)y_2,\ x\in M,\ y_1,y_2\in N$.*
3) *$P\mathfrak{M}_\varphi\subset N\cap\mathfrak{M}_\varphi$.*
4) *$\varphi(P(x))=\varphi(P(1)x),\ x\in\mathfrak{M}_\varphi$.*

For the proof of Theorem 2 we send the reader to [4] (see also [5] chap. III, § 5, exercise 9. f).

2. The main result

Let M be a W^*-algebra and $a\in M$. For every integer $n\geqq 1$ we denote by $e_n(a)$ the greatest projection $e\in M$ such that $(1-a)^n e=0$. We denote also $e_\infty(a)=\sup_n e_n(a)$.

Now we suppose that a belongs to the norm-closed two-sided ideal $\mathscr{I}$ generated by the finite projections. Since for every integer $n\geqq 1$ the equality $e_n(a)=\big(1-(1-a)^n\big)e_n(a)$ holds, we have $e_n(a)\in\mathscr{I}$. Consequently $e_n(a)$ are finite projections. This statement is sharpened by the following

Lemma 3. *If $a\in\mathscr{I}$ then $e_\infty(a)$ is finite.*

Proof. First, we remark that $e_\infty(a)$ belongs to the w-closure of $\mathscr{I}$. So we can suppose that M is semi-finite. Let φ be a faithful normal semi-finite trace on M.

By Theorem 1, there exists a family $(z_\iota)_{\iota\in I}$ of orthogonal central projections in M with $\sum_{\iota\in I} z_\iota=1$ such that $az_\iota\in\overline{\mathfrak{M}}_\varphi$ for all ι. Since $e_\infty(az_\iota)=e_\infty(a)z_\iota$ and it is sufficient to prove that every $e_\infty(a)z_\iota$ is finite, we can suppose that $a\in\overline{\mathfrak{M}}_\varphi$.

Let N be the W^*-subalgebra of M generated by the projections $e_n(a)$, $1\leqq n<\infty$. Then N is commutative and its unit is $e_\infty(a)$. Since $e_n(a)=\big(1-(1-a)^n\big)e_n(a)$, $1\leqq n<\infty$, the projections $e_n(a)$ belong to $\overline{\mathfrak{M}}_\varphi$. By [5], chap. I, § 1, exercise 6, they belong even to $\mathfrak{M}_\varphi$. Hence $N\cap\mathfrak{M}_\varphi$ is w-dense in N.

By Theorem 2 there exists a positive linear map $P:M\to N$ such that

1) $P(1)=e_\infty(a)$,
2) $P(y_1xy_2)=y_1P(x)y_2,\quad x\in M,\ y_1,y_2\in N$,
3) $P\mathfrak{M}_\varphi\subset N\cap\mathfrak{M}_\varphi$.

For every $2\leqq n<\infty$ we have

$$(1-a)^{n-1}\big(1-e_{n-1}(a)\big)(1-a)e_n(a)=0.$$

Since $1-e_{n-1}(a)$ is the right support of $(1-a)^{n-1}$, it follows that

$$(1-e_{n-1}(a))(1-a)e_n(a)=0$$

Consequently

$$P(1-a)(e_n(a)-e_{n-1}(a)) = (e_\infty(a)-e_{n-1}(a)P(1-a)e_n(a) =$$
$$= P((e_\infty(a)-e_{n-1}(a))(1-a)e_n(a)) = 0.$$

This equality implies

$$P(1-a)e_1(a) = P(1-a)e_2(a) = \dots = P(1-a)e_\infty(a) = P(1-a).$$

Since

$$P(1-a)e_1(a) = P((1-a)e_1(a)) = 0,$$

we conclude

$$P(1-a) = 0.$$

Now we have

$$e_\infty(a) = P(1) = P(a)\in \overline{N\cap \mathfrak{M}_\varphi}.$$

By [5], chap. I, § 1, exercise 6, $e_\infty(a)\in N\cap\mathfrak{M}_\varphi \subset \mathfrak{M}_\varphi$. Hence $e_\infty(a)$ is a finite projection.

For every $a\in M$ and for every integer $n\geqq 1$ we denote by $f_n(a)$ the left support of $(1-a)^n$. Put $f_\infty(a) = \inf_n f_n(a)$. Lemma 3 and the results of BREUER ([2], Theorems 1 and 2) imply the following theorem, an analogue of the Riesz decomposition theorem in the theory of compact operators:

Theorem 3. *Let M be a W^*-algebra, $\mathcal{J}$ the norm-closed two-sided ideal generated by all finite projections in M, and a an element of $\mathcal{J}$. Let the projections $e_n(a)$, $f_n(a)$, be defined as above. Then*

(i) *$e_\infty(a)$ is a finite projection;*

(ii) *for every $1\leqq n<\infty$ there exist $x_n\in M$ and a finite projection $p_n\in M$ such that $(1-a)^n x_n = 1-p_n$;*

(iii) *$ae_\infty(a)=e_\infty(a)ae_\infty(a)$ and $af_\infty(a)=f_\infty(a)af_\infty(a)$;*

(iv) *$e_\infty(a)\wedge f_\infty(a) = 0$ and $e_\infty(a)\vee f_\infty(a) = 1$;*

(v) *for every $1\leqq n\leqq\infty$ we have $e_n(a) \sim e_n(a^*) = 1-f_n(a)$.*

References

[1] C. Apostol and L. Zsidó, Ideals in W^*-algebras and the function η of A. Brown and C. Pearcy, *Revue Roum. Math. Pures et Appl.* (to appear).

[2] M. Breuer, Fredholm theories in von Neumann algebras I, *Math. Annalen*, **178** (1968), 243—254.

[3] J. Dixmier, Sur certains espaces considérés par M. H. Stone, *Summa Brasil. Math.*, **2** (1951), 151—182.

[4] J. Dixmier, Formes linéaires sur un anneau d'opérateurs. *Bull. Soc. Math. France*, **81** (1953), 9—39.

[5] J. Dixmier, *Les algèbres d'opérateurs dans l'espace Hilbertien (Algèbres de von Neumann)*, deuxième édition, Gauthier-Villars (Paris, 1969).

[6] V. I. Ovtshinnikov, On completely continuous operators relative to a von Neumann algebra (in Russian), *Funkcionalniĭ anal. i pril.*, **6** (1972), 37—40.

[7] F. Riesz et B. Sz.-Nagy, *Leçons d'analyse fonctionnelle*, 3e éd. Gauthier-Villars et Akadémiai Kiadó (Paris—Budapest, 1955).

(Received July 28, 1972)

On the consequences of permutation identities

By G. POLLÁK in Szeged

To Professor B. Szőkefalvi-Nagy on his 60th birthday

The aim of this note is to give a description of all permutation identities valid in a permutative semigroup [2]. YAMADA [4] was the first to consider permutation identities in semigroups. The best result in the field was attained by PERKINS [2] who proved that any commutative semigroup variety is finitely based. In the same work he gives an example showing that no similar proposition holds for varieties satisfying $xyzt=xzyt$. On the other hand, any permutative semigroup variety satisfying an identity of the form $x^{m+d}=x^m$ is finitely based. We give another class of (hereditary) finitely based varieties. As a matter of fact, this can be obtained from a result of PUTCHA and YAQUB [3] claiming that a semigroup in which a permutation identity of rather general type holds satisfies all permutation identities for products containing sufficiently many factors. From our results it would be easy to determine exactly the necessary number of factors, and to give a "standard" form of finite bases of identities (up to bases of permutation groups).

1. The consequence group

Following YAMADA [4], we call an identity of the form

(1) $$x_1 \dots x_n = x_{1\sigma} \dots x_{n\sigma}$$

a *permutation identity* if σ is a permutation of the set $\{1, \dots, n\}$. The number n will be called the *length* of identity (1).

Let $\mathfrak{S}$ be a semigroup variety. Denote the set of all permutation identities of length n which hold in $\mathfrak{S}$ by G_n and the set of the corresponding permutations by Γ_n. Obviously, Γ_n is a subgroup of the symmetric group Σ_n. The set of permutation identities of length $n+1$ which follow from G_n will be denoted by G_n' and the corresponding set of permutations by Γ_n'. Again, Γ_n' is a group called the *(first) consequence group of* Γ_n. The *qth consequence group* $\Gamma_n^{(q)}$ can be defined in a similar way through the permutation identities of length $n+q$ which follow from G_n. We remark though trivial that $\Gamma_n^{(q_1+q_2)}=\Gamma_n^{(q_1)(q_2)}$.

Our main task consists in finding out how Γ'_n depends on Γ_n. For this purpose we shall first look for a comfortable system of generators of Γ'_n.

Suppose (1) holds in $\mathfrak{S}$. The subsequent $n+2$ identities follow immediately:

$$x_1 \dots x_{n+1} = x_{1\sigma} \dots x_{n\sigma} x_{n+1},$$

$$x_1 \dots x_{n+1} = x_1 x_{1\sigma+1} \dots x_{n\sigma+1}, \tag{2}$$

$$x_1 \dots x_{n+1} \equiv u_1^{(i)} \dots u_n^{(i)} = u_{1\sigma}^{(i)} \dots u_{n\sigma}^{(i)} \qquad (i = 1, \dots n)$$

where

$$u_j^{(i)} = \begin{cases} x_j & \text{if } j < i, \\ x_i x_{i+1} & \text{if } j = i, \\ x_{j+1} & \text{if } j > i. \end{cases}$$

The corresponding elements $\sigma', \sigma'', \lambda_1, \dots, \lambda_n$ of the consequence group are given by the equations

$$j\sigma' = \begin{cases} j\sigma & \text{if } j \leqq n, \\ n+1 & \text{if } j = n+1 \end{cases} \tag{3_1}$$

$$j\sigma'' = \begin{cases} 1 & \text{if } j = 1, \\ (j-1)\sigma+1 & \text{if } 2 \leqq j \leqq n+1; \end{cases} \tag{3_2}$$

$$j\lambda_i = \begin{cases} j\sigma & \text{if } j \leqq i\sigma^{-1}, \quad j\sigma \leqq i, \\ j\sigma+1 & \text{if } j < i\sigma^{-1}, \quad j\sigma > i, \\ (j-1)\sigma & \text{if } j > i\sigma^{-1}, \quad (j-1)\sigma < i, \\ (j-1)\sigma+1 & \text{if } j > i\sigma^{-1}, \quad (j-1)\sigma \geqq i \end{cases} \tag{3_3}$$

for $i=1, \dots, n$.

Lemma 1. *The consequence group Γ'_n of Γ_n is generated by the elements* (3_1), (3_2), (3_3) *where σ ranges over Γ_n.*

It suffices to show that all identities in G'_n are consequences of the identities (2) where σ ranges over Γ_n. Now let

$$x_1 \dots x_{n+1} = x_{1\tau} \dots x_{(n+1)\tau} \tag{4}$$

be an identity in G'_n, i.e. a consequence of G_n. This means that there exists a sequence of words $(x_1 \dots x_{n+1} \equiv) a_0, a_1, \dots, a_k (\equiv x_{1\tau} \dots x_{(n+1)\tau})$ such that $a_r \equiv b_r u_1^{(r)} \dots u_n^{(r)} c_r$, $a_{r+1} \equiv b_r u_{1\sigma(r)}^{(r)} \dots u_{n\sigma(r)}^{(r)} c_r$ where b_r, c_r are arbitrary and $u_j^{(r)}$ nonempty words, $\sigma(r) \in \Gamma_n$. Denote the length of the word y by $l(y)$. Then $l(a_r)=l(a_{r+1})$ for all $r<k$ and thus, by induction, $l(a_r)=n+1$. On the other hand $l(a_r)=l(b_r)+\sum_{j=1}^{n} l(u_j^{(r)})+l(c_r)$, and, since $l(u_j^{(r)})>0$, there are only three possibilities: 1) $l(b_r)=0$, $l(c_r)=1$, $l(u_1^{(r)})=\cdots$

$\dots = l(u_n^{(r)}) = 1$ and $a_r = a_{r+1}$ follows from an identity of type (2_1); 2) $l(b_r) = 1$, $l(c_r) = 0$, $l(u_1^{(r)}) = \dots = l(u_n^{(r)}) = 1$ and $a_r = a_{r+1}$ follows from an identity of type (2_2); 3) $l(b_r) = l(c_r) = 0$, $l(u_i^{(r)}) = 2$ for exactly one i, $l(u_j^{(r)}) = 1$ for $j \neq i$ and $a_r = a_{r+1}$ follows from one of the identities (2_3), q.e.d.

The permutations (3) are not very easy to handle, therefore we shall use the system σ', $\lambda_i \lambda_{i+1}^{-1}$ $(1 \leqq i \leqq n-1)$, $\lambda_n \sigma'^{-1}$, $\sigma'' \lambda_1^{-1}$, equivalent to (3), instead. Introduce the notation

$$\gamma(i, j) = \begin{cases} (i\, i+1 \dots j) & \text{if } i \leqq j, \\ (i\, i-1 \dots j) & \text{if } i > j. \end{cases}$$

Thus, $\gamma(j, i) = \gamma(i, j)^{-1}$. It is straightforward to check the formulae

$$(5) \qquad \begin{aligned} \lambda_i &= \gamma(n+1, i\sigma^{-1})\sigma'\gamma(i, n+1) \quad \text{for } 1 \leqq i \leqq n, \\ \sigma'' &= \gamma(n+1, 1)\sigma'\gamma(1, n+1). \end{aligned}$$

Hence

$$(6) \qquad \begin{aligned} \lambda_i \lambda_{i+1}^{-1} &= \gamma(n+1, i\sigma^{-1})\sigma'\gamma(i, n+1)\gamma(n+1, i+1)\sigma'^{-1}\gamma((i+1)\sigma^{-1}, n+1) = \\ &= \gamma(n+1, i\sigma^{-1})\sigma' \cdot (i\, n+1) \cdot \sigma'^{-1}\gamma((i+1)\sigma^{-1}, n+1) = \\ &= \gamma(n+1, i\sigma^{-1}) \cdot (i\sigma^{-1}\, n+1) \cdot \gamma((i+1)\sigma^{-1}, n+1) = \\ &= \gamma(n+1, i\sigma^{-1}+1)\gamma((i+1)\sigma^{-1}, n+1) = \gamma((i+1)\sigma^{-1}, i\sigma^{-1}+1) \end{aligned}$$

for $1 \leqq i \leqq n-1$ and

$$(6') \qquad \begin{aligned} \lambda_n \sigma'^{-1} &= \gamma(n+1, n\sigma^{-1})\sigma'\gamma(n, n+1)\sigma'^{-1} = \gamma(n+1, n\sigma^{-1}) \cdot (n\sigma^{-1}\, n+1) = \\ &= \gamma(n+1, n\sigma^{-1}+1), \end{aligned}$$

$$(6'') \qquad \sigma'' \lambda_1^{-1} = \gamma(n+1, 1)\sigma'\gamma(1, n+1)\gamma(n+1, 1)\sigma'^{-1}\gamma(1\sigma^{-1}, n+1) = \gamma(1\sigma^{-1}, 1).$$

Remark that, by (3_1), (3_2) and (5),

$$(7) \qquad \begin{aligned} ((i+1)\sigma^{-1})\sigma'\sigma''^{-1} &= i\sigma^{-1}+1 \quad \text{for } 1 \leqq i \leqq n-1, \\ (n+1)\sigma'\sigma''^{-1} &= n\sigma^{-1}+1, \\ (1\sigma^{-1})\sigma'\sigma''^{-1} &= 1, \end{aligned}$$

and, since the symbols $n+1$, $1\sigma^{-1}$, $(i+1)\sigma^{-1}$ $(1 \leqq i \leqq n-1)$ are exactly the integers $1, \dots, n+1$ in a different order, we have obtained

Lemma 2. *Γ_n' is generated by the elements σ', $\gamma(i, i\sigma'\sigma''^{-1})$ $(i = 1, \dots, n+1)$ where σ ranges over Γ_n.*

The subgroup of Γ_n' generated by the cycles $\gamma(i, i\sigma'\sigma''^{-1})$ $(1 \leqq i \leqq n+1)$ will be denoted by Γ_n^*. As a generalization of Lemma 2, we have

Lemma 2'. *If*

(8) $$\sigma'\sigma''^{-1} = \nu_1 \dots \nu_s$$

is the decomposition of $\sigma'\sigma''^{-1}$ into disjoint cycles for some $\sigma \in \Gamma_n$ and i, j occur in the same ν_t then $\gamma(i,j) \in \Gamma_n^*$.

Indeed, for some power of $\sigma'\sigma''^{-1}$ we have $i(\sigma'\sigma''^{-1})^c = j$. If $c=1$ then $\gamma(i,j) \in \Gamma_n^*$ by its definition. Now let $c>1$ and suppose the assertion holds for $c-1$. Put $i(\sigma'\sigma''^{-1})^{c-1} = k$; then $\gamma(i,k) \in \Gamma_n^*$, $\gamma(k,j) = \gamma(k, k\sigma'\sigma''^{-1}) \in \Gamma_n^*$ and hence $\gamma(i,j) = \gamma(k,j) \cdot \gamma(i,k) \in \Gamma_n^*$.

2. Consequence groups of *f*-irreducible groups

The following subgroups of the symmetric group Σ_n will take important roles in what follows (A_n denotes, as usual, the alternating group):

$$\Sigma_{n,k} = \{\sigma | i\sigma = i \text{ for } i \geqq k\}, \qquad \bar{\Sigma}_{n,k} = \{\sigma | i\sigma = i \text{ for } i \leqq k\}, \qquad \Phi_k = \Sigma_{n,k} \otimes \bar{\Sigma}_{n,k},$$

$$\Sigma_n^{(e)} = \{\sigma | i\sigma = i \text{ for odd } i\}, \qquad \Sigma_n^{(o)} = \{\sigma | i\sigma = i \text{ for even } i\},$$

$$\Sigma_n^{(p)} = \Sigma_n^{(e)} \otimes \Sigma_n^{(o)} = \{\sigma | i\sigma \equiv i \pmod 2\}, \qquad A_n^{(p)} = \Sigma_n^{(p)} \cap A_n.$$

Observe that $\sigma \in \Phi_k$ iff the images $i\sigma$ of elements $i \leqq k$ precede those of elements $i \geqq k$ (in particular, $k\sigma = k$). Remark also $\Sigma_{n,n+1} = \bar{\Sigma}_{n,0} = \Sigma_n$.

The role of $A_n^{(p)}$ is clear from

Lemma 3. *$\Gamma_n' \subseteqq A_{n+1}$ iff $\Gamma_n \subseteqq A_n^{(p)}$.*

Proof. If $\Gamma_n \subseteqq A_n^{(p)}$ then σ', σ'' and $\sigma'\sigma''^{-1}$ are contained in $A_{n+1}^{(p)}$ for all $\sigma \in \Gamma_n$. Thus, $i \equiv i\sigma'\sigma''^{-1} \pmod 2$ for every i and $\gamma(i, i\sigma'\sigma''^{-1}) \in A_{n+1}$.

Conversely, suppose $\Gamma_n \nsubseteqq A_n^{(p)}$ and let $\sigma \in \Gamma_n \setminus A_n^{(p)}$. If $\sigma \notin A_n$ then $\sigma' \notin A_{n+1}$. If $\sigma \notin \Sigma_n^{(p)}$ suppose i is the least natural number such that

(9) $$i\sigma^{-1} \not\equiv i \pmod 2.$$

Then

(10) $$\gamma\left(i\sigma^{-1}, (i\sigma^{-1})\sigma'\sigma''^{-1}\right) \notin A_{n+1}.$$

Indeed, for $i=1$ we have $(1\sigma^{-1})\sigma'\sigma''^{-1} = 1$ and (10) follows from (9). If $i>1$ then $(i\sigma^{-1})\sigma'\sigma''^{-1} = (i-1)\sigma^{-1} + 1 \equiv i-1+1 = i \pmod 2$ and therefore $i\sigma^{-1} \not\equiv (i\sigma^{-1})\sigma'\sigma''^{-1} \pmod 2$ which proves (10).

The permutation group Γ_n will be called *fixelement-reducible* or *f-reducible* if $\Gamma_n \subseteqq \Phi_k$ for some $k \leqq n$, and *fixelement-irreducible (f-irreducible)* in the opposite case. Now we want to investigate the case where Γ_n is f-irreducible.

Lemma 4. *If Γ_n is f-irreducible then for every k ($1 \leqq k \leqq n$) there exist symbols i, j such that $i \leqq k < j$, $\gamma(i,j) \in \Gamma_n^*$.*

Proof. Since Γ_n is f-irreducible there exists $\sigma \in \Gamma_n \setminus \Phi_k$. If $k=1$ this means $1\sigma \neq 1$, so that $1\sigma^{-1} \neq 1$ and, by virtue of (7_3) and Lemma 2, we have $\gamma(1, 1\sigma^{-1})$ $(=\gamma(1\sigma^{-1}, 1)^{-1}) \in \Gamma_n^*$. Now put $k \geqq 2$. Then there exist elements i, l with $1 \leqq i \leqq k \leqq$ $\leqq l \leqq n$, $i\sigma > l\sigma$ (and therefore $i \neq 1\sigma^{-1}$). It is easy to see that one can even suppose $l\sigma = i\sigma - 1$. Now by (7) $i\sigma'\sigma''^{-1} = (i\sigma - 1)\sigma^{-1} + 1 = l+1$, so that $\gamma(i, l+1) \in \Gamma_n^*$, and the lemma is proved.

Corollary 2. *If Γ_n is f-irreducible then Γ_n^* is transitive.*

Indeed, $k\gamma(i,j) = k+1$; thus, every symbol ($< n+1$) can be carried over to every greater symbol and, taking into account the inverses of the γ's, it can be carried over to every element.

This corollary is majorized by the following lemma. The proof of the lemma, however, relies upon the corollary itself.

Lemma 5. *If Γ_n is f-irreducible then Γ_n^* is doubly transitive.*

Proof. Since Γ_n^* is already known to be transitive, we have to prove only that for every k ($1 < k < n+1$) there exists a permutation $\varrho_k \in \Gamma_n^*$ such that

$$k\varrho_k = k+1, \quad 1\varrho_k = 1. \tag{11}$$

Recall that $1\sigma^{-1} \neq 1, n+1$ for some $\sigma \in \Gamma_n$. If $\gamma(1, n+1) \in \Gamma_n^*$ then it has a power such that the permutation $\varrho_k = \gamma(n+1, 1)^{r_k} \cdot \gamma(1, 1\sigma^{-1})\gamma(1, n+1)^{r_k}$ satisfies (11) (for this, choose $k - 1\sigma^{-1} < r_k \leqq \min(k-1, n+1-1\sigma^{-1})$). For the rest of the proof suppose $\gamma(1, n+1) \notin \Gamma_n^*$. By Lemma 4, there exist symbols l, m such that $l \leqq k < m$, $\gamma(l, m) \in \Gamma_n^*$. If $1 < l$ put $\varrho_k = \gamma(l, m)$. If $l=1$ then, by assumption, $m < n+1$. Thus, there exist i, j with $i \leqq m < j$, $\gamma(i,j) \in \Gamma_n^*$. Put $\tau = \gamma(j, i)\gamma(1, m)\gamma(i, j)$ and

$$\varrho_k = \begin{cases} \tau\,(=\gamma(2, m+1)) & \text{if} \quad i = 1, \\ \gamma(i,j) & \text{if} \quad 1 < i \leqq k, \\ \gamma(1,m)\tau^{-1}\big(=(k\,k+1)(m\,m+1)\big) & \text{if} \quad i = k+1, \\ \tau^{-1}\gamma(1,m)\tau\big(=(i\,i+1)\,\gamma(2, m+1)(i\,i+1)\big) & \text{else.} \end{cases}$$

This proves our lemma.

Corollary 3. *If Γ_n is f-irreducible then Γ_n^* is primitive.*

Now we formulate the basic

Theorem 1. *If Γ_n is f-irreducible then its consequence group Γ_n' is* 1) *the subgroup Δ' of Σ_6 generated by $\gamma(1,4)$ and $\gamma(3,6)$ if $n=5$, $\Gamma_5 = \Delta = \{(1), (14)(25)\}$;* 2) *$A_{n+1}$ if $\Gamma_n \subseteqq A_n^{(p)}$;* 3) *$\Sigma_{n+1}$ else.*

Remark. Δ' is a group isomorphic to Σ_5. It can be obtained from the subgroup $\Sigma_{6,6}$ of Σ_6 (having 6 for invariant symbol) by an outer automorphism of the latter one.

Proof. The fact that the consequence group of Δ is Δ' can be checked by a straightforward calculation. Remark only that for $\sigma=(14)(25)$ we have $\sigma'\sigma''^{-1}==(14)(36)$ so that $\Gamma_n^*=\Delta'$ and $\sigma'=\gamma(6,3)\gamma(1,4)^2\gamma(3,6)\in\Gamma_n^*$.

In virtue of Lemma 3, all we need to prove is that Δ is the only f-irreducible group the consequence group of which does not contain the alternating group. In doing this we shall rely upon the following facts (see e.g. [1]):

I. If a subgroup Π_{m-q} of Σ_m has q invariant symbols and is transitive and primitive on the rest then any primitive subgroup of Σ_m which contains Π_{m-q} is $(q+1)$-fold transitive ([1], Theorem 5. 6. 2).

II. For $m>12$, $t>3\sqrt{m}-2$, the only t-fold transitive subgroups of Σ_m are Σ_m and A_m ([1], p. 68.).

III. If $m=kp+r$ where p is prime, $p>k$, $r>k$, $r>2$ then the only $(r+1)$-fold transitive subgroups of Σ_m are Σ_m and A_m ([1], Theorem 5. 7. 2).

Suppose Γ_n is f-irreducible. If Γ_n^* contains a transposition we have obviously $\Gamma_n^*=\Sigma_{n+1}$ (because of double transitivity). If Γ_n^* contains an element of the form $\gamma(k-2,k)$ then $A_{n+1}\subseteqq\Gamma_n^*$. Indeed, for $n=2$ the assertion is obvious. Let $n>2$. It suffices to show that if $k<n+1$ then $\gamma(k-1,k+1)\in\Gamma_n^*$ and if $k-2>1$ then $\gamma(k-3,k-1)\in\Gamma_n^*$ since these imply $\gamma(q-2,q)\in\Gamma_n^*$ for all $3\leqq q\leqq n+1$ and these cycles generate A_{n+1}. Let us prove the first part; the other one can be treated analogously. By Lemma 4, there exists a cycle $\gamma(i,j)\in\Gamma_n^*$ such that $i\leqq k\leqq j$. If $i\leqq k-2$ then $\gamma(j,i)\gamma(k-2,k)\gamma(i,j)=\gamma(k-1,k+1)\in\Gamma_n^*$. If $i=k-1$ or $i=k$ then $\gamma(j,i)\gamma(k-2,k)\gamma(i,j)=(k-2\;k\;k+1)$ or $(k-2\;k-1\;k+1)$, respectively. However, since $\gamma(k,k-2)\cdot(k-2\;k\;k+1)\cdot\gamma(k-2,k)=(k+1\;k-1\;k-2)$, in both cases we have $(k-1\;k\;k+1)=(k+1\;k-1\;k-2)(k-2\;k\;k+1)\in\Gamma_n^*$.

Now consider two different cycles $\gamma(i_1,j_1)$, $\gamma(i_2,j_2)\in\Gamma_n^*$ (e.g. $\gamma(1,1\sigma^{-1})$ for $1\sigma\neq1$ and $\gamma(n\tau^{-1}+1,n+1)$ for $n\tau\neq n$). We may suppose $i_1<j_1$, $i_1\leqq i_2<j_2$. If, moreover, $j_2<j_1$ then $\gamma(i_2-1,j_2-1)=\gamma(i_1,j_1)\gamma(i_2,j_2)\gamma(j_1,i_1)\in\Gamma_n^*$. If $j_2=j_1$ then $\gamma(i_2+1,j_2+1)=\gamma(j_1,i_1)\gamma(i_2,j_2)\gamma(i_1,j_1)\in\Gamma_n^*$. In both cases we have two cycles of the form $\gamma(k,l)$ and $\gamma(k-1,l-1)$. However then $\gamma(l,k)\gamma(l-1,k-1)\gamma(k,l)^2==\gamma(k-1,k+1)\in\Gamma_n^*$ and hence $\Gamma_n^*\supseteqq A_{n+1}$.

The case $i_1=i_2$, $j_1<j_2$ is symmetrical to the second subcase of the above one.

If $j_1=i_2$ then $\gamma(i_2,j_2)\gamma(i_1,j_1)=\gamma(i_1,j_2)$ and the pair $\gamma(i_1,j_2)$, $\gamma(i_2,j_2)$ gives the first case again.

If $j_1<i_2$ and $n\leqq6$ then at least one of both cycles is of length $\leqq3$ and the former argument yields $\Gamma_n^*\supseteqq A_{n+1}$ once more. Let $n>6$ (i.e. $n+1\geqq8$). One of both cycles is of length $\leqq\dfrac{n+1}{2}$; denote this one by $\gamma(i,j)$. If $j<n+1$ take $\gamma(k,l)\in\Gamma_n^*$ with

$k \leqq j < l$ (if $j = n+1$ then $i > 1$ and we ought to demand $l < i-1 \leqq k$). The permutations $\gamma(i, j)$, $\gamma(l, k)\gamma(i, j)\gamma(k, l)$ generate a subgroup Π_{j-i+2} of Γ_n^* having at least $\frac{n-1}{2}$ invariant symbols and being primitive on the rest. Thus, by I, Γ_n^* is $\frac{n+1}{2}$-fold transitive. If $n \geqq 27$ it holds $\frac{n+1}{2} > 3\sqrt{n+1}-2$ and $\Gamma_n^* \supseteqq A_{n+1}$ follows from II. For $7 \leqq n \leqq 26$ III allows the following maximal multiplicities of transitivity:

$r=3$ for $n+1 = 8, 9, 10, 13, 14, 16, 17, 20, 22, 25, 26$;
$r=4$ for $n+1 = 11, 15, 18, 19, 21, 23, 27$;
$r=5$ for $n+1 = 12, 24$.

In all cases $r < \frac{n+1}{2}$.

Thus, the only remaining possibility is $i_1 < i_2 < j_1 < j_2$ for every pair of cycles. Suppose there are three different cycles $\gamma(i_1, j_1)$, $\gamma(i_2, j_2)$, $\gamma(i_3, j_3) \in \Gamma_n^*$, $i_1 < i_2 < i_3 < j_1 < j_2 < j_3$. Then

$$\varrho_1 = \gamma(i_1, j_1)\gamma(j_3, i_3)\gamma(j_1, i_1)\gamma(i_3, j_3) = (i_3-1\, i_3)(j_1\, j_1+1) \in \Gamma_n^*,$$

$$\varrho_2 = \gamma(i_2, j_2)\gamma(j_3, i_3)\gamma(j_2, i_2)\gamma(i_3, j_3) = (i_3-1\, i_3)(j_2\, j_2+1) \in \Gamma_n^*,$$

$$\varrho_1\varrho_2 = (j_1\, j_1+1)(j_2\, j_2+1) \in \Gamma_n^*,$$

$$\gamma(i_1, j_1)\varrho_2\varrho_1\gamma(j_1, i_1)\varrho_1\varrho_2 = \gamma(j_1-1, j_1+1) \in \Gamma_n^*.$$

Thus, we may suppose Γ_n^* contains only two cycles: $\gamma_1 = \gamma(i_1, j_1)$ and $\gamma_2 = \gamma(i_2, j_2)$ However, this implies $\sigma'\sigma''^{-1} = (i_1\, j_1)(i_2\, j_2)$ or (1) for every $\sigma \in \Gamma_n$ (recall $\sigma'\sigma''^{-1} \in A_{n+1}$). But $\sigma'\sigma''^{-1}$ determines σ uniquely; indeed, σ' does and the conditions $\sigma'^{-1}\gamma(n+1, 1)\sigma' = \sigma'\sigma''^{-1}\gamma(n+1, 1)$, $(n+1)\sigma' = n+1$ determine σ'. Hence Γ_n is a two-element group and $\sigma(\neq 1)$ is of order 2. As Γ_n is f-irreducible, $1\sigma \neq 1$, $n\sigma \neq n$ and, consequently, $i_1 = 1$, $j_2 = n+1$, $j_1 = 1\sigma^{-1} = 1\sigma$, $i_2 = n\sigma+1$ and $\sigma'\sigma''^{-1} = (1\, j_1)\dots(i_2-1\; n)(2\; j_1+1)\dots(i_2\; n+1) = (1\, j_1)(i_2\; n+1)$. Hence one finds by a routine induction $\sigma = (1\, j_1)(2\; j_1+1)\dots(n-j_1+1\; n)$; $i_2 = n-j_1+2$. If $j_1 = n$ then $i_2 = 2$ and $\gamma_2^{-1}\gamma_1^{-1}\gamma_2^2 = (1\,2\,3)$. Now let $j_1 < n$, $i_2 > 2$. Form the following elements:

$$\alpha = \gamma_2^{-1}\gamma_1\gamma_2 = (1 \dots i_2-1\; i_2+1 \dots j_1+1),$$

$$\beta = \gamma_2^{-1}\gamma_1\gamma_2^2 = (1 \dots i_2-1\; i_2+2 \dots j_1+2),$$

$$\delta = \alpha\beta^{-1} = (i_2-1\; i_2+1)(j_1+1\; j_1+2),$$

$$\varepsilon = \delta^{-1}\gamma_1\delta\gamma_1^{-1} = (i_2-2\; i_2)(i_2-1\; i_2+1),$$

and, if $i_2 > 3$,

$$\zeta = \gamma_1\varepsilon\gamma_1^{-1} = (i_2-3\; i_2-1)(i_2-2\; i_2).$$

Then

$$\varepsilon\zeta = (i_2-3\ i_2-1\ i_2+1),$$

$$\gamma_2\gamma_1\gamma_2\gamma_1^{-2}\gamma_2\varepsilon\zeta\gamma_2^{-1}\gamma_1^2\gamma_2^{-1}\gamma_1^{-1}\gamma_2^{-1} = (i_2-2\ i_2-1\ i_2) \quad \text{if} \quad i_2 = j_1-1,$$

$$\gamma_2\gamma_1^{-2}\gamma_2\varepsilon\zeta\gamma_2^{-2}\gamma_1^2\gamma_2^{-1} = (i_2-1\ i_2\ i_2+1) \quad \text{else.}$$

On the other hand, if $i_2=3, j_1>4$ put $\eta = \delta^{-1}\gamma_1^{-1}\delta\gamma_1 = (24)(35)$. Then

$$\eta\varepsilon = (1\ 3\ 5),$$

$$\gamma_2\gamma_1^{-1}\gamma_2\gamma_1^{-2}\gamma_2\gamma_1\eta\varepsilon\gamma_1^{-1}\gamma_2^{-1}\gamma_1^2\gamma_2^{-1}\gamma_1\gamma_2^{-1} = (2\ 3\ 4).$$

Thus, we have reduced the problem to the case $i_2=3, j_1=4$ corresponding to $\Gamma_n=\Delta$. The theorem is proved.

From the proof it turns out that $\Gamma_n^*\supseteqq A_{n+1}$ if $\Gamma_n\neq\Delta$ and $\Gamma_n^*=\Delta'$ if $\Gamma_n=\Delta$. Combining this with Lemma 3 and the plus information on Γ_n^* its proof comprises, we have

Colollary 4. *If Γ_n is f-irreducible then*

$$\Gamma_n^* = \begin{cases} \Delta' & \text{if } \Gamma_n = \Delta, \\ A_{n+1} & \text{if } \Gamma_n \subseteqq \Sigma_n^{(p)}, \\ \Sigma_{n+1} & \text{else.} \end{cases}$$

Thus, $\Gamma_n^*=\Gamma_n'$ or else $\Gamma_n\subseteqq\Sigma_n^{(p)}$, $\Gamma_n\nsubseteqq A_n$.

3. The general case

The case of f-reducible groups can be reduced now to that of f-irreducible ones. Let $\Gamma_n = \bigcap_{t=1}^{s} \Phi_{k_t}$ $(1\leqq k_1<\dots<k_s\leqq n)$ and $\Gamma_n\nsubseteqq\Phi_k$ for any other k. Put, furthermore, $k_0=0$, $k_{s+1}=n+1$, and denote $\Sigma_{n,k_t}\cap\bar\Sigma_{n,k_{t+1}}$ by P_t, $k_{t+1}-k_t-1$ by j_t $(t=0, \dots, s)$. Then $\Gamma_n \subseteqq \prod_{t=0}^{s} P_t$ and it is easy to see that the mapping $\varphi_t: P_t\to\Sigma_{j_t}$ defined by $j(\varrho\varphi_t) = (j+k_t)\varrho-k_t$ is an isomorphism. Denote by Λ_t the projection of Γ_n in P_t. Then Γ_n is a subdirect product of $\Lambda_0, \dots, \Lambda_s$. Obviously, $\Lambda_t\varphi_t$ is f-irreducible. Every $\sigma\in\prod_{t=0}^{s} P_t$ has a unique factorization

$$\sigma = \sigma_0\dots\sigma_s \quad (\sigma_t\in P_t); \tag{12}$$

in particular, $\sigma_t\in\Lambda_t$ if $\sigma\in\Gamma_n$. Analogously, put $P_t' \doteq \Sigma_{n+1,k_t}\cap\bar\Sigma_{n+1,k_{t+1}+1}\subseteqq\Sigma_{n+1}$; then every element of $\prod_{t=0}^{s} P_t'$ has a unique factorization

$$\tau = \tau_0\dots\tau_s \qquad (\tau_t\in P_t'). \tag{12'}$$

Define $\varphi'_t: P'_t \to \Sigma_{j_t+1}$ by the same rule as we did φ_t (only for $i \leqq j_t+1$ instead of $i \leqq j_t$).

Those subdirect factors Λ_t contained in $\Sigma_n^{(p)}$ but not in $A_n^{(p)}$ behave in a manner slightly different from the rest. Introduce therefore the notation

$$T = \{t \mid 0 \leqq t \leqq s \wedge \Lambda_t \subseteqq \Sigma_n^{(p)} \wedge \Lambda_t \nsubseteq A_n^{(p)}\}$$

and the projection $\mu: \Gamma_n \to \prod_{t \in T} \Lambda_t$ which maps σ onto $\prod_{t \in T} \sigma_t$.

The consequence group of an f-reducible group is now fully described by

Theorem 2. *Let $\Gamma_n \subseteqq \bigcap_{t=1}^{s} \Phi_{k_t}$ $(1 \leqq k_1 < \dots < k_s \leqq n)$ and $\Gamma_n \nsubseteq \Phi_k$ for any other k. Then*

$$\Gamma'_n = (\Gamma_n \mu)' \times \prod_{t \notin T} \Lambda'_t. \tag{13}$$

Furthermore, $\Lambda'_t = (\Lambda_t \varphi_t)' \varphi_t'^{-1}$ for $1 \leqq t \leqq s$ and

$$(\Gamma_n \mu)' = \left\{ \prod_{t \in T} \tau_t \mid \tau_t \in P'_t \wedge (\exists \sigma \in \Gamma_n) \, \forall_t (\tau_t \in A_{n+1} \Leftrightarrow \sigma_t \in A_n) \right\}. \tag{14}$$

Proof. Since $i\sigma = i\sigma_t$ for $k_t \leqq i \leqq \max(k_{t+1}, n)$, it follows $i\sigma' = i\sigma'_t$, $i\sigma'' = (i-1)\sigma + 1 = (i-1)\sigma_t + 1 = i\sigma''_t$ for $k_t < i \leqq k_{t+1}$. As the domains of σ'_t, σ''_t and σ'_u, σ''_u are disjoint for $t \neq u$, we have $\sigma' = \sigma'_0 \dots \sigma'_s$, $\sigma'' = \sigma''_0 \dots \sigma''_s$, $\sigma' \sigma''^{-1} = (\sigma'_0 \sigma_0''^{-1}) \dots (\sigma'_s \sigma_s''^{-1})$. Moreover, $i\sigma' \sigma''^{-1} = i\sigma'_t \sigma_t''^{-1}$ implies $\gamma(i, i\sigma' \sigma''^{-1}) = \gamma(i, i\sigma'_t \sigma_t''^{-1}) \in \Lambda'_t$. If $t \in T$ put $\bar{\sigma} = \prod_{t \in T} \sigma_t$; the same argument yields $\bar{\sigma}' = \prod_{t \in T} \sigma'_t$, $\gamma(i, i\sigma' \sigma''^{-1}) = \gamma(i, i\sigma'_t \sigma_t''^{-1}) = \gamma(i, i\bar{\sigma}' \bar{\sigma}''^{-1}) \in (\Gamma_n \mu)'$. Thus, Γ'_n is contained in the right side of (13).

Before proceeding to the converse, we turn to the statement $\Lambda'_t = (\Lambda_t \varphi_t)' \varphi_t'^{-1}$. We can see that φ'_t is an isomorphism, $(\varrho\varphi_t)' = \varrho' \varphi'_t$, $(\varrho\varphi_t)'' = \varrho'' \varphi'_t$ for $\varrho \in P_t$ and, if $k_t < i \leqq k_{t+1}$,

$$\gamma(i, i\varrho' \varrho''^{-1}) \varphi'_t = \gamma(i - k_t, i\varrho' \varrho''^{-1} - k_t) = \gamma(i - k_t, (i - k_t)((\varrho' \varrho''^{-1}) \varphi'_t) - k_t) =$$
$$= \gamma(i - k_t, (i - k_t)(\varrho\varphi_t)' (\varrho\varphi_t)''^{-1}).$$

Hence $\Lambda'_t \varphi'_t = (\Lambda_t \varphi_t)'$ and $\Lambda^*_t \varphi'_t = (\Lambda_t \varphi_t)^*$. [1]

Now $\Lambda_t \varphi_t$ is f-irreducible and therefore $(\Lambda_t \varphi_t)^*$ is generated by the cyclet $\gamma(j, j\lambda' \lambda''^{-1})$ $(1 \leqq j \leqq j_t + 1, \lambda \in \Lambda_t \varphi_t)$. The same holds for $(\Lambda_t \varphi_t)'$ if $\Lambda_t \varphi_t \nsubseteq \Sigma_{j_t}^{(p)}$ or $\Lambda_t \varphi_t \subseteqq A_{j_t}^{(p)}$, i.e. for $t \notin T$. Hence, Λ'_t is also generated by the corresponding cycles $\gamma(i, i\varrho' \varrho''^{-1})$ $(k_t < i \leqq k_{t+1}, \varrho \in \Lambda_t)$. By definition of Λ_t, there exists $\sigma \in \Gamma_n$ such thas σ_t in (12) equals to ϱ. However then $\gamma(i, i\varrho' \varrho''^{-1}) = \gamma(i, i\sigma' \sigma''^{-1}) \in \Gamma'_n$. Hence $\Lambda_t \subseteqq \Gamma'_n$.

Now suppose $t \in T$. Then $\Lambda_t \varphi_t \subseteqq \Sigma_{j_t}^{(p)}$, $\Lambda_t \varphi_t \nsubseteq A_{j_t}^{(p)}$ which imply $(\Lambda_t \varphi_t)' = \Sigma_{j_t - 1}$, $(\Lambda_t \varphi_t)^* = A_{j_t + 1}$. This last gives $\Lambda^*_t \cong A_{j_t + 1}$ and so $\Lambda^*_t = P'_t \cap A_n$. As in the fore-

[1]) Remark that this immediately gives $\Gamma^*_n = \prod_{t=0}^{s} \Lambda^*_t$.

going paragraph for Λ_t', now one can verify $\Lambda_t^*\subseteqq\Gamma_n'$. Thus $\Gamma_n'\supseteqq\prod_{t\in T}\Lambda_t^*\cong\prod_{t\in T}A_{j_t+1}$. Furthermore, $(\Gamma_n\mu)'$ is generated by the elements of Λ_t^* $(t\in T)$ and by those of the form $(\sigma\mu)'$ $(\sigma\in\Gamma_n)$. Since $\Lambda_t^*\subseteqq A_{n+1}$ and $(\sigma\mu)_t=\sigma_t'$ for $t\in T$, these generators and hence $(\Gamma_n\mu)'$, too, are contained in the right side of (14). On the other hand if for $\tau=\prod_{t\in T}\tau_t$, $\tau_t\in P_t'$ there exists $\sigma\in\Gamma_n$ such that $\tau_t\in A_{n+1}\Leftrightarrow\sigma_t\in A_n$ then $\lambda=(\sigma\mu)'^{-1}\tau\in\in\prod_{t\in T}(P_t'\cap A_{n+1})=\prod_{t\in T}\Lambda_t^*$ which is obviously contained in the right side of (14); so is $(\sigma\mu)'$ and hence the same holds for $\tau=(\sigma\mu)'\lambda$.

Finally, we have seen $\Lambda_t^*\subseteqq\Gamma_n'$ for all t; thus, in order to prove $(\Gamma_n\mu)'\subseteqq\Gamma_n'$ it suffices to show $(\sigma\mu)'\in\Gamma_n'$ for $\sigma\in\Gamma_n$. But $\sigma'=(\sigma\mu)'\prod_{t\notin T}\sigma_t'$ and $\prod_{t\notin T}\sigma_t'\in\Gamma_n'$ has been proved earlier. This completes the proof.

It follows from this theorem that the consequence group of an arbitrary group is a direct product of a certain number of permutation groups of type Δ' and of one further group which is an extension of a direct product of alternating groups by an elementary 2-group. The second consequence group is a direct product of symmetric groups.

Suppose $l\geqq 1$ is the maximal number of consecutive integers $1<k, k+1, \dots$ $\dots, k+l-1<n$ such that $\Gamma_n\subseteqq\bigcap_{t=0}^{l-1}\Phi_{k+t}$, $\Gamma_n\nsubseteqq\Phi_{k-1}$, $\Gamma_n\nsubseteqq\Phi_{k+l}$. Then the $(l+1)$st consequence group is the first one being isomorphic to a symmetric group and if neither 1 nor n is invariant under Γ_n then $\Gamma_n^{(l+1)}$ *is* a symmetric group. This last holds for either Γ_n' or Γ_n'' if Γ_n is f-irreducible.

4. A class of finitely based varieties

The results of this part follow already from the theorem of Putcha and Yaqub cited in the introduction. However, in order to simplify the proof, we shall make use of a result of Perkins, too.

First of all remark that *every variety defined by a set of permutation identities is finitely based.* This follows immediately from the above results.

Combining this fact with the result of Perkins already mentioned, claiming that any uniformly periodic permutative variety is finitely based, we obtain

Theorem 3. *Let $\mathfrak{S}$ be a semigroup variety such that in $\mathfrak{S}$ hold two (not necessarily different) permutation equations*

$$x_1\dots x_n=x_{1\sigma}\dots x_{n\sigma}\qquad(\sigma\in\Sigma_n,\quad 1\sigma\neq 1),$$

$$x_1\dots x_m=x_{1\tau}\dots x_{m\tau}\qquad(\tau\in\Sigma_m,\quad m\tau\neq m).$$

Then $\mathfrak{S}$ is finitely based.

Proof. Put $N = \max(m, n)$; then neither 1 nor N are invariant under Γ_N. Thus, there exists a number l such that the l-th consequence group of Γ_N is Σ_{N+l}. Now there are two cases. Either all identities in $\mathfrak{S}$ are balanced[2]); then the identities of length $< N+l$ form a base. Or a non-balanced identity holds in $\mathfrak{S}$; however then holds an identity of the form $x^{k+d} = x^k$, too, and $\mathfrak{S}$ is finitely based in virtue of Perkins's result.

References

[1] M. HALL, *The theory of groups* (New York, 1959).
[2] P. PERKINS, Bases for equational theories of semigroups, *J. Algebra,* **11** (1968), 298—314.
[3] M. S. PUTCHA—A. YAQUB, Semigroups satisfying permutation identities, *Semigroup Forum,* **3** (1971), 68—73.
[4] M. YAMADA, *The structure of separative bands,* Dissertation (Univ. of Utah, 1962).

(Received October 24, 1972)

[2]) That is, each variable occurs on both sides the same number of times.

On logarithmic concave measures and functions

By ANDRÁS PRÉKOPA in Budapest

Dedicated to Béla Szőkefalvi-Nagy on his 60th birthday

1. Introduction. The purpose of the present paper is to give a new proof for the main theorem proved in [3] and develop further properties of logarithmic concave measures and functions. Having in mind the applications of our theory to mathematical programming, we restrict ourselves to functions and measures in finite dimensional Euclidean spaces.

A function f defined on R^n is said to be logarithmic concave if for every pair of vectors $\mathbf{x}_1, \mathbf{x}_2 \in R^n$ and for every $0 < \lambda < 1$ we have

$$f(\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \geqq (f(\mathbf{x}_1))^\lambda (f(\mathbf{x}_2))^{1-\lambda}. \tag{1.1}$$

A measure defined on the measurable subsets of R^n is logarithmic concave if for every pair A, B of convex subsets of R^n and for every $0 < \lambda < 1$, we have the following inequality

$$P(\lambda A + (1-\lambda) B) \geqq (P(A))^\lambda (P(B))^{1-\lambda}, \tag{1.2}$$

where the sign + means Minkowski addition of sets.

If the function f is logarithmic concave in R^n and $f \not\equiv 0$ then it can be written as $f(\mathbf{x}) = e^{-Q(\mathbf{x})}$ $(\mathbf{x} \in R^n)$ where $Q(\mathbf{x})$ is convex in the entire space and the value $+\infty$ is also allowed for the function Q. The set where f is positive, is convex and f is clearly continuous in the interior of this set.

The above-mentioned main theorem is repeated below in its original form.

Theorem 1. *Let Q be a convex function defined on the entire n-dimensional space. Suppose that $Q(\mathbf{x}) \geqq a$ where a is some real number. Let $\psi(z)$ be a function defined on the infinite interval $[a, \infty)$. Suppose that $\psi(z)$ is nonnegative, nonincreasing, differentiable and $-\psi'(z)$ is logarithmic concave. Consider the function $f(\mathbf{x}) = \psi(Q(\mathbf{x}))$ $(\mathbf{x} \in R^n)$ and suppose that it is a probability density i.e.*

$$\int_{R_n} f(\mathbf{x})\, d\mathbf{x} = 1. \tag{1.3}$$

Denote by $P(C)$ the integral of $f(\mathbf{x})$ over the measurable subset C of R^n. If A and B are any two convex sets in R^n and $0<\lambda<1$, then the inequality (1. 2) *holds.*

This theorem remains true without the assumption (1. 3). In fact the theorem is obviously true if the integral on the left hand side in (1. 3) is an arbitrary non-negative number. If this integral equals infinity then first we apply the theorem for the following function

$$f_T(\mathbf{x}) = f(\mathbf{x}) \text{ if } \|\mathbf{x}\| < T \text{ and } f_T(\mathbf{x}) = 0 \text{ otherwise},$$

where T is a positive number. The integral of f_T is finite over the space R^n hence we have (1. 2) with the measure P_T generated by f_T. Since

$$P(C) = \lim_{T\to\infty} P_T(C)$$

for every measurable set $C \subset R^n$, the inequality (1. 2) is satisfied with the measure generated by f too.

A second remark concerning Theorem 1 is the following: any function $\psi(z)$ satisfying the requirements of the theorem is itself logarithmic concave. In fact the finiteness of the integral of the function $\psi(Q(\mathbf{x}))$ over the space R^n implies that $\lim_{z\to\infty} \psi(z)=0$, hence

$$\psi(z) = \int_z^\infty [-\psi'(x)]\,dx \qquad (z \geqq a). \tag{1.4}$$

Consider the measure defined on the measurable subsets of R^1 generated by the logarithmic concave function

$$g(x) = -\psi'(x) \quad \text{if} \quad x \geqq a \quad \text{and} \quad g(x)=0 \quad \text{otherwise}.$$

The logarithmic concavity of this function implies that (see Theorem 3 in [3]) for any interval A of R^1, the following function of the variable z

$$\int_{A+z} g(x)\,dx \quad (-\infty < z < \infty) \tag{1.5}$$

is logarithmic concave. Since the functions (1.4) and (1. 5) coincide for $z \geqq a$, if $A=[0, \infty)$, our statement is proved. The function $\psi(z)$ can be written as

$$\psi(z) = e^{-s(z)} \, (z \geqq a),$$

where $s(z)$ is convex and nondecreasing. Any convex and nondecreasing function of a convex function is also convex hence $s(Q(\mathbf{x}))= S(\mathbf{x})$ is a convex function in R^n and

$$f(\mathbf{x}) = e^{-S(\mathbf{x})} \, (\mathbf{x} \in R^n).$$

In view of these two remarks, Theorem 1 can be reformulated in the following form, including the case of unbounded measures.

Theorem 2. *If the measure P, defined on the measurable subsets of R^n, is generated by a logarithmic concave function, then the measure P is also logarithmic concave.*

2. Integral inequalities. The proof of Theorem 1 published in [3] is based on the theorem of Brunn and on the following integral inequality proved also in [3]: if f, g are nonnegative and Borel measurable functions defined on R^1 and

$$r(t) = \sup_{x+y=2t} f(x)g(y) \qquad (-\infty < t < \infty)$$

then we have

$$\int_{-\infty}^{\infty} r(t)\,dt \geqq \left(\int_{-\infty}^{\infty} f^2(x)\,dx\right)^{\frac{1}{2}} \left(\int_{-\infty}^{\infty} g^2(y)\,dy\right)^{\frac{1}{2}}, \tag{2.1}$$

where the value $+\infty$ is also allowed for the integrals.

L. LEINDLER generalized this inequality in the following manner [2]. Let $f_1, \ldots, f_k$ be nonnegative and Borel measurable functions defined on R^1 and define the function $r(t)$ $(t \in R^1)$ by the equality

$$r(t) = \sup_{\lambda_1 x_1 + \cdots + \lambda_k x_k = t} f_1(x_1) \ldots f_k(x_k),$$

where $\lambda_1, \ldots, \lambda_k$ are positive constants satisfying the equality $\lambda_1 + \cdots + \lambda_k = 1$. Then the function $r(t)$ $(t \in R^1)$ is also Borel measurable and we have the following inequality:

$$\int_{\infty-}^{\infty} r(t)\,dt \geqq \left(\int_{-\infty}^{\infty} f_1^{\frac{1}{\lambda_1}}(x_1)\,dx_1\right)^{\lambda_1} \ldots \left(\int_{-\infty}^{\infty} f_k^{\frac{1}{\lambda_k}}(x_k)\,dx_k\right)^{\lambda_k}. \tag{2.2}$$

Now we generalize the inequality (2. 2) for functions of n variables. This generalisation is formulated in the following

Theorem 3. *Let $f_1, \ldots, f_k$ be nonnegative and Borel measurable functions defined on R^n and let*

$$r(\mathbf{t}) = \sup_{\lambda_1 \mathbf{x}_1 + \cdots + \lambda_k \mathbf{x}_k = \mathbf{t}} f_1(\mathbf{x}_1) \ldots f_k(\mathbf{x}_k) (\mathbf{t} \in R^n),$$

where $\lambda_1, \ldots, \lambda_k$ are positive constants satisfying the equality $\lambda_1 + \cdots + \lambda_k = 1$. Then the function $r(\mathbf{t})$ $(\mathbf{t} \in R^n)$ is also Borel measurable and we have the following inequality:

$$\int_{R^n} r(\mathbf{t})\,d\mathbf{t} \geqq \left(\int_{R^n} f_1^{\frac{1}{\lambda_1}}(\mathbf{x}_1)\,d\mathbf{x}_1\right)^{\lambda_1} \ldots \left(\int_{R^n} f_k^{\frac{1}{\lambda_k}}(\mathbf{x}_k)\,d\mathbf{x}_k\right)^{\lambda_k}. \tag{2.3}$$

Proof. The proof of the measurability of the function $r(\mathbf{t})$ ($\mathbf{t}\in R^{(n)}$) goes in an entirely similar way as in the case of $n=1$, $k=2$ (see the proof of Theorem 1 in [3]).

We prove (2. 3) by induction. Suppose that it holds for $n-1$ and prove that it holds for n. Let $x_{1i}, x_{2i}, \dots, x_{ki}, t_i$ ($i=1, \dots, n$) denote the components of the vectors $\mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{t}$, respectively. Fixing the second, ..., nth components so that

$$t_i = \lambda_1 x_{1i} + \lambda_2 x_{2i} + \cdots + \lambda_k x_{ki} \qquad (i = 2, \dots, n), \tag{2. 4}$$

it follows that

$$r(t_1, t_2, \dots, t_n) \geqq \sup_{\lambda_1 x_{11} + \lambda_2 x_{21} + \cdots + \lambda_k x_{k1} = t_1} f_1(x_{11}, x_{12}, \dots, x_{1n}) \cdots f_k(x_{k1}, x_{k2}, \dots, x_{kn}).$$

By the application of the inequality (2. 2) it follows from here that

$$\int_{-\infty}^{\infty} r(t_1, t_2, \dots, t_n)\, dt_1 \geqq$$

$$\geqq \Big(\int_{-\infty}^{\infty} f_1^{\frac{1}{\lambda_1}}(x_{11}, x_{12}, \dots, x_{1n})\, dx_{11} \Big)^{\lambda_1} \cdots \Big(\int_{-\infty}^{\infty} f_k^{\frac{1}{\lambda_k}}(x_{k1}, x_{k2}, \dots, x_{kn})\, dx_{k1} \Big)^{\lambda_k}.$$

Taking into account (2. 4) we can write further that

$$\begin{aligned} \int_{-\infty}^{\infty} r(t_1, t_2, \dots, t_n)\, dt_1 \geqq \sup_{\substack{\lambda_1 x_{12} + \dots + \lambda_k x_{k2} = t_2 \\ \vdots \\ \lambda_1 x_{1n} + \dots + \lambda_k x_{kn} = t_n}} & \Big(\int_{-\infty}^{\infty} f_1^{\frac{1}{\lambda_1}}(x_{11}, x_{12}, \dots, x_{1n})\, dx_{11} \Big)^{\lambda_1} \cdots \\ & \cdots \Big(\int_{-\infty}^{\infty} f_k^{\frac{1}{\lambda_k}}(x_{k1}, x_{k2}, \dots, x_{kn})\, dx_{k1} \Big)^{\lambda_k}. \end{aligned} \tag{2. 5}$$

Now we apply the inductive assumption that the inequality (2. 3) holds for functions of $n-1$ variables. This implies that the integral on the right hand side of (2. 5) with respect to $t_2, \dots, t_n$ is greater than or equal to the following product

$$\begin{aligned} & \Big(\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_1^{\frac{1}{\lambda_1}}(x_{11}, \dots, x_{1n})\, dx_{11} \dots dx_{1n} \Big)^{\lambda_1} \cdots \\ & \qquad \cdots \Big(\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_k^{\frac{1}{\lambda_k}}(x_{k1}, \dots, x_{kn})\, dx_{k1} \dots dx_{kn} \Big)^{\lambda_k}. \end{aligned} \tag{2. 6}$$

Looking at (2. 5) we immediately see that the integral of $r(\mathbf{t})$ over the space R_n is greater than or equal to the product standing in (2. 6). Thus the theorem is proved.

Remark. In what follows we need only that special case of the integral inequality (2. 3) where $k=2$ and the functions f_1, f_2 are logarithmic concave. The proof of this special case is very easy on the basis of the integral inequality (2. 1).

Below we give a sketch of this proof. We may restrict ourselves to the case of $n=1$ since Theorem 3 shows that the generalization for the case of $n>1$ is simple. Let $N=2^m$ where m is a positive integer and let $i+j=N$. By a subsequent application of (2. 1) we get

$$\int_{-\infty}^{\infty} \sup_{\frac{1}{N}(x_1+\dots+x_N)=t} f_1^{\frac{N}{i}}(x_1)\dots f_1^{\frac{N}{i}}(x_i) f_2^{\frac{N}{j}}(x_{i+1})\dots f_2^{\frac{N}{j}}(x_N)\,dt \geqq$$

$$\geqq \Big(\int_{-\infty}^{\infty} f_1^{\frac{N}{i}}(x)\,dx\Big)^{\frac{i}{N}} \Big(\int_{-\infty}^{\infty} f_2^{\frac{N}{j}}(x)\,dx\Big)^{\frac{j}{N}}.$$

The logarithmic concavity of f_1 and f_2 implies that the integrand on the left hand side is smaller than or equal to the following function

$$\sup_{\frac{i}{N}u+\frac{j}{N}v=t} f_1(u) f_2(v) \qquad (-\infty<t<\infty)$$

thus we have (2. 2) for $k=2$ and $\lambda=\frac{i}{N}$, $1-\lambda=\frac{j}{N}$. The assertion for arbitrary $\lambda(0<\lambda<1)$ follows from here by a continuity argument.

3. New proof and sharpening of Theorem 2. On the basis of Theorem 4 the proof of Theorem 2 is very simple. To do this let us define the functions f_1, f_2, f_3 as follows:

$$f_1(\mathbf{x}) = f(\mathbf{x}) \text{ if } \mathbf{x}\in A \text{ and } f_1(\mathbf{x}) = 0 \text{ otherwise,}$$

$$f_2(\mathbf{x}) = f(\mathbf{x}) \text{ if } \mathbf{x}\in B \text{ and } f_2(\mathbf{x}) = 0 \text{ otherwise,}$$

$$f_3(\mathbf{x}) = f(\mathbf{x}) \text{ if } \mathbf{x}\in \lambda A+(1-\lambda)B \text{ and } f_3(\mathbf{x}) = 0 \text{ otherwise.}$$

The logarithmic concavity of the function f implies that for every $\mathbf{t}$ we have

$$f_3(\mathbf{t}) \geqq \sup_{\lambda\mathbf{x}+(1-\lambda)\mathbf{y}=\mathbf{t}} (f_1(\mathbf{x}))^{\lambda}(f_2(\mathbf{y}))^{1-\lambda}.$$

Hence, applying Theorem 4, we obtain

$$\int_{\lambda A+(1-\lambda)B} f(\mathbf{t})\,d\mathbf{t} = \int_{R^n} f_3(\mathbf{t})\,d\mathbf{t} \geqq \Big(\int_{R^n} f_1(\mathbf{x})\,d\mathbf{x}\Big)^{\lambda}\Big(\int_{R^n} f_2(\mathbf{y})\,d\mathbf{y}\Big)^{1-\lambda} =$$

$$= \Big(\int_A f(\mathbf{x})\,d\mathbf{x}\Big)^{\lambda}\Big(\int_B f(\mathbf{y})\,d\mathbf{y}\Big)^{1-\lambda},$$

which is the required inequality.

Theorem 4. *Let P be a measure defined on the measurable subsets of R^n and generated by the logarithmic concave function f. Let A, B be two convex subsets of R^n with the property that $0<P(A)<\infty$, $0<P(B)<\infty$. Suppose that for every $\lambda(0<\lambda<1)$ the sets A and B can be decomposed as $A=A_1\cup A_2$, $B=B_1\cup B_2$, where $A_1\cap A_2=\emptyset$, $B_1\cap B_2=\emptyset$ so that the following conditions are satisfied.*

a) A_1, B_1 *are bounded closed convex sets,* A_2, B_2 *are convex sets.*
b) *The following relations hold*

$$[\lambda A_1+(1-\lambda)B_1]\cup[\lambda A_2+(1-\lambda)B_2]=\lambda A+(1-\lambda)B, \tag{3.1}$$

$$[\lambda A_1+(1-\lambda)B_1]\cap[\lambda A_2+(1-\lambda)B_2]=\emptyset, \tag{3.2}$$

$$[\lambda A_1+(1-\lambda)B_1]\cap A_1=\emptyset, \tag{3.3}$$

$$[\lambda A_1+(1-\lambda)B_1]\cap B_1=\emptyset. \tag{3.4}$$

c) *For the measures of the decomposing sets the following relations hold*

$$P(A_1)>0, \quad P(B_1)>0, \tag{3.5}$$

$$\frac{P(A_2)}{P(A_1)}=\frac{P(B_2)}{P(B_1)}. \tag{3.6}$$

d) f *is strictly logarithmic concave in the convex hull of* $A_1\cup B_1$ *i.e. the strict inequality holds in* (1. 1) *whenever* $\mathbf{x}_1, \mathbf{x}_2$ *are elements of the convex hull of* $A_1\cup B_1$ *and* $\mathbf{x}_1\neq\mathbf{x}_2$.

Under these conditions for every $\lambda\,(0<\lambda<1)$ *we have*

$$P(\lambda A+(1-\lambda)B)>(P(A))^\lambda(P(B))^{1-\lambda}.$$

Proof. Let λ be a number satisfying $0<\lambda<1$ and consider the subdivisions of the sets A, B belonging to this λ. Since A_1, B_1 are disjoint closed convex sets and f is strictly logarithmic concave in the convex hull of $A_1\cup B_1$, (3. 3) and (3. 4) imply that

$$f(\mathbf{t})>\sup_{\substack{\lambda\mathbf{x}+(1-\lambda)\mathbf{y}=\mathbf{t}\\ \mathbf{x}\in A_1,\,\mathbf{y}\in B_1}}(f(\mathbf{x}))^\lambda(f(\mathbf{y}))^{1-\lambda}. \tag{3.7}$$

Let $f_1(\mathbf{x})=f(x)$ if $x\in A_1$ and $f_1(\mathbf{x})=0$ otherwise, $f_2(\mathbf{y})=f(\mathbf{y})$ if $\mathbf{y}\in B_1$ and $f_2(\mathbf{y})=0$ otherwise. Then for every $\mathbf{t}\in\lambda A_1+(1-\lambda)B_1$ we have by (3. 7):

$$f(\mathbf{t})>\sup_{\lambda\mathbf{x}+(1-\lambda)\mathbf{y}=\mathbf{t}}(f_1(\mathbf{x}))^\lambda(f_2(\mathbf{y}))^{1-\lambda}. \tag{3.8}$$

If $\mathbf{t}\notin\lambda A_1+(1-\lambda)B_1$ then the right hand side in (3. 8) equals 0. Hence it follows that

$$\begin{aligned} P(\lambda A_1+(1-\lambda)B_1) &= \int_{\lambda A_1+(1-\lambda)B_1} f(\mathbf{t})\,d\mathbf{t} > \\ &> \int_{\lambda A_1+(1-\lambda)B_1}\sup_{\lambda\mathbf{x}+(1-\lambda)\mathbf{y}=\mathbf{t}}(f_1(\mathbf{x}))^\lambda(f_2(\mathbf{y}))^{1-\lambda}\,d\mathbf{t} = \\ &= \int_{R^n}\sup_{\lambda\mathbf{x}+(1-\lambda)\mathbf{y}=\mathbf{t}}(f_1(\mathbf{x}))^\lambda(f_2(\mathbf{y}))^{1-\lambda}\,d\mathbf{t} \geqq \\ &\geqq \Big(\int_{R^n} f_1(\mathbf{x})\,d\mathbf{x}\Big)^\lambda\Big(\int_{R^n} f_2(\mathbf{y})\,d\mathbf{y}\Big)^{1-\lambda} = (P(A_1))^\lambda(P(B_1))^{1-\lambda}. \end{aligned} \tag{3.9}$$

Continuing the reasoning it follows from (3. 1) and (3. 2), (3. 9) and Theorem 2 that for every $\lambda (0<\lambda<1)$

$$P(\lambda A+(1-\lambda)B) = P(\lambda A_1+(1-\lambda)B_1)+P(\lambda A_2+(1-\lambda)B_2) >$$
$$> (P(A_1))^\lambda (P(B_1))^{1-\lambda}+(P(A_2))^\lambda (P(B_2))^{1-\lambda}.$$

Taking into account (3. 5) we can write

$$(P(A))^\lambda (P(B))^{1-\lambda} = (P(A_1))^\lambda (P(B_1))^{1-\lambda}+(P(A_2))^\lambda (P(B_2))^{1-\lambda}.$$

Thus the theorem is proved.

One of the most important application of Theorem 5 is expressed by the following

Theorem 5. *Let f be a logarithmic concave probability density in R^n. Denote by F the probability distribution function belonging to the density f. If f is positive and strictly logarithmic concave in an open convex set D then F is also strictly logarithmic concave in the set D.*

Proof. Let $\mathbf{u}$, $\mathbf{v}$ be elements of the interior of the set D and suppose that $\mathbf{u}\neq\mathbf{v}$. By the definition of the function F we can write that

(3. 10)
$$F(\mathbf{u}) = P(A), \quad \text{where} \quad A = \{\mathbf{x}|\mathbf{x}\leqq \mathbf{u}\},$$
$$F(\mathbf{v}) = P(B), \quad \text{where} \quad B = \{\mathbf{x}|\mathbf{x}\leqq \mathbf{v}\}.$$

Given a $\lambda (0<\lambda<1)$, the following equality is obviously true

(3. 11)
$$F(\lambda\mathbf{u}+(1-\lambda)\mathbf{v}) = P(\lambda A+(1-\lambda)B).$$

Let us define the sets A_1, B_1, A_2, B_2 in the following way

$$A_1 = \left\{\mathbf{x}|\mathbf{x}\leqq \mathbf{u},\ \sum_{i=1}^n x_i \geqq \sum_{i=1}^n u_i-\varepsilon\right\}, \qquad B_1 = \left\{\mathbf{x}|\mathbf{x}\leqq \mathbf{v},\ \sum_{i=1}^n x_i \geqq \sum_{i=1}^n v_i-\delta\right\},$$
$$A_2 = A-A_1, \quad B_2 = B-B_1,$$

where ε and δ are fixed positive numbers. Obviously $P(A_1)>0$, $P(B_1)>0$.

Conditions (3. 1) and (3. 2) are satisfied for every positive ε, δ while conditions (3. 3), (3. 4) are satisfied for sufficiently small positive numbers ε, δ. This statement follows from the following equalities:

$$\lambda A_1+(1-\lambda)B_1 = \left\{\mathbf{x}|\mathbf{x}\leqq \lambda\mathbf{u}+(1-\lambda)\mathbf{v},\ \sum_{i=1}^n x_i \geqq \sum_{i=1}^n (\lambda u_i+(1-\lambda)v_i)-\lambda\varepsilon-(1-\lambda)\delta\right\},$$

$$\lambda A_2+(1-\lambda)B_2 = \left\{\mathbf{x}|\mathbf{x}\leqq \lambda\mathbf{u}+(1-\lambda)\mathbf{v},\ \sum_{i=1}^n x_i < \sum_{i=1}^n (\lambda u_i+(1-\lambda)v_i)-\lambda\varepsilon-(1-\lambda)\delta\right\}.$$

Let us fix an ε_0 and a δ_0 having this property. Then if (3. 6) holds true, we are ready.

If on the other hand (3. 6) is not satisfied, then since $P(A_1)$ is continuous in ε, $P(B_1)$ is continuous in δ and

$$\lim_{\varepsilon\to 0} P(A_1) = \lim_{\delta\to 0} P(B_1) = 0,$$

we can find positive numbers ε_1, δ_1 such that $\varepsilon_1 \leqq \varepsilon_0$, $\delta_1 \leqq \delta_0$ and (3. 6) is satisfied with these. Thus in view of (3. 10) and (3. 11) our theorem follows from Theorem 4.

4. Further properties of logarithmic concave functions. In this section we mention three theorems concerning logarithmic concave functions. The proofs are based on the integral inequality (2. 3).

Theorem 6. *Let $f(\mathbf{x}, \mathbf{y})$ be a function of $n+m$ variables where $\mathbf{x}$ is an n-component and $\mathbf{y}$ is an m-component vector. Suppose that f is logarithmic concave in R^{n+m} and et A be a convex subset of R^m. Then the function of the variable $\mathbf{x}$:*

$$\int_A f(\mathbf{x}, \mathbf{y})\, d\mathbf{y}$$

lis logarithmic in the entire space R^n.

Proof. Let $\mathbf{x}_1, \mathbf{x}_2 \in R^n$ and $0<\lambda<1$. Define the functions $f_1(\mathbf{y})$, $f_2(\mathbf{y})$, $f_3(\mathbf{y})$ as follows:

$$f_1(\mathbf{y}) = f(\mathbf{x}_1, \mathbf{y}) \text{ if } \mathbf{y}\in A, \text{ and } f_1(\mathbf{y}) = 0 \text{ otherwise,}$$

$$f_2(\mathbf{y}) = f(\mathbf{x}_2, \mathbf{y}) \text{ if } \mathbf{y}\in A, \text{ and } f_2(\mathbf{y}) = 0 \text{ otherwise,}$$

$$f_3(\mathbf{y}) = f(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2, \mathbf{y}) \text{ if } \mathbf{y}\in A, \text{ and } f_3(\mathbf{y}) = 0 \text{ otherwise.}$$

Since f is logarithmic concave in R^{n+m} and A is a convex set in R^m, we have

$$f_3(\mathbf{y}) \geqq \sup_{\lambda\mathbf{u}+(1-\lambda)\mathbf{v}=\mathbf{y}} (f_1(\mathbf{u}))^\lambda (f_2(\mathbf{v}))^{1-\lambda}.$$

Hence by Theorem 3 it follows that

$$\int_A f(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2, \mathbf{y})\, d\mathbf{y} = \int_{R^m} f_3(\mathbf{y})\, d\mathbf{y} \geqq$$

$$\geqq \Big(\int_{R^m} f_1(\mathbf{u})\, d\mathbf{u}\Big)^\lambda \Big(\int_{R^m} f_2(\mathbf{v})\, d\mathbf{v}\Big)^{1-\lambda} = \Big(\int_A f(\mathbf{x}_1, \mathbf{y})\, d\mathbf{y}\Big)^\lambda \Big(\int_A f(\mathbf{x}_2, \mathbf{y})\, d\mathbf{y}\Big)^{1-\lambda}.$$

Thus the theorem is proved.

Theorem 7. *Let f, g be logarithmic concave functions defined in the space R^n. Then the convolution of these functions i.e.*

$$\int_{R^n} f(\mathbf{x}-\mathbf{y})\, g(\mathbf{y})\, d\mathbf{y}$$

is also logarithmic concave in the entire space R^n.

Proof. The theorem is a consequence of Theorem 7. In fact the function $f(\mathbf{x}-\mathbf{y})g(\mathbf{y})$ is a logarithmic concave function of the $2n$ variables contained in the vectors $\mathbf{x}$ and $\mathbf{y}$. Applying Theorem 6 for this function and for the convex set $A=R^n$, we obtain the assertion of Theorem 7.

For the case $n=1$ the assertion was proved by IBRAGIMOV [1] in 1956. The following theorem is an immediate consequence of Theorem 6. It is mentioned separately for completeness.

Theorem 8. *If f is a logarithmic concave multivariate probability density, then all marginal densities are also logarithmic concave.*

Bibliography

[1] И. А. Ибрагимов, О композиции одновершиных распределений, *Теория вероятностей и ее применения*, **1** (1956), 283—288.
[2] L. LEINDLER, On a certain converse of Hölder's inequality. II, *Acta Sci. Math.*, **33** (1972), 217—223.
[3] A. PRÉKOPA, Logarithmic concave measures with application to stochastic programming, *Acta Sci. Math.*, **32** (1971), 301—316.
[4] B. SZ.-NAGY, *Introduction to Real Functions and Orthogonal Expansions* (Budapest, 1964).

(Received October 25, 1972)

Einige Teilbarkeitskriterien

Von LÁSZLÓ RÉDEI in Budapest

Prof. Béla Szőkefalvi-Nagy zum 60. Geburtstag gewidmet

§ 1

Für natürliche Zahlen $k(>1)$, a, n gilt das oft verwendete Teilbarkeitskriterium

$$k^a-1|k^n-1 \Leftrightarrow a|n. \tag{1}$$

Da n beiderseits offenbar nur mod a in Betracht kommt, genügt es (1) für den Fall $1 \leqq n \leqq a$ zu beweisen. In diesem Fall sind aber beide Seiten von (1) gleichbedeutend mit $a=n$. Folglich ist (1) allgemein richtig. Man sieht, daß in diesem Beweis auch von der Ordnungsrelation $<$ (oder, was auf dasselbe hinausläuft, von der Betragsbewertung des rationalen Zahlkörpers) Gebrauch gemacht wurde.

Für nichtkonstante Polynome $f(x)$ über einem Körper gilt ähnlich

$$f(x)^a-1|f(x)^n-1 \Leftrightarrow a|n. \tag{2}$$

Der Beweis ist dem vorigen ähnlich, mit dem geringen Unterschied, daß nach Reduktion auf den Fall $1 \leqq n \leqq a$ Gradvergleich (d.h. im wesentlichen Gradbewertung) hilft.

(1) und (2) lassen für beliebige Integritätsbereiche eine gemeinsame Verallgemeinerung zu (auch wenn keine Bewertung zur Verfügung steht).

Satz 1. *Für ein Element $\varkappa$ eines kommutativen Integritätsbereiches R mit Einselement ε und der Einheitengruppe E und für natürliche Zahlen a, n gilt das Teilbarkeitskriterium*

$$\varkappa^a-\varepsilon|\varkappa^n-\varepsilon \Leftrightarrow \varkappa^{(a,n)}=\varepsilon \wedge \left(\varkappa^{(a,n)} \neq \varepsilon \,\&\, \frac{\varkappa^a-\varepsilon}{\varkappa^{(a,n)}-\varepsilon} \in E\right), \tag{3}$$

wobei $\wedge$ und & als „oder" bzw. „und" zu lesen sind und (a, n) den größten gemeinsamen Teiler von a und n bezeichnet. (Der auftretende Quotient liegt in R, da sein Nenner ein Teiler des Zählers ist.)

(Der Leser sieht, daß (1) und (2) Spezialfälle von Satz 1 sind.)

Den Beweis von Satz 1 beginnen wir mit der Bemerkung, daß für jedes $\xi \in R$

$$\varkappa^a - \varepsilon | \xi \Leftrightarrow \varkappa^a - \varepsilon | \varkappa \xi \tag{4}$$

besteht. Hiervon ist der Teil $\Rightarrow$ trivialerweise richtig. Um den Teil $\Leftarrow$ zu beweisen, setzen wir die rechte Seite von (4) voraus. Dann gilt

$$\varkappa^a - \varepsilon | \varkappa^a \xi,$$

also gilt wegen $\varkappa^a \xi = (\varkappa^a - \varepsilon)\xi + \xi$ auch die linke Seite von (4). Somit ist (4) richtig. (Man sieht, daß im Beweis die Nullteilerfreiheit von R nicht ausgenutzt wurde.)

Nun ist der Teil $\Leftarrow$ von (3) trivialerweise richtig. Um den Teil $\Rightarrow$ zu beweisen, setzen wir die linke Seite von (3) voraus. Dann gelten

$$\varkappa^a - \varepsilon | \varkappa^{au} - \varepsilon, \quad \varkappa^a - \varepsilon | \varkappa^{nv} - \varepsilon$$

für alle natürlichen Zahlen u, v. Folglich gilt

$$\varkappa^a - \varepsilon | \varkappa^{au} - \varkappa^{nv}.$$

Für $au > nv$ entsteht hieraus nach (4)

$$\varkappa^a - \varepsilon | \varkappa^{au-nv} - \varepsilon.$$

Wählt man u und v so, daß

$$au - nv = (a, n)$$

gilt, so gewinnt man

$$\varkappa^a - \varepsilon | \varkappa^{(a,n)} - \varepsilon.$$

Da hieraus trivialerweise die rechte Seite von (3) folgt, ist Satz 1 bewiesen.

§ 2

Als ein „nächster" Schritt nach (1) *gilt für natürliche Zahlen* $k(>1)$, a, b, n *das Teilbarkeitskriterium*

$$(k^a - 1)(k^b - 1) | k^n - 1 \Leftrightarrow [a, b](k^{(a,b)} - 1) | n, \tag{5}$$

wobei $[a, b]$ *das kleinste gemeinsame Vielfache von* a *und* b *ist.*

Wir bemerken, daß (1) zum Beweis von (5) natürlich nicht ausreicht.

Ohne wesentlich größere Mühe beweisen wir den allgemeineren:

Satz 2. *Für ein Element* $\varkappa$ *eines Integritätsbereiches* R *mit dem Einselement* ε *und der Einheitengruppe* E *und für natürliche Zahlen* a, b, n *gilt das Teilbarkeitskriterium*

$$(\varkappa^a - \varepsilon)(\varkappa^b - \varepsilon) | \varkappa^n - \varepsilon \Leftrightarrow \varkappa^n = \varepsilon \wedge \left(\varkappa^n \neq \varepsilon \,\&\, \frac{\varkappa^a - \varepsilon}{\varkappa^{(a,n)} - \varepsilon}, \frac{\varkappa^b - \varepsilon}{\varkappa^{(b,n)} - \varepsilon} \in E \,\& \right.$$
$$\left. \&\, \varkappa^{(a,b,n)} - \varepsilon \left| \frac{(a, b, n) n}{(a, n)(b, n)} \varepsilon \right. \right). \tag{6}$$

(Der Leser sieht leicht, daß (5) in der Tat ein Spezialfall von Satz 2 ist.)

Im Fall $\varkappa^n = \varepsilon$ sind beide Seiten von (6) trivialerweise richtig. Deshalb setzen wir fortan

$$\varkappa^n \neq \varepsilon \tag{7}$$

voraus. Dann reduziert sich die Behauptung (6) auf

$$(\varkappa^a - \varepsilon)(\varkappa^b - \varepsilon) | \varkappa^n - \varepsilon \Leftrightarrow \frac{\varkappa^a - \varepsilon}{\varkappa^{(a,n)} - \varepsilon}, \frac{\varkappa^b - \varepsilon}{\varkappa^{(b,n)} - \varepsilon} \in E \,\&\, \varkappa^{(a,b,n)} - \varepsilon | \frac{(a,b,n)n}{(a,n)(b,n)} \varepsilon. \tag{8}$$

Wenn die linke Seite von (8) erfüllt ist, so sind $\varkappa^a - \varepsilon$ und $\varkappa^b - \varepsilon$ Teiler von $\varkappa^n - \varepsilon$, ferner folgt aus (7) offenbar

$$\varkappa^{(a,n)}, \varkappa^{(b,n)} \neq \varepsilon,$$

also muß nach Satz 1

$$\frac{\varkappa^a - \varepsilon}{\varkappa^{(a,n)} - \varepsilon}, \frac{\varkappa^b - \varepsilon}{\varkappa^{(b,n)} - \varepsilon} \in E \tag{9}$$

gelten. Deshalb setzen wir fortan auch (9) voraus. Dann reduziert sich die Behauptung (8) weiter auf

$$(\varkappa^a - \varepsilon)(\varkappa^b - \varepsilon) | \varkappa^n - \varepsilon \Leftrightarrow \varkappa^{(a,b,n)} - \varepsilon | \frac{(a,b,n)n}{(a,n)(b,n)} \varepsilon. \tag{10}$$

(Im folgenden Beweis von (10) wird von Satz 1 kein Gebrauch mehr gemacht.)

Wir schicken den Beweis für den Fall

$$(a, b, n) = 1 \tag{11}$$

voraus. Gleich bemerken wir, daß aus (11)

$$(a,n)(b,n) | n, \quad ((a,n),(b,n)) = 1 \tag{12}$$

folgen.

Wir benötigen eine kleine Vorbereitung. Hierzu nehmen wir zwei teilerfremde natürliche Zahlen r, s und setzen

$$f(x) = \frac{(x^{rs} - \varepsilon)(x - \varepsilon)}{(x^r - \varepsilon)(x^s - \varepsilon)}, \tag{13}$$

wobei Zähler und Nenner als Elemente des Polynomringes $R[x]$ zu deuten sind. Da aber $x - \varepsilon$ im Ideal $(x^r - \varepsilon, x^s - \varepsilon)$ liegt, gilt wegen (13)

$$f(x) \in R[x]$$

offenbar. Da ferner der Ersetzungswert eines Polynoms von der Form

$$\frac{x^i - \varepsilon}{x - \varepsilon} \in R[x] \qquad (i = 1, 2, \ldots)$$

für $x=\varepsilon$ gleich $i\varepsilon$ ist, folgt aus (13) $f(\varepsilon)=\varepsilon$. Andererseits ist $f(\varkappa)\equiv f(\varepsilon)$ (mod $\varkappa-\varepsilon$), also gilt

$$f(\varkappa) \equiv \varepsilon \,(\text{mod}\, \varkappa - \varepsilon). \tag{14}$$

Nun ist die linke Seite von (10) wegen (9) gleichbedeutend mit

$$(\varkappa^{(a,n)}-\varepsilon)(\varkappa^{(b,n)}-\varepsilon)|\varkappa^n-\varepsilon.$$

Wegen (7) und (12) formt sich diese Teilbarkeit identisch in

$$(\varkappa^{(a,n)}-\varepsilon)(\varkappa^{(b,n)}-\varepsilon)(\varkappa-\varepsilon)|\frac{\varkappa^n-\varepsilon}{\varkappa^{(a,n)(b,n)}-\varepsilon}(\varkappa^{(a,n)(b,n)}-\varepsilon)(\varkappa-\varepsilon)$$

um. Verwendet man (13) mit $r=(a,n)$, $s=(b,n)$ (die hierzu nötige Bedingung $(r,s)=1$ ist nach (12_2) erfüllt), so entsteht (nach Kürzung) die weitere äquivalente Umformung

$$\varkappa-\varepsilon\,\Big|\,\frac{\varkappa^n-\varepsilon}{\varkappa^{(a,n)(b,n)}-\varepsilon}f(\varkappa).$$

Da sich der zweite Faktor der rechten Seite wegen (14) streichen läßt und der erste Faktor kongruent

$$\frac{n}{(a,n)(b,n)}\varepsilon$$

mod $\varkappa-\varepsilon$ ist, ist (10) für den Fall (11) bewiesen.

Der allgemeine Fall läßt sich leicht auf den vorigen zurückführen. Zu diesem Zweck setzen wir

$$d=(a,b,n), \quad \lambda=\varkappa^d.$$

Dann ist (10) gleichbedeutend mit

$$(\lambda^{\frac{a}{d}}-\varepsilon)(\lambda^{\frac{b}{d}}-\varepsilon)|\lambda^{\frac{n}{d}}-\varepsilon \Leftrightarrow \lambda-\varepsilon|\frac{\frac{n}{d}}{\left(\frac{a}{d},\frac{n}{d}\right)\left(\frac{b}{d},\frac{n}{d}\right)}\varepsilon,$$

ist also nach vorigem Resultat richtig. Somit ist Satz 2 bewiesen.

(Eingegangen am 11. März 1972)

A new law of the iterated logarithm for multiplicative systems

By P. RÉVÉSZ in Budapest

To Professor B. Sz.-Nagy on his 60th birthday

Introduction

The sequence $\varphi_1, \varphi_2, \dots$ of random variables on $(X, \mathcal{F}, \mathbf{P})$ is called a multiplicative system (MS) if

$$\int \varphi_{i_1}\varphi_{i_2}\dots\varphi_{i_k} = 0 \qquad (i_1 < i_2 < \dots < i_k;\ k = 1, 2, \dots),$$

it is called an equinormed strongly multiplicative system (ESMS) if

$$\int \varphi_i = 0,\ \int \varphi_i^2 = 1 \qquad (i = 1, 2, \dots),$$

$$\int \varphi_{i_1}^{r_1}\varphi_{i_2}^{r_2}\dots\varphi_{i_k}^{r_k} = \int \varphi_{i_1}^{r_1}\int \varphi_{i_2}^{r_2}\dots\int \varphi_{i_k}^{r_k} \qquad (i_1 < i_2 < \dots < i_k;\ k = 1, 2, \dots),$$

where r_i $(i = 1, 2, \dots, k)$ can be equal to 1 or 2.

Several theorems state that the properties of a MS resp. ESMS are very similar to those of independent systems.

The best known laws of the iterated logarithm for a MS are the following:

Theorem A. (S. TAKAHASHI[1]) [1].) *Let $\varphi_1, \varphi_2, \dots$ be a uniformly bounded* MS *for which*

(1) $$\int \varphi_i^2 = \int \varphi_i^2\varphi_j^2 = 1 \qquad (i = 1, 2, \dots;\ j = 1, 2, \dots,\ i \neq j).$$

Then

$$\varlimsup_{n\to\infty} \frac{\varphi_1 + \varphi_2 + \dots + \varphi_n}{\sqrt{2n \log\log n}} \leqq 1 \qquad a.e.$$

In fact Takahashi assumed (instead of (1)) that

(2) $$\frac{\varphi_1^2 + \varphi_2^2 + \dots + \varphi_{[\theta^n]}^2}{[\theta^n]} \to 1 \qquad \text{a.e.}$$

for any $\theta > 1$. (Clearly (1) implies (2).)

[1]) See also [2].

Theorem B. (GAPOSHKIN [3].) *Let* $\varphi_1, \varphi_2, \ldots$ *be a uniformly bounded* ESMS *and let* $a_1, a_2, \ldots$ *be a sequence of real numbers for which*

$$a_n = o\left(\frac{A_n}{\sqrt{\log\log n}}\right) \quad \text{and} \quad A_n \to \infty$$

where $A_n^2 = \sum_{k=1}^{n} a_k^2$. *Then*

$$\varlimsup_{n\to\infty} \frac{\sum_{k=1}^{n} a_k \varphi_k}{\sqrt{2A_n^2 \log\log A_n}} \leqq 1 \qquad \text{a.e.}$$

Theorem C. (RÉVÉSZ [4].) *Let* $\varphi_1, \varphi_2, \ldots$ *be a uniformly bounded* MS *for which*

$$\int \varphi_{i_1}^2 \varphi_{i_2}^2 \ldots \varphi_{i_k}^2 = 1 \qquad (i_1 < i_2 < \cdots < i_k;\ k = 1, 2, \ldots).$$

Then

$$\varlimsup_{n\to\infty} \frac{\varphi_1 + \varphi_2 + \cdots + \varphi_n}{\sqrt{2n \log\log n}} \geqq 1 \qquad \text{a.e.}$$

In this paper we intend to find a common generalization of Theorems A and B. Our theorem can be formulated as follows:

Theorem 1. *Let* $\varphi_1, \varphi_2, \ldots$ *be a uniformly bounded* MS *and let* $a_1, a_2, \ldots$ *be a sequence of real numbers for which*

(3)
$$a_n = o\left(\frac{A_n}{\sqrt{\log\log n}}\right) \quad \text{and} \quad A_n \to \infty,$$

where $A^2(n) = A_n^2 = \sum_{k=1}^{n} a_k^2$. *Further, let* $M_k = M_k(\theta)$, *defined by*

(4)
$$A_{M_k - 1}^2 < \theta^k \leqq A_{M_k}^2$$

and suppose

(5)
$$\varliminf_{k\to\infty} \frac{T_{M_k}^2}{A_{M_k}^2} > 0 \qquad \text{a.e.}$$

for any $\theta > 1$, *where* $T^2(n) = T_n^2 = \sum_{k=1}^{n} a_k^2 \varphi_k^2$. *Then*

$$\varlimsup_{n\to\infty} \frac{S_n}{\sqrt{2T_n^2 \log\log A_n}} \leqq 1 \qquad \text{a.e.,}$$

where $S(n) = S_n = \sum_{k=1}^{n} a_k \varphi_k$.

Remark. This theorem is clearly a generalization of Theorem A. The fact that it is also a generalization of Theorem B is shown in the consequence of Lemma 4.

The proof of this Theorem is essentially based on that of TAKAHASHI [1].

§ 1 contains some inequalities. The proof of Theorem 1 is prepared in § 2.

§ 1. Inequalities

Theorem D. (Azuma [5].) *Let $\varphi_1, \varphi_2, \dots, \varphi_n$ be a uniformly bounded* MS *($|\varphi_i| \leqq K_1$, $i = 1, 2, \dots, n$) and let $a_1, a_2, \dots, a_n$ be a sequence of real numbers, further let λ be a positive number. Then*

$$\int e^{\lambda S} \leqq \exp\left(\frac{\lambda^2 A^2 K_1^2}{2}\right),$$

where $S = \sum_{k=1}^{n} a_k \varphi_k$ and $A^2 = \sum_{k=1}^{n} a_k^2$.

We reproduce the proof here because the original one contains a minor missprint.

Proof. Since e^x is a convex function, for $|x| \leqq 1$ and $a \neq 0$ we have

$$e^{ax} \leqq e^{|a|} \frac{|a| + ax}{2|a|} + e^{-|a|} \frac{|a| - ax}{2|a|} = \operatorname{ch}(|a|) + \frac{ax}{|a|} \operatorname{sh}(|a|).$$

Hence

$$\int e^{\lambda S} = \int \prod_{k=1}^{n} \exp\left(a_k \lambda K_1 \frac{\varphi_k}{K_1}\right) \leqq$$

$$\leqq \int \prod_{k=1}^{n} \left[\operatorname{ch}(|a_k| \lambda K_1) + \frac{a_k}{|a_k|} \frac{\varphi_k}{K_1} \operatorname{sh}(|a_k| \lambda K_1)\right] =$$

$$= \prod_{k=1}^{n} \operatorname{ch}(|a_k| \lambda K_1) = \prod_{k=1}^{n} \sum_{m=0}^{\infty} \frac{(\lambda K_1 |a_k|)^{2m}}{(2m)!} \leqq \prod_{k=1}^{n} \sum_{m=0}^{\infty} \frac{(\lambda K_1 |a_k|)^{2m}}{2^m m!} =$$

$$= \prod_{k=1}^{n} \exp\left(\frac{\lambda^2 K_1^2 a_k^2}{2}\right) = \exp\left(\frac{\lambda^2 K_1^2 A^2}{2}\right),$$

i.e., Theorem D is proved.

Theorem 2. *Let $\varphi_1, \varphi_2, \dots, \varphi_n$ ($|\varphi_i| \leqq K_1$; $i = 1, 2, \dots, n$) be a sequence of uniformly bounded random variables and let $a_1, a_2, \dots, a_n$ be a sequence of real numbers. Then*

$$\exp\left(\lambda S - \frac{\lambda^2 T^2}{2}\left(1 + 2\lambda K_1 \max_{1 \leqq k \leqq n} |a_k|\right)\right) \leqq \prod_{k=1}^{n} (1 + \lambda a_k \varphi_k)$$

where

$$S = \sum_{k=1}^{n} a_k \varphi_k, \quad T^2 = \sum_{k=1}^{n} a_k^2 \varphi_k^2, \quad A^2 = \sum_{k=1}^{n} a_k^2,$$

and λ is a positive number for which

$$\lambda K_1 \max_{1 \leqq k \leqq n} |a_k| \leqq \tfrac{1}{2}.$$

Proof. Since

$$e^x \leqq (1+x)\exp\left(\frac{x^2}{2}+|x^3|\right) \quad \text{if} \quad |x| \leqq \tfrac{1}{2}$$

we have

$$\exp(\lambda S) \leqq \prod_{k=1}^{n} (1+\lambda a_k \varphi_k)\exp\left(\frac{\lambda^2 a_k^2 \varphi_k^2}{2}+\lambda^3 |a_k \varphi_k|^3\right) \leqq$$

$$\leqq \exp\left(\frac{\lambda^2 T^2}{2}(1+2\lambda K_1 \max_{1\leqq k\leqq n} |a_k|)\right)\prod_{k=1}^{n}(1+\lambda a_k \varphi_k)$$

which implies our Theorem.

Theorem 3. *Let* $\varphi_1, \varphi_2, \ldots, \varphi_n$ *be a uniformly bounded* MS $(|\varphi_i|\leqq K_1;$ $i=1,2,\ldots,n)$ *and let* $a_1, a_2, \ldots, a_n$ *be a sequence of real numbers, further let* y *be a positive number. Then* $\mathbf{P}\{|S|\geqq yK_1 A\sqrt{2}\}\leqq 2e^{-y^2}$, *where* $S=\sum_{k=1}^{n} a_k\varphi_k$ *and* $A^2 = \sum_{k=1}^{n} a_k^2$.

Proof. Set $\lambda=(\sqrt{2}y)/(K_1 A)$. Then by Theorem D we have

$$\int e^{\lambda|S|} \leqq \int e^{\lambda S}+\int e^{-\lambda S} \leqq 2\exp\left(\frac{\lambda^2 A^2 K_1^2}{2}\right),$$

and the Markov inequality gives

$$\mathbf{P}(|S| \geqq yK_1 A\sqrt{2}) = \mathbf{P}(e^{\lambda|S|} \geqq \exp(\lambda y K_1 A\sqrt{2})) \leqq$$

$$\leqq 2\exp\left(\frac{\lambda^2 A^2 K_1^2}{2}-\lambda y K_1 A\sqrt{2}\right) = 2\exp(y^2-2y^2) = 2e^{-y^2},$$

which proves our Theorem 3.

Consequence of Theorem 3. *Let* $\varphi_1, \varphi_2, \ldots$ $(|\varphi_i|\leqq K_1;\ i=1,2,\ldots)$ *be a sequence of uniformly bounded random variables for which*

$$\int \varphi_{i_1}^2\varphi_{i_2}^2\ldots\varphi_{i_k}^2 = 1 \qquad (i_1<i_2<\cdots<i_k;\ k=1,2,\ldots)$$

and let $a_1, a_2, \ldots$ *be a sequence of real numbers satisfying condition* (3). *Then*

$$\mathbf{P}\left\{\left|\frac{a_1^2\varphi_1^2+\cdots+a_n^2\varphi_n^2}{a_1^2+\cdots+a_n^2}-1\right| \geqq \varepsilon\right\} \leqq 2\exp\left(-\frac{\varepsilon^2}{2(K_1^2+1)^2}\frac{\left(\sum_{k=1}^{n} a_k^2\right)^2}{\sum_{k=1}^{n} a_k^4}\right) \leqq$$

$$\leqq 2\exp\left(-\log\log\left(\sqrt{\sum_{k=1}^{n} a_k^2}\right)\right)$$

for any $\varepsilon>0$ *if* n *is large enough.*

Proof. Clearly $\{\varphi_k^2-1\}$ is a MS. Hence by Theorem 3,

$$\mathbf{P}\left\{\left|\frac{a_1^2\varphi_1^2+\cdots+a_n^2\varphi_n^2}{a_1^2+\cdots+a_n^2}-1\right|\geqq\varepsilon\right\}=\mathbf{P}\left\{\left|\sum_{k=1}^n a_k^2(\varphi_k^2-1)\right|\geqq\varepsilon\sum_{k=1}^n a_k^2\right\}=$$

$$=\mathbf{P}\left\{\left|\sum_{k=1}^n a_k^2(\varphi_k^2-1)\right|\geqq\sqrt{2}\,\frac{\varepsilon\sum_{k=1}^n a_k^2}{\sqrt{2}\,(K_1^2+1)\sqrt{\sum_{k=1}^n a_k^4}}(K_1^2+1)\sqrt{\sum_{k=1}^n a_k^4}\right\}\leqq$$

$$\leqq 2\exp\left(-\frac{\varepsilon^2}{2(K_1^2+1)^2}\frac{\left(\sum_{k=1}^n a_k^2\right)^2}{\sum_{k=1}^n a_k^4}\right).$$

Since (3) implies

$$\max_{1\leqq k\leqq n}|a_k|=o\left(\frac{A_n}{\sqrt{\log\log A_n}}\right)\qquad\left(A_n^2=\sum_{k=1}^n a_k^2\right),$$

we have

$$\frac{\left(\sum_{k=1}^n a_k^2\right)^2}{\sum_{k=1}^n a_k^4}\geqq\frac{\left(\sum_{k=1}^n a_k^2\right)^2}{\left(\sum_{k=1}^n a_k^2\right)\left(\max_{1\leqq k\leqq n}|a_k|\right)^2}\geqq\frac{4(K_1^2+1)^2}{\varepsilon^2}\log\log\left(\sqrt{\sum_{k=1}^n a_k^2}\right)$$

f n is large enough, and this proves the consequence.

Theorem 4. *Let $\varphi_1,\varphi_2,\ldots,\varphi_n$ be a uniformly bounded* MS $(|\varphi_i|\leqq K_1$; *$i=1,2,\ldots,n)$ and let $a_1,a_2,\ldots,a_n$ be a sequence of real numbers. Then*

(6) $$\mathbf{P}\left\{\max_{1\leqq m\leqq n}|S(m)|\geqq K_2\sqrt{A^2\log\log A}\right\}\leqq K_3\exp(-2\log\log A)$$

where

$$S(m)=\sum_{k=1}^m a_k\varphi_k,\quad A^2=\sum_{k=1}^n a_k^2.$$

K_2 and K_3 are suitable positive constants.

Before the proof of this theorem we introduce some notations: Let $a_1,a_2,\ldots$ be a sequence of real numbers and let

$$I=\{m,m+1,\ldots,n\}=[m,n]\qquad(m\leqq n)$$

be the interval of the integers between m and n. Let VI be a partition of I. In its definition we distinguish two cases:

Case 1. There exists ϱ such that $m\leqq\varrho\leqq n$ and $a_\varrho^2\geqq\frac{1}{2}\sum_{i=m}^n a_i^2$.

Case 2. Such an integer does not exist.

In Case 1,

$$\mathrm{V}I = \{\{m, m+1, \ldots, \varrho-1\}, \{\varrho\}, \{\varrho+1, \varrho+2, \ldots, n\}\}.$$

(Of course it can happen that one of these intervals is empty.)

In Case 2,

$$\{\mathrm{V}I = \{m, m+2, \ldots, \tau\}, \{\tau+1, \tau+2, \ldots, n\}\},$$

where τ is defined by

$$\left| \sum_{i=m}^{\tau} a_i^2 - \sum_{i=\tau+1}^{n} a_i^2 \right| = \min.$$

Now let P be a sequence of intervals:

$$P = \{[m_1, n_1], [m_2, n_2], \ldots, [m_s, n_s]\} = \{I_1, I_2, \ldots, I_s\}$$
$$(m_1 \leqq n_1 < m_2 \leqq n_2 < \cdots < m_s \leqq n_s).$$

Then we define UP as the subsequence of P containing those elements (of P) which have more than 1 element (integer).

Finally let

$$\mathrm{V}P = \{\mathrm{V}I_1, \mathrm{V}I_2, \ldots, \mathrm{V}I_s\}.$$

Now construct the sequence $P_0, P_1, \ldots$ as follows:

$$P_0 = \{[1, n]\} \quad \text{and} \quad P_{t+1} = \mathrm{V}UP_t \qquad (t = 0, 1, 2, \ldots).$$

We mention the following two simple properties of the sequence $P_0, P_1, \ldots$.

Property 1. If μ_t is the number of the elements of P_t then $\mu_t \leqq 3.2^{t-1}$ $(t = 1, 2, \ldots)$.

Property 2. If $I_{tj} \in P_t$ then

$$A^2(t, j) \leqq (\tfrac{3}{4})^{t-1} A^2(I) \quad \left(A^2(t, j) = \sum_{k \in I_{tj}} a_k^2;\ A^2(I) = \sum_{k=1}^{n} a_k^2 \right).$$

Now we can turn to the

Proof of Theorem 4. Clearly we have

$$\max_{1 \leqq m \leqq n} |S(m)| \leqq 2 \sum_{t=0}^{\infty} \max_{1 \leqq j \leqq \mu_t} |S(t, j)|$$

where

$$S(t, j) = \sum_{k \in I_{t,j}} a_k \varphi_k; \qquad \{I_{t1}, I_{t2}, \ldots, I_{t\mu_t}\} = P_t.$$

Set

$$y_t = \sqrt{2 \log \log A + 2t}, \qquad A^2 = \sum_{k=1}^{n} a_k^2,$$

$$F_t = \bigcup_{j=1}^{\mu_t} \{|S(t, j)| \geqq \sqrt{2}\, y_t K_1 A(t, j)\}, \qquad E = \bigcup_{t=0}^{\infty} F_t.$$

Then by Theorem 3 we have

$$\mathbf{P}\left\{|S(t,j)| \geqq \sqrt{2}\, y_t K_1 A(t,j)\right\} \leqq 2e^{-y_t^2} = 2\,\frac{e^{-2\log\log A}}{e^{2t}}$$

hence

$$\mathbf{P}(F_t) \leqq \sum_{j=1}^{\mu_t} 2\,\frac{e^{-2\log\log A}}{e^{2t}} \leqq \frac{3}{2^t}\, e^{-2\log\log A}$$

and

(7) $$\mathbf{P}(E) \leqq \sum_{t=0}^{\infty} \mathbf{P}(F_t) \leqq 6e^{-2\log\log A}.$$

Clearly if $x \notin F_t$ then

$$\max_{1\leqq j\leqq \mu_t} |S(t,j)| \leqq \sqrt{2}\, K_1 \sqrt{2\log\log A + 2t}\, \max_{1\leqq j\leqq \mu_t} A(t,j) \leqq$$

$$\leqq \sqrt{2}\, K_1 \sqrt{2t}\, \max_{1\leqq j\leqq \mu_t} A(t,j) + \sqrt{2}\, K_1 \sqrt{2\log\log A}\, \max_{1\leqq j\leqq \mu_t} A(t,j) \leqq$$

$$\leqq \sqrt{2}\, K_1 \sqrt{2t}\, \left(\sqrt{\tfrac{3}{4}}\right)^{t-1} A + \sqrt{2}\, K_1 \sqrt{2\log\log A}\, \left(\sqrt{\tfrac{3}{4}}\right)^{t-1} A$$

and if $x \notin E$ then

(8) $$\max_{1\leqq m\leqq n} |S(m)| \leqq \left[4K_1 \sum_{t=0}^{\infty} \sqrt{t}\, \left(\sqrt{\tfrac{3}{4}}\right)^{t-1} + K_1 \sqrt{\log\log A}\, \sum_{t=0}^{\infty} \left(\sqrt{\tfrac{3}{4}}\right)^{t-1}\right] A \leqq$$

$$\leqq K_2 \sqrt{A^2 \log\log A}.$$

(7) and (8) imply (6).

§ 2. The proof of Theorem 1

First we prove several lemmas.

Lemma 1. *Under the conditions and notations of Theorem* 1 *we have*

$$\mathbf{P}\left\{\frac{S(M_k)}{\sqrt{2A^2(M_k)\log\log A(M_k)}} \geqq \frac{T^2(M_k)}{2CA^2(M_k)}\left(1 + 2\lambda K_1 \max_{1\leqq j\leqq M_k} |a_j|\right) + (1+\varepsilon)\frac{C}{2}\right\} =$$

$$= O\left(\frac{1}{k^{1+\varepsilon}}\right)$$

for any $C > 0$ where

$$\lambda = \lambda(C) = \sqrt{\frac{2\log\log A(M_k)}{C^2 A^2(M_k)}}.$$

Proof. Set

$$y = (1+\varepsilon)\, C \sqrt{\frac{A^2(M_k)\log\log A(M_k)}{2}}.$$

Since condition (3) implies

$$\max_{1 \leqq j < N} |a_j| = o\left(\frac{A_N}{\sqrt{\log\log A_N}}\right),$$

we have

$$\lambda K_1 \max_{1 \leqq j < M_k} |a_j| = \sqrt{\frac{2\log\log A(M_k)}{C^2 A^2(M_k)}}\, K_1 \max_{1 \leqq j \leqq M_k} |a_j| \leqq \tfrac{1}{2}$$

(if k is large enough). Furthermore, Theorem 2 implies

$$\mathbf{P}\left\{\frac{S(M_k)}{\sqrt{2A^2(M_k)\log\log A(M_k)}} \geqq \frac{T^2(M_k)}{2A^2(M_k)C}(1+2\lambda K_1 \max_{1 \leqq j \leqq M_k} |a_j|) + (1+\varepsilon)\frac{C}{2}\right\} =$$

$$= \mathbf{P}\left\{S(M_k) \geqq \frac{\lambda}{2} T^2(M_k)(1+2\lambda K_1 \max_{1 \leqq j \leqq M_k} |a_j|) + y\right\} =$$

$$= \mathbf{P}\left\{\exp\left(\lambda S(M_k) - \frac{\lambda^2}{2} T^2(M_k)(1+2\lambda K_1 \max_{1 \leqq j \leqq M_k} |a_j|)\right) \geqq e^{\lambda y}\right\} \leqq$$

$$\leqq e^{-\lambda y} = O\left(\frac{1}{k^{1+\varepsilon}}\right),$$

i.e., Lemma 1 is proved.

Lemma 2. *Under the conditions of Theorem* 1 *for any* $\varrho > 0$ *one can find a set* $F(\in \mathscr{F})$, *a positive number* $\mathscr{K}$ *and an integer* n_0 *such that*

$$\mathbf{P}(F) \leqq \varrho$$

and $(T_n^2/A_n^2)^{\frac{1}{2}} \geqq \mathscr{K}$ *hold on* $\bar{F}$ *if* $n \geqq n_0$.

Proof. This lemma is a trivial consequence of (5).

Lemma 3. *Define the event* $\mathfrak{A}_k$ *by*

$$\mathfrak{A}_k = \left\{\frac{S(M_k)}{\sqrt{2T^2(M_k)\log\log A(M_k)}} \geqq 1+\delta\right\} \qquad (\delta > 0).$$

Then (under the conditions of Theorem 1) only finitely many $\mathfrak{A}_k$ *can occur with probability* 1.

Proof. By Lemma 1 among the events

$$\mathfrak{B}_k(C) = \left\{\frac{S(M_k)}{\sqrt{2T^2(M_k)\log\log A(M_k)}} \geqq \right.$$

$$\left. \geqq \frac{1}{2C}\sqrt{\frac{T^2(M_k)}{A^2(M_k)}}(1+2\lambda(C)K_1 \max|a_j|) + (1+\varepsilon)\frac{C}{2\sqrt{\frac{T^2(M_k)}{A^2(M_k)}}}\right\}$$

only finitely many will occur. Let now $\{\gamma_k\}$ be a sequence of random variables taking the values $C_1, C_2, \ldots, C_R$ ($k=1, 2, \ldots$; $C_i>0$; $i=1, 2, \ldots, R$). Then among the events $\mathfrak{B}_k(\gamma_k)$ only finitely many will occur (with probability 1) too.

Define a uniform partition of the interval $(\mathscr{K}, K_1)$ (where $\mathscr{K}$ is defined in Lemma 2):

$$C_1 = \mathscr{K}+\frac{K_1-\mathscr{K}}{R}, \quad C_2 = \mathscr{K}+2\frac{K_1-\mathscr{K}}{R}, \ldots, C_R = K_1$$

and let

$$\gamma_k = \begin{cases} \mathscr{K}+i\dfrac{K_1-\mathscr{K}}{R} & \text{if} \quad \mathscr{K}+(i-1)\dfrac{K_1-\mathscr{K}}{R} \leqq \sqrt{\dfrac{T^2(M_k)}{A^2(M_k)}} \leqq \mathscr{K}+i\dfrac{K_1-\mathscr{K}}{R}, \\ 0 & \text{if} \quad \sqrt{\dfrac{T^2(M_k)}{A^2(M_k)}} \leqq \mathscr{K}. \end{cases}$$

Then

$$\overline{\mathfrak{B}_k(\gamma_k)} \cap \bar{F} \subset \left\{\frac{S(M_k)}{\sqrt{2T^2(M_k)\log\log A(M_k)}} \leqq \frac{1}{2\gamma_k}\sqrt{\frac{T^2(M_k)}{A^2(M_k)}}\left(1+2\lambda(C_k)\max_{1\leqq j\leqq M_k}|a_j|\right)+\right.$$

$$+(1+\varepsilon)\frac{\gamma_k}{2\sqrt{\dfrac{T^2(M_k)}{A^2(M_k)}}} \leqq$$

$$\leqq \tfrac{1}{2}\left(1+2\sqrt{\frac{2\log\log A(M_k)}{\mathscr{K}^2A^2(M_k)}}K_1\max|a_j|\right)+(1+\varepsilon)\frac{\sqrt{\dfrac{T^2(M_k)}{A^2(M_k)}}+\dfrac{1}{R}}{2\sqrt{\dfrac{T^2(M_k)}{A^2(M_k)}}} \leqq$$

$$\left.\leqq \tfrac{1}{2}(1+o(1))+\frac{(1+\varepsilon)}{2}\left(1+\frac{1}{R\mathscr{K}}\right) \leqq 1+\delta\right\}$$

if k is large enough and ε and R are chosen in a suitable way. This proves our Lemma 3.

Lemma 4. *Set*

$$F_k = \left\{\max_{M_k\leqq N<M_{k+1}}|S(N)-S(M_k)| \geqq \varepsilon K_2\sqrt{A^2(M_k)\log\log A(M_k)}\right\}$$

for any $\varepsilon>0$. Then (under the conditions of Theorem 1) among the events F_k only finitely many occur with probability 1.

Proof. Since

$$\frac{\sqrt{[A^2(M_{k+1}-1)-A^2(M_k)]\log\log\sqrt{A^2(M_{k+1}-1)-A^2(M_k)}}}{\sqrt{A^2(M_k)\log\log A(M_k)}} \leqq \varepsilon$$

(if k is large enough and θ is chosen near to 1), by Theorem 4 we have

$$\mathbf{P}(F_k) \leqq \mathbf{P}\Big\{ \max_{M_k \leqq N < M_{k+1}} |S(N)-S(M_k)| \geqq$$

$$\geqq K_2 \sqrt{[A^2(M_{k+1}-1)-A^2(M_k)]\log\log\sqrt{A^2(M_{k+1}-1)-A^2(M_k)}}\Big\} \leqq$$

$$\leqq K_3 \exp\big(-2\log\log\sqrt{A^2(M_{k+1}-1)-A^2(M_k)}\big) = O\left(\frac{1}{k^2}\right)$$

i.e., Lemma 4 is proved.

This lemma, the Consequence of Theorem 3 and the simple relation

$$\sqrt{\sum_{j=1}^{M_k} a_j^4 \log\log\sqrt{\sum_{l=1}^{M_k} a_l^4}} = o\left(\sum_{j=1}^{M_k} a_j^2\right)$$

immediately imply

Consequence of Lemma 4. *Let $\varphi_1, \varphi_2, \ldots$ be a sequence of uniformly bounded random variables for which*

$$\int \varphi_{i_1}^2 \varphi_{i_2}^2 \ldots \varphi_{i_k}^2 = 1 \qquad (i_1 < i_2 < \cdots < i_k;\; k = 1, 2, \ldots)$$

and let $a_1, a_2, \ldots$ be a sequence of real numbers satisfying condition (3). *Then*

$$\mathbf{P}\left(\frac{T_n^2}{A_n^2} \to 1\right) = 1.$$

Finally, Theorem 1 is a trivial consequence of Lemmas 3 and 4.

References

[1] S. Takahashi, Notes on the law of the iterated logarithm (in print).
[2] P. Révész, A note to a paper of S. Takahashi (in print).
[3] В. Ф. Гапошкин, К закону повторного логарифма для сильно мултипликативных систем, *Теория вероятностей*, **14** (1969), 511—516.
[4] P. Révész, The law of the iterated logarithm for multiplicative systems, *Indiana Univ. Math. J.*, **21** (1972), 557—564.
[5] K. Azuma, Weighted sums of certain dependent random variables, *Tohoku Math. J.*, **19** (1967), 357—367.

(Received February 3, 1972)

Analytic relations between functional models for contractions

By ION SUCIU in Bucharest (Romania)

To Professor Béla Sz.-Nagy on his 60th anniversary

1. Introduction. The Sz.-Nagy—C. Foiaş functional calculus with bounded analytic functions leads to several results in the study of contractions by means of classical theorems from the analytic function theory.

In this paper, we are going to show how a generalization to functional calculi of two contractions (Theorem 1) of the Harnack inequalities for positive harmonic functions allows us to establish some analytic relations between their Sz.-Nagy—Foiaş functional models (Theorems 2, 3).

We shall use the terminology and notations of [7]. The unitary dilation of the contraction T on the Hilbert space $\mathfrak{H}$ will be denoted by a triplet $[\mathfrak{K}, V, U]$ where $\mathfrak{K}$ is a Hilbert space, V is the isometric embedding of $\mathfrak{H}$ into $\mathfrak{K}$ and U a unitary operator on $\mathfrak{K}$ such that

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

and

$$T^n = V^* U^n V \qquad (n = 0, 1, 2, \ldots)$$

All notations used in [7] for the geometric structure of the unitary dilation will be rewritten here according to this convention. For example

$$\mathfrak{K}_+ = \bigvee_{n=0}^{\infty} U^n V \mathfrak{H}$$

$$\mathfrak{L} = \overline{(U - VTV^*) V \mathfrak{H}}, \quad \mathfrak{L}^* = \overline{(U - VT^* V^*) V \mathfrak{H}}, \quad \mathfrak{L}_* = \overline{(I - UVT^* V^*) V \mathfrak{H}}.$$

D will stand for the unit disc $\{|z| < 1\}$ of the complex plane and X for the unit circle $\{|z| = 1\}$. $C(X)$ will denote the C^*-algebra of all continuous complex valued functions on X and A the subalgebra of $C(X)$ containing all functions in $C(X)$

which have analytic extension in D. For $f \in C(X)$ we shall write

$$f(T) = V^* f(U) V.$$

Then $f \to f(T)$ $(f \in C(X))$ is a linear positive map of $C(X)$ into $B(\mathfrak{H})$ the restriction of which to A is an algebra homomorphism of A into $B(\mathfrak{H})$ such that, for any polynomial p in A, $p(T)$ has its usual meaning.

2. Harnack part. Recall that for an integer j the symbol $T^{(j)}$ stands for T^j if $j \geqq 0$ and for T^{*-j} if $j < 0$. The main result of this section is:

Theorem 1. ([6]) *Let T_1, T_2 be two contractions on a Hilbert space $\mathfrak{H}$. Let $[\mathfrak{K}^1, V_1, U_1]$, $[\mathfrak{K}^2, V_2, U_2]$ be their unitary dilations and a a number such that $0 < a < 1$. The following assertions are equivalent:*

(i) *for any polynomial p in A for which* $\operatorname{Re} p \geqq 0$ *we have*

$$a \operatorname{Re} p(T_1) \leqq \operatorname{Re} p(T_2) \leqq 1/a \operatorname{Re} p(T_1);$$

(ii) *for any positive function u in $C(X)$ we have*

$$au(T_1) \leqq u(T_2) \leqq 1/a\ u(T_1);$$

(iii) *for any positive integer n, any positive $n \times n$-matrix (u_{ij}) over $C(X)$ and any finite system $h_1, \dots, h_n$ of elements in $\mathfrak{H}$ we have*

$$a \sum_{i,j} (u_{ij}(T_1) h_j, h_i) \leqq \sum_{i,j} (u_{ij}(T_2) h_j, h_i) \leqq 1/a \sum_{i,j} (u_{ij}(T_1) h_j, h_i);$$

(iv) *for any positive integer n and any finite system $h_1, \dots, h_n$ of elements in $\mathfrak{H}$ we have*

$$a \sum_{i,j} (T^{(j-i)} h_j, h_i) \leqq \sum_{i,j} (T_2^{(j-i)} h_j, h_i) \leqq 1/a \sum_{i,j} (T_1^{(j-i)} h_j, h_i);$$

(v) *there exists a linear boundedly invertible operator S from $\mathfrak{K}^2$ onto $\mathfrak{K}^1$, such that $\|S\| \leqq 1/\sqrt{a}$ and*

$$SV_2 = V_1, \quad SU_2 = U_1 S.$$

Proof. The implication (i)$\Rightarrow$(ii) follows from the fact that the real parts of the polynomials in A are uniformly dense in the set of real functions in $C(X)$.

The implication (ii)$\Rightarrow$(iii) comes from the Naĭmark dilation theorem as follows: according to (ii) $f \to f(T_2) - af(T_1)$, $(f \in C(X))$, is a positive linear map of $C(X)$ in $B(\mathfrak{H})$. Let $[\mathfrak{K}, V, \pi)$ be the spectral dilation of this map. Thus $\mathfrak{K}$ is a Hilbert space, V is a bounded operator from $\mathfrak{H}$ into $\mathfrak{K}$ and π a representation of $C(X)$ in $B(\mathfrak{K})$ such that

$$f(T_2) - af(T_1) = V^* \pi(f) V \qquad (f \in C(X))$$

(ese for example [1], [5]). Let $(u_{ij})=(g_{ij})^*(g_{ij})$ be a positive $n\times n$: matrix over $C(X)$ and $h_1, \dots, h_n \in \mathfrak{H}$. We have

$$\sum_{i,j}(u_{ij}(T_2)-au_{ij}(T_1)h_j, h_i) = \sum_{i,j}(V^*\pi(u_{ij})Vh_j, Vh_i) =$$

$$= \sum_{i,j}(V^*\pi(\sum_k \bar{g}_{ki}g_{kj})Vh_j, h_i) = \sum_k\sum_{ij}(\pi(g_{ki})^*\pi(g_k)Vh_j, Vh_i) =$$

$$= \sum_k \|\sum_j \pi(g_{kj})Vh_j\|^2 \geqq 0.$$

One obtains the second inequality in (iii) by simmetry.

Taking (iii) with $(u_{ij})=(g_{ij})^*(g_{ij})$, where $g_{1j}(z)=z^j$, $j=1, 2, \dots, n$ and $g_{ij}(z)=0$ for $i\geqq 2$ we obtain (iv).

Let us prove the implication (iv)⇒(v). For any positive integer n and $h_1, \dots, h_n \in H$ we have

$$a\|\sum_j U_1^j V_1 h_j\|^2 = a\sum_{i,j}(V_1^* U_1^{j-i} V_1 h_j, h_i) = a\sum_{ij}(T_1^{(j-i)}h_j, h_i) \leqq$$

$$\leqq \sum_{ij}(T_2^{(j-i)}h_j, h_i) = \sum_{ij}(V_2^* U_2^{j-i} V_2 h_j, h_i) = \|\sum_j U_2^j V_2 h_j\|^2.$$

Thus there exists a bounded operator S from $\mathfrak{R}^2$ into $\mathfrak{R}^1$ such that $\|S\|\leqq 1/\sqrt{a}$ and

$$S\sum_{j=1} U_2^j V_2 h_j = \sum_{j=1} U_1^j V_1 h_j.$$

The second inequality in (iv) shows that S^{-1} exists and $\|S^{-1}\|\leqq 1/\sqrt{a}$. It is clear that

$$SV_2=V_1, \quad SU_2=U_1 S.$$

Since the implication (ii)⇒(i) is obvious, it remains to prove the implication (v)⇒(ii). To do this, let $K=a(S^{-1})^*S^{-1}$. Then $0\leqq K\leqq I$ and it is easy to see that $Kf(U_1)=f(U_1)K$ for any $f\in C(X)$.

Moreover

$$af(T_2) = aV_2^* f(U_2)V_2 = V_2^* aS^{-1}f(U_1)SV_2 =$$

$$= V_1^* a(S^{-1})^* S^{-1} f(U_1)V_1 = V_1^* Kf(U_1)V_1.$$

Let Z be the positive square root of $I-K$. Then Z commutes with $f(U_1)$ for any $f\in C(X)$. Hence for all positive u in $C(X)$ and h in H we have

$$((u(T_1)-au(T_2))h, h) = ((V_1^* u(U_1)V_1 - V_1^* Ku(U_1)V_1)h, h) =$$

$$= (V_1^*(I-K)u(U_1)V_1 h, h) = (V_1^* Z^2 u(U_1)V_1 h, h) = (u(U_1)ZV_1 h, ZV_1 h) \geqq 0.$$

Hence

$$au(T_2) \leqq u(T_1).$$

The second inequality in (ii) is obtained again by symmetry.

The proof of the theorem is complete.

The inequalities contained in Theorem 1 generalize the Harnack inequalities for positive harmonic functions.

We say that T_1 ant T_2 are *Harnack equivalent* if they satisfy one of the (equivalent) assertions of Theorem 1. (Note that T_1 is always Harnack equivalent with $T_2 = T_1$). This equivalence relation determines on the set of all contractions on $\mathfrak{H}$ equivalence classes. Such a class will be called a *Harnack part.* The concept is analogous to that of Gleason parts of the complex homomorphisms of a function algebra (see for example [2]).

Corollary 1. *Two contractions T_1, T_2 are Harnack equivalent if and only if T_1^*, T_2^* are.*

Corollary 2. *If T_1 and T_2 are Harnack equivalent then U_1 and U_2 are unitary equivalent.*

Proof. Using standard arguments we can show that if $S = |S|U$ is the polar decomposition of S then the fact that S has a bounded inverse implies that U is a unitary operator from $\mathfrak{K}^2$ onto $\mathfrak{K}^1$ and $UU_2 = U_1 U$.

Note that, in general, $UV_2 \neq V_1$, thus the two unitary dilations do not coincide.

Corollary 3. *If T is an isometric operator on $\mathfrak{H}$ then the Harnack part containing T reduces to $\{T\}$.*

Proof. Suppose that T_1 is in the same Harnack part as $T_2 = T$ and let $[\mathfrak{K}^1, V_1, U_1]$, $[\mathfrak{K}^2, V_2, U_2]$ be the unitary dilations of T_1, T_2, respectively. Let S be the operator defined in Theorem 1. Since T_2 is an isometry we have $V_2 T_2 = U_2 V_2$. Therefore

$$V_1 T_2 = S V_2 T_2 = S U_2 V_2 = U_1 V_1$$

Hence

$$T_2 = V_1^* V_1 T_2 = V_1^* U_1 V_1 = T_1 .$$

Corollary 4. *Let T_1, T_2 be in the same Harnack part. Then T_1 and T_2 have the same unitary part. In particular, if T_1 is completely non unitary then so is T_2.*

Proof. The maximal subspaces of $\mathfrak{H}$ which reduce T_i $(i=1, 2)$, to unitary operators are

$$\mathfrak{H}_i = \{h \in \mathfrak{H} : U_i V_i h \in V_i H, \qquad n = 0, \pm 1, \pm 2 \ldots\}.$$

For $h \in \mathfrak{H}_1$ and $n = 0, \pm 1, \pm 2, \ldots$ we have

$$U_2^n V_2 h = S^{-1} U_1^n S V_2 h = S^{-1} U_1^n V_1 h \in V_2 \mathfrak{H}.$$

Thus $\mathfrak{H}_1 \subset \mathfrak{H}_2$ and by symmetry $\mathfrak{H}_1 = \mathfrak{H}_2$. Moreover, for $h \in \mathfrak{H}_1 = \mathfrak{H}_2$ we have

$$V_2 T_1 h = S^{-1} V_1 T_1 h = S^{-1} U_1 V_1 h = U_2 S^{-1} V_1 h = U_2 V_2 h = V_2 T_2 h.$$

Thus $T_1 h = T_2 h$.

In [3] C. FOIAŞ proves that the set $B_0 = \{T \in B(\mathfrak{H}), \|T\| < 1\}$ forms a Harnack part, the Harnack part of the contraction 0. Using this result and Corollary 4 one can also prove (see [3]) that there exist Harnack parts different from B_0 and which contain more than one element.

3. Analyticity of the operator S. Suppose that T_1, T_2 are in the same Harnack part and let S be the operator defined in Theorem 1. Since $SV_2 = V_1$, $SU_2 = U_1 S$, $S_2^* U = U_1^* S$, we have

$$S\mathfrak{K}_+^2 = S \bigvee_{n=0}^{\infty} U_2^n V_2 \mathfrak{H} = \bigvee_{n=0}^{\infty} U_1^n S V_2 \mathfrak{H} = \bigvee_{n=0}^{\infty} U_1^n V_1 \mathfrak{H} = \mathfrak{K}_+^1,$$

$$S\mathfrak{K}_-^2 = S \bigvee_{n=0}^{\infty} U_2^* V_2 \mathfrak{H} = \bigvee_{n=0}^{\infty} U_1^* S V_2 \mathfrak{H} = \bigvee_{n=0}^{\infty} U_1^* V_1 \mathfrak{H} = \mathfrak{K}_-^1.$$

Thus

(3.1) $$S\mathfrak{K}_+^2 = \mathfrak{K}_+^1, \quad S\mathfrak{K}_-^2 = \mathfrak{K}_-^1$$

From (3.1) it follows that

$$S^* M_+(\mathfrak{L}^1) = S^*(\mathfrak{K}^1 \ominus \mathfrak{K}_-^1) \subset \mathfrak{K}^2 \ominus \mathfrak{K}_-^2 = M_+(\mathfrak{L}^2).$$

Hence

(3.2) $$S^* M_+(\mathfrak{L}^1) = M_+(\mathfrak{L}^2).$$

Since $S^* U_1^* = U_2^* S$ (3.2) implies

(3.3) $$S^* M(\mathfrak{L}^1) = M(\mathfrak{L}^2).$$

On the other hand

$$S\mathfrak{R}^2 = S \bigcap_{n=0}^{\infty} U_2^n \mathfrak{K}_+^2 = \bigcap_{n=0}^{\infty} U_1^n K_+^1 = \mathfrak{R}^1.$$

Thus

(3.4) $$S\mathfrak{R}^2 = \mathfrak{R}^1.$$

If $h \in \mathfrak{H}$, then $P_{\mathfrak{R}^2} V_2 h = \lim U_2^n T_2^{*n} h$. Thus $SP_{\mathfrak{R}^2} V_2 h = \lim SU_2^n T_2^{*n} h =$ $= \lim U_1^n V_1 T_2^{*n} h$. But $\|U_1^n V_1 T_2^{*n} h\| = \|V_1 T_2^{*n} h\| = \|V_2 T_2^{*n} h\| = \|U_2^n V_2 T_2^{*n} h\|$. Thus $\|SP_{\mathfrak{R}^2} V_2 T_2^* h\| = \lim \|U_1^n V T_2^{*n} h\| = \lim \|U_2^n V_2 T_2^{*n} h\| = \|P_{\mathfrak{R}^2} V_2 h\|$. Which together prove

(3.5) $$\|SP_{\mathfrak{R}^2} V_2 h\| = \|P_{\mathfrak{R}^2} V_2 h\|, \qquad (h \in H).$$

Put

$$\mathfrak{M}^2 = \overline{P_{\mathfrak{R}^2} V_2 \mathfrak{H}}, \quad \mathfrak{M}^1 = S\mathfrak{M}^2.$$

Since

$$U_2^* P_{\mathfrak{R}^2} V_2 h = P_{\mathfrak{R}^2} V_2 T_2^* h \qquad (h \in \mathfrak{H})$$

it follows that $U_2^*\mathfrak{M}^2 \subset \mathfrak{M}^2$ and $U_1^*\mathfrak{M}_1 = U_1^* S\mathfrak{M}^2 = SU_2^*\mathfrak{M}^2 \subset S\mathfrak{M}^2 = \mathfrak{M}^1$. Set

$$T_1' = U_1^*|\mathfrak{M}^1, \quad T_2' = U_2^*|\mathfrak{M}^2, \quad S' = S|\mathfrak{M}^2.$$

According to (3. 5), S' is a unitary operator from $\mathfrak{M}^2$ onto $\mathfrak{M}^1$ and

$$S'T_2' = T_1'S'.$$

It is easy to verify that $U_2^*|\mathfrak{R}^2$ and consequently $U_1^*|\mathfrak{R}^1$ are minimal unitary dilations of T_2' and T_1', respectively. Since $SU_2^* = U_1^* S$, and S extends S', by using standard arguments we can conclude that $S|\mathfrak{R}^2$ is a unitary operator from $\mathfrak{R}^2$ onto $\mathfrak{R}^1$.

From (3. 1) it follows that the operator $S_+ = S|\mathfrak{K}_2^+$ from $\mathfrak{K}_2^+$ onto $\mathfrak{K}_1^+$ has a bounded inverse. Since

$$\mathfrak{L}^i = \mathfrak{K}_+^i \ominus U_i \mathfrak{K}_+^i$$

for any $l \in \mathfrak{L}_*^1$ and $k \in \mathfrak{K}_2^+$ we have

$$(S_+^* l, U_2 k) = (l, S_+ U_2 k) = (l, SU_2 k) = (l, U_1 Sk) = 0.$$

Thus $S_+^* \mathfrak{L}_*^1 \subset \mathfrak{L}_+^2$ and by symmetry we obtain

$$S_+^* \mathfrak{L}_*^1 = \mathfrak{L}_*^2. \tag{3. 6}$$

So we have proved the following

Theorem 2. *Let T_1, T_2 be two Harnack equivalent contractions on $\mathfrak{H}$ and let S be the operator defined in Theorem* 1. *Then*

(i) $S^* M(\mathfrak{L}^1) = M(L^2), \quad S^* M_+(\mathfrak{L}^1) = M_+(\mathfrak{L}^2);$

(ii) *S_+ is a bounded operator from $M_+(\mathfrak{L}_*^2) \oplus \mathfrak{R}^2$ onto $M_+(\mathfrak{L}_*^1) \oplus \mathfrak{R}^1$ which has bounded inverse and*

$$S_+^* \mathfrak{L}_*^1 = \mathfrak{L}_*^2;$$

(iii) *$S|\mathfrak{R}^2$ is a unitary operator from $\mathfrak{R}^2$ onto $\mathfrak{R}^1$.*

From assertions (i) and (ii) of Theorem 2 it follows that $\mathfrak{K}^1 = M(\mathfrak{L}^1)$ $(\mathfrak{K}^1 = M(\mathfrak{L}_*^1))$ if and only if $\mathfrak{K}^2 = M(\mathfrak{L}^2)$ $(\mathfrak{K}^1 = M(\mathfrak{L}_*^2))$. In virtue of Theorem 1. 2, ch. II in [7] we obtain

Corollary 5. *If T_1 and T_2 are Harnack equivalent then T_1 is of class $C_{\cdot 0}$ $(C_{0\cdot}, C_{00})$ if and only if T_2 has this property.*

From assertion (ii) of Theorem 2 and Corollary 1 we conclude

Corollary 6. *If T_1 and T_2 are Harnack equivalent then they have the same defect indices.*

Suppose now that $\mathfrak{H}$ is separable. Taking the Fourier representations of the bilateral shift involved, Theorem 2 allows us to say (according to Lemma 3. 1 Ch.

V in [7]) that in these representations $S^*|M(\mathfrak{L}^1)$ is a bounded analytic function $\{\mathfrak{L}^1, \mathfrak{L}^2, S^*(\lambda)\}$. In the $C_{\cdot 0}$ case S is a bounded analytic function too, namely $\{\mathfrak{L}^2_*, \mathfrak{L}^1_*, S(\lambda)\}$.

In this last case we can establish an analytic relation between characteristic functions as follows.

Theorem 3. *Let T_1 and T_2 be two Harnack equivalent contractions on H. Suppose T_1 (and consequently T_2) belongs to the class $C_{\cdot 0}$. Let $\{\mathfrak{L}^1, \mathfrak{L}^1_*, \theta_1(\lambda)\}$, $\{\mathfrak{L}^2, \mathfrak{L}^2_*, \theta_2(\lambda)\}$ be the characteristic functions of T_1, T_2 respectively. Then there exist bounded, boundedly invertible, analytic functions $\{\mathfrak{L}^2_*, \mathfrak{L}^1_*, S(\lambda)\}$ and $\{\mathfrak{L}^1, \mathfrak{L}^2, \Sigma(\lambda)\}$ such that we have*

$$S(e^{it})^*\theta_1(e^{it}) = \theta_2(e^{it})\sum(e^{it}) \quad a.e.$$

Proof. Let $\{\mathfrak{L}^2_*, \mathfrak{L}^1_*, S(\lambda)\}$ be the bounded analytic function constructed above. From Theorem 2 it follows that

$$S(e^{it})^*\theta_1(e^{it})H^2(\mathfrak{L}^1)\subset\theta_2(e^{it})H^2(\mathfrak{L}^2).$$

Thus we can define the operator Σ by

$$S(e^{it})^*\theta_1(e^{it})u(t) = \theta_2(e^{it})(\sum u)(t) \quad (u\in H^2(\mathfrak{L}^1)).$$

It is easy to verify that the operator Σ commutes with the multiplication with e^{it}. It results that Σ arises as multiplication operator from a bounded analytic function $\{\mathfrak{L}^1, \mathfrak{L}^2, \Sigma(\lambda)\}$. The fact that these functions are boundedly invertible results directly from Lemma 3. 2 ch. V in [7] and Theorem 2 above.

References

[1] W. B. Arveson, Subalgebras of C*-algebras, *Acta Math.*, **123** (1969), 141—224.
[2] E. Bishop, Representing measures for points in a uniform algebra, *Bull. Amer. Math. Soc.*, **70** (1964), 121—122.
[3] C. Foiaş, On the Harnack parts of contractions (to appear).
[4] M. A. Naĭmark, Positive definite operator valued functions on commutative groups, *Bull. (Izvestia) Acad. Sci. USSR (ser. math.)*, **7** (1943), 237—244.
[5] W. F. Stinespring, Positive functions on C*-algebras, *Proc. Amer. Math. Soc.*, **6** (1966), 211—216.
[6] I. Suciu, Harnack inequalities for a functional calculus, *Coll. Math. Soc. János Bolyai.* 5, *Hilbert space operators*, Tihany (Hungary), 1970.
[7] B. Sz.-Nagy—C. Foiaş, *Harmonic analysis of operators in Hilbert space* (Budapest, 1970).

(Received October 28, 1972)

On the convergence of Hermite—Fejér interpolation based on the roots of the Legendre polynomials

By J. SZABADOS in Budapest*)

To Professor B. Sz.-Nagy on his sixtieth birthday

Let $f(x)$ be an arbitrary continuous function in the interval $[-1, 1]$. If $1 > x_1 > > x_2 > \dots > x_n > -1$ are the roots of the Legendre polynomial $P_n(x)$ of degree n then the so-called Hermite—Fejér interpolating polynomials

$$H_n(f, x) = \sum_{k=1}^{n} \frac{1-2xx_k+x_k^2}{1-x_k^2}\left(\frac{P_n(x)}{P_n'(x_k)(x-x_k)}\right)^2$$

of degree $\leqq 2n-1$ satisfy

$$H_n(f, x_k) = f(x_k), \quad H_n'(f, x_k) = 0 \qquad (k = 1, \dots, n).$$

It is well known (see FEJÉR [1]) that

$$\lim_{n\to\infty} H_n(f, x) = f(x) \qquad (|x| < 1)$$

for all continuous $f(x)$, and the convergence is uniform in each closed subinterval of $(-1, 1)$. Our first result improves this statement by giving an estimate for the rate of convergence. In what follows, $\omega_f(t)$ will denote the modulus of continuity of $f(x)$.

Theorem 1. *Let $f(x)$ be a continuous function in $[-1, 1]$ then*

$$|f(x) - H_n(f, x)| =$$

$$= \max\left(\left|f(1) - \tfrac{1}{2}\int_{-1}^{1} f(x)\,dx\right|, \left|f(-1) - \tfrac{1}{2}\int_{-1}^{1} f(x)\,dx\right|\right) \cdot O\left(\frac{1}{n\sqrt{1-x^2}}\right) +$$

$$+ O\left(\omega_f\left(\frac{\log n}{n}\right)\right) \qquad (|x| < 1).$$

*) Research supported by NSF under Grant GP-9493 while the author visited the Ohio State University, Columbus, Ohio.

Proof. Let $n \geqq 3$,

$$m = \left[\frac{n}{\log n}\right], \tag{1}$$

and $q_m(x)$ be the best approximating polynomial of degree $\leqq m$ to $f(x)$ in $[-1, 1]$. Then by the Jackson theorem

$$\|f(x) - q_m(x)\| = O\left(\omega_f\left(\frac{1}{m}\right)\right) \tag{2}$$

($\|.\|$ means the maximum-norm of the corresponding function in $[-1, 1]$). By the linearity and positivity of the operator H_n we have

$$\begin{aligned} |f(x) - H_n(f, x)| &\leqq |f(x) - q_m(x)| + |q_m(x) - H_n(q_m, x)| + \\ &+ |H_n(q_m - f, x)| = O\left(\omega_f\left(\frac{1}{m}\right)\right) + |q_m(x) - H_n(q_m, x)|. \end{aligned} \tag{3}$$

Assume first that $0 \leqq x < 1$. By $m \leqq 2n-1$ (see (1)) we obtain (cf. e.g. SZEGŐ [2], (14. 1. 9))

$$\begin{aligned} q_m(x) - H_n(q_m, x) &= \sum_{k=1}^{n} q_m'(x_k) \frac{P_n(x)^2}{P_n'(x_k)^2 (x - x_k)} = \\ &= \tfrac{1}{2} P_n(x)^2 \sum_{k=1}^{n} \frac{2(1+x_k) q_m'(x_k)}{P_n'(x_k)^2 (1 - x_k^2)} + P_n(x)^2 (1-x) \sum_{k=1}^{n} \frac{q_m'(x_k)}{P_n'(x_k)^2 (1 - x_k)(x - x_k)}. \end{aligned} \tag{4}$$

Here the first sum is the Gauss—Jacobi quadrature for the polynomial $(1+x)q_m'(x)$ of degree $\leqq m \leqq 2n-1$, thus it is equal to

$$\int_{-1}^{1} (1+x) q_m'(x)\, dx = 2q_m(1) - \int_{-1}^{1} q_m(x)\, dx,$$

which, in turn, tends to $2f(1) - \int_{-1}^{1} f(x)\, dx$ as n (and by (1) m) tend to infinity (see (2)). Therefore, by the inequality

$$P_n(x)^2 \leqq \frac{1}{n\sqrt{1-x^2}} \qquad (|x| < 1) \tag{5}$$

(cf. SZEGŐ [2], Theorem 7. 3. 3), we get from (4)

$$\begin{aligned} q_m(x) - H_n(q_m, x) &= \left| f(1) - \tfrac{1}{2} \int_{-1}^{1} f(x)\, dx \right| O\left(\frac{1}{n\sqrt{1-x^2}}\right) + \\ &+ O\left(\frac{(1-x^2)^{3/4}}{\sqrt{n}} \sum_{k=1}^{n} \frac{|q_m'(x_k)|}{P_n'(x_k)^2 (1-x_k)} \left|\frac{P_n(x)}{x - x_k}\right|\right) \qquad (0 \leqq x < 1). \end{aligned} \tag{6}$$

Using a theorem of S. B. STEČKIN [3] which states that for an arbitrary polynomial $q_m(x)$ of degree $\leqq m$

$$|q_m'(x)| = O\left(\frac{m}{\sqrt{1-x^2}}\right)\cdot \omega_{q_m}\left(\frac{1}{m}\right) \qquad (|x|<1)$$

holds, we get by (2) and $\omega_g(t)\leqq 2\|g\|$ that

$$(7)\qquad |q_m'(x)| = O\left(\frac{m}{\sqrt{1-x^2}}\right)\left[\omega_f\left(\frac{1}{m}\right)+\omega_{q_m-f}\left(\frac{1}{m}\right)\right] = O\left(\frac{m\omega_f\left(\frac{1}{m}\right)}{\sqrt{1-x^2}}\right) \qquad (|x|<1).$$

We also need the following estimates in connection with the Legendre polynomials:

$$(8)\qquad \frac{2k-1}{2n+1}\pi \leqq \theta_k = \arccos x_k \leqq \frac{2k}{2n+1} \qquad (k=1,\dots,n)$$

(SZEGŐ [2], Theorem 6. 21. 2), and

$$(9)\qquad P_n'(x_k)\sim\begin{cases}\theta_k^{-3/2}\sqrt{n}\sim k^{-3/2}n^2 & (1\leqq k\leqq n/2),\\ (\pi-\theta_k)^{-3/2}\sqrt{n}\sim(n-k)^{-3/2}n^2 & (n/2\leqq k\leqq n).\end{cases}$$

(SZEGŐ [2], (8. 9. 7)). If we denote $x=\cos\theta\geqq 0$ and $|x-x_j| = \min_{1\leqq k\leqq n}|x-x_k|$ then (5) and (8) imply

$$(10)\qquad \left|\frac{P_n(x)}{x-x_k}\right| = \begin{cases} O(|P_n'(x_j)|) & \text{if } k=j,\\ O\left(\frac{n^2}{\sqrt{j}\,|j^2-k^2|}\right) & \text{if } 1<j+k\leqq n,\ k\neq j,\\ O\left(\frac{n^2}{\sqrt{j}\,|j-k|(2n-j-k)}\right) & \text{if } n<j+k\leqq 2n.\end{cases}$$

Collecting our estimates (7)—(10) we obtain by (1)

$$\frac{(1-x^2)^{3/4}}{n^{1/2}}\sum_{k=1}^{n}\frac{|q_m'(x_k)|}{P_n'(x_k)^2(1-x_k)}\left|\frac{P_n(x)}{x-x_k}\right| = O\left(\frac{j^{3/2}m}{n^2}\right)\omega_f\left(\frac{1}{m}\right)\left[\frac{1}{j^{-3/2}n^2j^3n^{-3}}+\right.$$

$$+\left(\frac{n^2}{\sqrt{j}}\sum_{\substack{1<j+k\leqq n\\ k\neq j}}\frac{1}{k^{-3}n^4k^3n^3|j^2-k^2|}+\right.$$

$$\left.\left.+\sum_{n<j+k\leqq\frac{3}{2}n}\frac{1}{(n-k)^{-3}n^4(n-k)^3n^{-3}|j-k|(2n-j-k)}\right)\right] =$$

$$= O\left(\frac{1}{\log n}\right)\omega_f\left(\frac{\log n}{n}\right)\left(1+\sum_{\substack{k=1\\ k\neq j}}^{n}\frac{1}{|j-k|}\right) = O\left(\omega_f\left(\frac{\log n}{n}\right)\right) \qquad (0\leqq x<1).$$

This together with (3) and (6) means that

$$|f(x)-H_n(f,x)| = \left|f(1)-\tfrac{1}{2}\int_{-1}^{1} f(x)\,dx\right| O\left(\frac{1}{n\sqrt{1-x^2}}\right)+O\left(\omega_f\left(\frac{\log n}{n}\right)\right)$$
$$(0 \leqq x < 1).$$

A similar estimate holds for $-1<x<0$ and the proof of Theorem 1 is complete.

As for the endpoints ± 1, FEJÉR [1] proved that

$$\lim_{n\to\infty} H_n(f, \pm 1) = \tfrac{1}{2}\int_{-1}^{1} f(x)\,dx. \tag{11}$$

G. FREUD [4] raised the question (in a much more general form) whether the *necessary* condition

$$f(\pm 1) = \tfrac{1}{2}\int_{-1}^{1} f(x)\,dx \tag{12}$$

for

$$\lim_{n\to\infty} \|f(x)-H_n(f,x)\| = 0, \tag{13}$$

obtained from (11), is *sufficient* as well. Recently, A. SCHÖNHAGE [5] has given an answer in the affirmative by proving that (12) implies (13). The following result (which is an easy corollary to our Theorem 1) is an improvement of the Schönhage theorem (namely, it contains an estimate for the rate of convergence).

Theorem 2. *Let $f(x)$ be a continuous function in* $[-1, 1]$ *for which* (12) *holds. Then*

$$\|f(x)-H_n(f,x)\| = O\left(\omega_f\left(\frac{\log n}{n}\right)\right).$$

References

[1] L. FEJÉR, Über Interpolation, *Göttinger Nachrichten* (1916), 66—91.
[2] G. SZEGŐ, *Orthogonal Polynomials*, Amer. Math. Soc. Coll. Publ., Vol. XXIII, 1959.
[3] S. B. STEČKIN, Generalization of some inequalities of S. N. Bernstein, *DAN SSSR*, **60/9** (1948), 1511—1514 (Russian).
[4] G. FREUD, On Hermite—Fejér interpolation sequences, *Acta Math. Acad. Sci. Hung.*, **23** (1972), 175—178.
[5] A. SCHÖNHAGE, Zur Konvergenz der Stufenpolynome über den Nullstellen der Legendre Polynome, *Proceedings of the Conference on Abstract Spaces and Approximation*, held in Oberwolfach, 1971, Birkhäuser Verlag (to appear).

(Received April 1, 1972)

Über das im Operatorring enthaltene allgemeine Radikal eines Untermoduls

Von FERENC A. SZÁSZ in Budapest

Professor B. Szőkefalvi-Nagy zu seinem 60. Geburtstag gewidmet

Das Radikal einer algebraischen Struktur gibt die Singularität der Struktur in gewissem Sinne an. Oft kann man für eine Struktur mit einem Radikal (0) einen Struktursatz bestätigen, der die Struktur mit Invarianten charakterisiert. Ein klassisches Beispiel ist dafür der Wedderburn—Artinsche Struktursatz der Artinschen Ringe mit Radikal Null (vgl. N. JACOBSON [6, Chapter III]).

Als Verallgemeinerungen der Vektorräume über einem Schiefkörper, spielen die Operatormoduln über einem Ring in der Mathematik (so z. B. in der Algebra, sowie auch in der Funktionalanalysis) eine wichtige Rolle. Wie es bekannt ist, definiert die Aufgabe 5. 28 des Buches von A. KERTÉSZ [7, Seite 141] ein Radikal eines beliebigen A-Rechtsmoduls M folgendermaßen:

$$I(M) = [x;\ \ x \in A,\ Mx \subseteqq \varphi(M)];$$

wobei $\Phi(M)$ den Frattinischen A-Untermodul von M bedeutet. Man kann leicht zeigen, daß dieses Radikal $I(M)$ immer ein Ideal des Operatorringes A ist, derart, daß $I(M)$ das Jacobsonsche Radikal (siehe H. JACOBSON [6, Chapter I, II]) des Ringes A enthält.

Andererseits hat O. STEINFELD [10] folgendes bewiesen: Sind $\mathscr{I}$ ein Ideal eines Ringes A, und P ein Primideal von $\mathscr{I}$, so ist der Idealquotient

$$P:I = [a;\ \ a \in A,\ \mathscr{I}a + a\mathscr{I} \subseteqq P]$$

ebenfalls ein Primideal des Ringes A. Weiterhin hat O. STEINFELD [11] einige Eigenschaften der Abbildung $x \to (x:y)$ untersucht, wobei $(x:y)$ ein Rechtsresiduum in einem verbandsgeordneten Gruppoid ist. Die Verbandsvereinigung zweier Rechtsresiduuma braucht i. allg. kein Rechtsresiduum zu sein, und Verfasser [13] gibt hinreichende Bedingungen an, damit diese Vereinigung wiederum ein Rechtsresiduum ist.

Das Ziel dieser Arbeit is zweifach. Einerseits werden wir den Kertészschen Begriff des Radikals eines Moduls in zweien Stufen verallgemeinern, in dem wir (1) auch A-Untermoduln statt eines A-Rechtsmoduls und (2) auch Amitsur—Kurošsche allgemeine Radikale und Modulquotienten statt nur der Modulquotienten benützen werden. Andererseits werden wir sowohl das Ergebnis von SREINFELD [10], als auch einige Resultate von ION D. ION [5] verallgemeinern.

Bezüglich der benützten Begriffe verweisen wir auf G. BIRKHOFF [1], N. I. DIVINSKY [2], L. FUCHS [3], L. FUCHS [4, Seiten 189—191 und 195—208], A. KERTÉSZ [7], L. LESIEUR and R. CROISOT [8] weiterhin auf F. SZÁSZ [12] und G. SZÁSZ [14]. Es soll bemerkt werden, daß unsere Beweismethoden manchmal den Methoden von ION D. ION [5] und J. A. RILEY [9] ähnlich sind.

Definition 1. Für die A-Untermoduln L_1 und L_2 eines A-Rechtsmoduls M sei

$$L_1:L_2 = [a; a \in A, \quad L_2 a \subseteqq L_1]$$

Dieses nennen wir den Modulquotienten von L_1 und L_2. Offenbar ist $L_1:L_2$ ein ideal von A.

Definition 2. Sind $\mathbf{R}$ eine allgemeine (Amitsur—Kurošsche) Radikaleigenschaft eines Ringes A, und N ein A-Untermodul eines A-Rechtsmoduls M, so heißt die maximale solche Untermenge $\mathbf{R}(N)$ von A, für die $\mathbf{R}(N)/(N:M) = \mathbf{R}(A)/(N:M)$ gilt, das $\mathbf{R}$-Radikal des A-Untermoduls N.

Behauptung 3. $\mathbf{R}(N)$ *ist ein Ideal von* A, *und es gilt* $(N:M) \subseteqq \mathbf{R}(N)$.

Beweis. Die erste Aussage folgt aus dem zweiten Isomorphismussatz und die zweite Aussage ist trivial.

Behauptung 4. (1) *Ist* $(N:M) = \mathbf{R}(A)$, *so gilt* $\mathbf{R}(N) = \mathbf{R}(A)$. (2) *Sind insbesondere* $\mathbf{R} = \mathbf{J}$ *das Jacobsonsche Radikal, und* $N = \Phi(M)$ *der Frattinische* A-*Untermodul von* M, *so gilt* $\mathbf{J}(N) \supseteqq \mathbf{J}(A)$ (vgl. Aufgabe 5. 28 von KERTÉSZ [7]).

Beweis. (1) Im Falle $(N:M) = \mathbf{R}(A)$ bringt die Definition 2 und Axiom (C) des Buches von Divinsky [2, Seite 3] $\mathbf{R}(N) = \mathbf{R}(A)$ zur Folge. (2) Weiterhin erhält man $\mathbf{J}(A) \subseteqq (\Phi(M):M) = (N:M)$ und wegen $\mathbf{J}(N) \supseteqq (N:M)$ ergibt sich auch $\mathbf{J}(A) \subseteqq \mathbf{J}(N)$.

Definition 5. Sind $\mathbf{R}$ eine allgemeine Radikaleigenschaft von Ringen, und N ein A-Untermodul eines A-Rechtsmoduls M, so heißt N ein R-primärer Untermodul in M, wenn für jedes Ideal I von A, aus der Bedingung

$$mI \subseteqq N \qquad (m \in M)$$

immer $m \in N$ oder $I \subseteqq \mathbf{R}(N)$ folgt.

B e h a u p t u n g 6. *Der A-Untermodul N eines A-Rechtsmoduls M ist für ein Radikal* **R** *dann und nur dann* **R**-*primär in M, wenn für jeden A-Untermodul K von M und für jedes* $a \in A$ *aus* $Ka \subseteqq N$ *immer* $K \subseteqq N$ *oder* $a \in \mathbf{R}(N)$ *folgt.*

B e w e i s. Es seien N ein **R**-primärer A-Untermodul im A-Rechtsmodul M und $k \in K$ ein solches Element, für welches $k \notin N$ gilt. Bezeichne (a) das durch ein Element $a \in A$ erzeugte zweiseitige Hauptideal von A. Dann folgt $k(a) \subseteqq N$ aus $k(Aa) = (kA)a \subseteqq \subseteqq Ka \subseteqq N$, und daher wegen $k \notin N$ für $I = (a)$ auch $(a) \subseteqq \mathbf{R}(N)$ Hiernach ergibt sich $a \in \mathbf{R}(N)$.

Umgekehrt, nehmen wir an, daß die Bedingungen der Behauptung 6 erfüllt sind, und wir zeigen, daß N wirklich **R**-primär in M ist. Es seien $m \in M$, und $K = mA \subsetneqq \subsetneqq N$. Da für jedes $a \in I$ das Umfassen $KI = mA$. $I \subseteqq mI \subseteqq N$ hat $Ka \subseteqq N$ zur Folge, erhält man $a \in \mathbf{R}(N)$ und somit auch $I \subseteqq \mathbf{R}(N)$, w. z. b. w.

B e h a u p t u n g 7. *Es seien* **R** *eine Radikaleigenschaft von Ringen und I ein beliebiges Ideal eines Ringes A. Nehmen wir an, daß* $\mathbf{R}(A/I)$ *für jedes Ideal I von A nilpotent ist. Ist N ein* **R**-*primärer A-Untermodul im A-Rechtsmodul M, so ist das Radikal* $P = \mathbf{R}(N)$ *von N ein Primideal von A.*

B e w e i s. Sind I_1 und I_2 solche Ideale von A, für die $I_1 \cdot I_2 \subseteqq P$ und $I_1 \subsetneqq P$ bestehen, so existiert ein Exponent e, für den $M \cdot (I_1 \cdot I_2)^e \subseteqq N$ und $M \cdot (I_1 \cdot I_2)^{e-1} \subsetneqq N$ erfüllt sind. Es seien $K_1 = M \cdot (I_1 I_2)^{e-1}$, $k_1 \in K_1$ und $k_1 \notin N$. Dann erhält man $(k_1 I_1) I_2 = k_1 (I_1 I_2) \subseteqq K_1 I_1 I_2 \subseteqq N$, und wegen $I_1 \subsetneqq P$ auch $k_1 I_1 \subsetneqq N$. Es sei k_2 ein beliebiges Element von $K_2 = k_1 I_1$ mit $k_2 \notin N$. Dann gilt $k_2 I_2 \subseteqq N$, folglich, wegen $k_2 \notin N$, auch $\mathscr{I}_2 \subseteqq P = \mathbf{R}(N)$, w. z. b. w.

B e h a u p t u n g 8. *Es seien A ein rechtartinscher Ring oder ein rechtsnoetherscher Ring, und* **R** *das Baer—Koethesche obere Nilradikal. Ist N ein* **R**-*primärer A-Untermodul des A-Rechtsmoduls M, so ist das* **R**-*Radikal* $P = \mathbf{R}(N)$ *von N ein Primideal von A.*

B e w e i s. Nach einem Satz von J. LEVITZKI [2] ist $\mathbf{R}(A/\mathscr{I})$ nilpotent für rechtsnoethersche Ringe A. Weiterhin ist $\mathbf{R}(A/\mathscr{I})$ ebenfalls nilpotent nach einem Satz von CH. HOPKINS [2] für rechtsartinsche Ringe. Deshalb ist es genügend Behauptung 7 für diese Fälle anzuwenden.

D e f i n i t i o n 9. Es seien **R** eine Radikaleigenschaft von Ringen, und N ein A-Untermodul eines A-Rechtsmoduls M, weiterhin $P = \mathbf{R}(N)$ das **R**-Radikal von N. Dann heißt N ein P-primärer Untermodul, wenn $P = \mathbf{R}(N)$ ein Primideal von A ist.

S a t z 10. *Es sei* $\mathbf{R}(A/I)$ *nilpotent für jedes Ideal* $\mathscr{I}$ *von A. P ist ein Primideal von A, und N ist ein P-primärer A-Untermodul von M dann und nur dann, wenn die folgenden Bedingungen gleichzeitig erfüllt sind:*

(1) *Für jedes Ideal I von A und jedes $m \in M$ folgt immer aus $mI \subseteqq N$ die Relation $m \in N$ oder $I \subseteqq P$;*

(2) $(N:M) \subseteqq P$;

(3) *Es gibt einen Exponenten e für jedes Hauptideal $(x) \subseteqq P$, derart, daß $M \cdot (x)^e \subseteqq N$ besteht.*

Beweis. Ist N ein P-primärer A-Untermodul, so folgt (1) und (2) unmittelbar aus der Definition 9. Da $\mathbf{R}(A/I)$ nach der Voraussetzung stets nilpotent ist, folgt aus $(x) \subseteqq P = \mathbf{R}(N)$ die Existenz eines Exponenten e derart, daß $M \cdot (x)^e \subseteqq N$ gilt. Dies bedeutet aber, daß auch (3) erfüllt ist.

Umgekehrt, nehmen wir an, daß die Bedingungen (1), (2) und (3) für N bestehen. Wir werden zeigen, daß $P = \mathbf{R}(N)$ gilt, und, daß P ein Primideal von A ist.

Ist nämlich x ein beliebiges Element von P, so gilt wegen (3) $M \cdot (x)^e \subseteqq N$ mit einem geeigneten Exponent e. Es sei P_α ein beliebiges Primideal von A mit der Bedingung

$$(N:M) \subseteqq P_\alpha .$$

Wegen $(x)^{e_\alpha} \subseteqq P_\alpha$ und $\bigcap_\alpha P_\alpha = \mathbf{R}(N)$ erhält man dann $x \in \mathbf{R}(N)$ und somit $P \subseteqq \mathbf{R}(N)$.

Das umgekehrte Umfassen $\mathbf{R}(N) \subseteqq P$ werden wir folgendermaßen zeigen. Ist y ein *beliebiges* Element von $\mathbf{R}(N)$, so gibt es wegen der Nilpotenz von $\mathbf{R}(A/I)$ einen Exponenten e, derart, daß

$$M \cdot (y)^e \subseteqq N \quad \text{und} \quad M \cdot (y)^{e-1} \nsubseteq N$$

bestehen. Ist hier $e = 1$, so gilt $(y) \subseteqq (N:M)$, woher man wegen der Bedingung (2) $(y) \subseteqq P$, folglich $\mathbf{R}(N) \subseteqq P$ erhält. Ist aber $e \geqq 2$, so seien $K = M \cdot (y)^{e-1}$, weiterhin $k \in K$ und $k \notin N$. Wegen

$$k \cdot (y) \subseteqq M \cdot (y)^e \subseteqq N$$

und wegen der Bedingung (1) ergibt sich $(y) \subseteqq P$, folglich $y \in P$ und daher auch $\mathbf{R}(N) \subseteqq P$, womit alles nötiges bewiesen ist.

Behauptung 11. *Es seien* $\mathbf{R}$ *das Baer—Koethesche obere Nilradikal und A ein rechtsnoetherscher oder ein rechtsartinscher Ring. Dann sind die Behauptungen* (1), (2) *und* (3) *des Satzes* 10 *notwendig und hinreichend, damit P ein Primideal von A und N ein P-primärer A-Untermodul von M sind.*

Beweis ist ähnlich dem Beweis der Behauptung 8, und somit kann er weggelassen werden.

Bemerkung 12. Der Ring Z der ganzen rationalen Zahlen ist Noethersch aber nicht Artinsch. Dagen ist ein Zeroring über einer Prüferschen quasizyklischen additiven Gruppe $C(p^\infty)$ Artinsch aber nicht Noethersch (siehe L. Fuchs [3]).

Behauptung 13. *Es sei* $\mathbf{R}(A/I)$ *nilpotent für jedes Ideal* $\mathscr{I}$ *von* A. *Für ein festes Primideal* P, *der Durchschnitt endlich vieler* P*-primärer* A*-Untermoduln* N_j $(j=1, 2, \ldots, k)$ *wiederum* P*-primär im Modul* M.

Beweis. Wegen $N = \bigcap_{j=1}^{e} N_j$ und wegen der Bedingung (2) des Satzes 10 erhält man $(N:M)\subseteqq(N_j:M)\subseteqq P$. Weiterhin gibt es wegen der Bedingung (3) des Satzes 10 für $x\in P$ geeignete Exponenten e_j mit

$$M\cdot(x)^{e_j} \subseteqq N \qquad (j = 1, 2, \ldots, k).$$

Es sei nun $e=\max(e_1, e_2, \ldots, e_k)$. Dann gilt auch $M\cdot(x)^e\subseteqq N$. Es sei weiterhin I ein beliebiges Ideal von A. Bestehen $m\in M$, $m\notin N$ und $mI\subseteqq N$, so gibt es einen Index j mit $m\notin N_j$, woraus wegen $mI\subseteqq N_j$ offenbar $I\subseteqq P=\mathbf{R}(N)$ folgt, w. z. b. w.

Jetzt beweisen wir die folgende Verallgemeinerung des erwähnten Ergebnisses von Steinfeld, die folgendermaßen lautet:

Satz 14. *Es seien* $\mathbf{R}$ *eine Radikaleigenschaft von Ringen derart, daß* $\mathbf{R}(A/I)$ *für jedes Ideal* I *eines Ringes* A *nilpotent ist,* P *ein Primideal von* A, N *ein* P*-primärer* A*-Untermodul des* A*-Rechtsmoduls* M *und* I *ein Ideal von* A *mit der Bedingung* $I\subseteqq P=\mathbf{R}(N)$. *Ist nun*

$$N^* = [m; m\in M, \quad m\mathscr{I} \subseteqq N],$$

so ist N^* *ein* A*-Untermodul von* M, *derart, daß* N^* *auch* P*-primär in* M *ist.*

Beweis. Wir werden den Satz 10 auf N^* anwenden. Ist x ein beliebiges Element von $(N^*:M)$, so bestehen

$$Mx \subseteqq N^* \quad \text{und} \quad Mx\mathscr{I} \subseteqq N.$$

Wegen $I\subseteqq P=\mathbf{R}(N)$ erhält man offenbar $Mx\subseteqq N$, woher $x\in(N:M)$, folglich auch $(N^*:M)\subseteqq(N:M)\subseteqq P$. Deshalb gilt die Bedingung (2) des Satzes 10 für N^*.

Es seien nun B ein Ideal von A und m ein Element von M, derart, daß

$$mB \subseteqq N^* \quad \text{und} \quad m\notin N^*$$

erfüllt sind. Nach der Definition von N^* ergibt sich:

$$mBI \subseteqq N,$$

und wegen $N\subseteqq N^*$ auch $mBI\subseteqq N^*$. Wegen $I\subseteqq P=\mathbf{R}(N)$ und wegen der Bedingung (1) des Satzes 10 hat $BI\subseteqq P$ offenbar $B\subseteqq P$ zur Folge, denn P ist ein Primideal von A.

Wegen der Bedingung (3) des Satzes 10 gibt es für jedes $x\in P=\mathbf{R}(N)$ einen Exponenten e mit $M\cdot(x)^e\subseteqq N$. Wegen $N\subseteqq N^*$ erhält man aber auch $M\cdot(x)^e\subseteqq N^*$, w. z. b. w.

Literaturverzeichnis

[1] G. BIRKHOFF, *Lattice Theory* (Providence, 1948).
[2] N. Divinsky, *Rings and Radicals* (London, 1965).
[3] L. FUCHS, *Abelian Groups* (Budapest, 1958).
[4] L. FUCHS, *Partially Ordered Algebraic Systems* (Oxford—London—New York—Paris, 1963).
[5] ION D. ION, O prezentare simplă a teoriei descompunerilor primare in necomutativ, *Analele Universității București, Ser. Matem.*, **16** (1967), 109—112.
[6] N. JACOBSON, *Structure of Rings*, 2. edition (Providence, 1964).
[7] A. KERTÉSZ, *Vorlesungen über Artinsche Ringe* (Budapest, 1968).
[8] L. LESIEUR—R. CROISOT, *Algèbre noetherienne noncommutative* (Paris, 1963).
[9] J. A. RILEY, Axiomatic primary and tertiary decomposition theory, *Trans. Amer. Math. Soc.*, **105** (1962), 177—201.
[10] O. STEINFELD, On ideal-quotients and prime ideals, *Acta Math. Acad. Sci. Hung.*, **4** (1953), 289—298.
[11] O. STEINFELD, On residuals in partially ordered semigroups, *Publ. Math. Debrecen*, **12** (1965), 107—116.
[12] F. SZÁSZ, *Radikale der Ringe* (im Erscheinen).
[13] F. SZÁSZ, On right residuals in lattice-ordered groupoids, *Math. Nachrichten*, 53 (1972) 69—75.
[14] G. SZÁSZ, *Introduction to Lattice Theory* (New York, 1963).

(Eingegangen am 1. April 1972)

Über die Verallgemeinerung einer Erweiterung von Ringen

Von J. SZENDREI in Szeged

Herrn Prof. B. Szőkefalvi-Nagy zum 60. Geburtstag gewidmet

1. R und R_1, R_2 bezeichnen einen beliebigen (assoziativen) Ring bzw. zwei echte Unterringe von R. Läßt sich jedes Element $r(\in R)$ als eine Summe von der Form

$$r = r_1 + r_2 \qquad (r_1 \in R_1,\ r_2 \in R_2),$$

aufschreiben, d.h. gilt

$$R^+ = R_1^+ + R_2^+,$$

wo das oben angesetzte „+" den Modul eines Ringes bezeichnet, so wird R *in Bezug auf die Addition* (im allgemeineren Sinne) *zerlegbar* genannt.

Als Umkehrung tritt das Erweiterungsproblem auf, aus gegebenen Ringen S_1, S_2 alle Ringe R mit

$$R^+ = R_1^+ + R_2^+, \quad S_1 \simeq R_1, \quad S_2 \simeq R_2$$

zu bestimmen.

Für diese Ringkonstruktion leistet der Ring Z der ganzen Zahlen ein einfaches Beispiel. Sind nämlich die ganzen Zahlen a_1, a_2 teilerfremd, so ist $Z^+ = (a_1)^+ + (a_2)^+$, d.h. Z ist (im allgemeineren Sinne) zerlegbar.

Im Spezialfall $R_1 \cap R_2 = 0$ hat J. Szép [3] dieses Erweiterungsproblem gelöst, und weitere Eigenschaften dieser Konstruktion wurden in den Arbeiten [3, 4] untersucht.

Es sei bemerkt, daß das entsprechende gruppentheoretische Problem, d.h. die Verallgemeinerung des Gruppenproduktes von Zappa—Casadio von L. Rédei und J. Szép [1] gelöst wurde.

2. Wir bezeichnen die Elemente von S_1 mit $0, a, b, \ldots$ und diejenigen von S_2 mit $o, \alpha, \beta, \ldots$. In der Menge **R** der (geordneten) Paare (a, α) $(a \in S_1, \alpha \in S_2)$ definieren wir die Addition und die Multiplikation auf die folgende Weise:

$$(a, \alpha) + (b, \beta) = (a + b, \alpha + \beta), \tag{1}$$

$$(a, \alpha)(b, \beta) = (ab + a^\beta + {}^\alpha b, \alpha^b + {}^a\beta + \alpha\beta), \tag{2}$$

wobei die Funktionen

$$a^{\beta},\ {}^{\alpha}b \in S_1,\ \alpha^{b},\ {}^{a}\beta \in S_2 \tag{3}$$

den folgenden Bedingungen

$$0^{\alpha} = a^{o} = {}^{\alpha}0 = {}^{o}a = 0, \tag{4}$$

$$o^{a} = \alpha^{0} = {}^{a}o = {}^{0}\alpha = o \tag{5}$$

unterworfen sind.

Die sämtlichen Lösungen unseres Problems (bis auf Isomorphie) ergeben sich als Faktorstrukturen von R nach gewissen Kongruenzrelationen.

Es gilt der folgende

Satz. *Sind $S_1^*(\subseteqq S_1)$, $S_2^*(\subseteqq S_2)$ isomorphe Unterringe von S_1, bzw. S_2 mit dem Isomorphismus*

$$A\colon s_* \to As_* \tag{6}$$

angegeben, so definiert man in R die Äquivalenzrelation

$$(a, \alpha) \equiv (b, \beta) \Leftrightarrow A(b-a) = -(\beta - \alpha). \tag{7}$$

Damit diese Äquivalenzrelation eine Kongruenzrelation C von **R** *und die zugehörige Faktorstruktur* $\mathbf{R}/C$ *ein Ring ist, ist notwendig und hinreichend, daß die Funktionen* (3) *(außer* (4) *und* (5)*) die folgenden Bedingungen erfüllen:*

$$A({}^{\alpha}a_*) = \alpha(Aa_*) + \alpha^{a}, \tag{8}$$

$$A^{-1}({}^{a}\alpha_*) = a(A^{-1}\alpha_*) - a^{\alpha_*}, \tag{9}$$

$$A(a_*^{\alpha}) = (Aa_*)\alpha - {}^{a_*}\alpha, \tag{10}$$

$$A^{-1}(\alpha_*^{a}) = (A^{-1}\alpha_*)a - {}^{\alpha_*}a, \tag{11}$$

$$A\left(a(b^{\gamma}) + a^{b_{\gamma}} - (ab)^{\gamma}\right) = -\left({}^{a}({}^{b}\gamma) - {}^{ab}\gamma\right), \tag{12}$$

$$A\left({}^{\alpha}(bc) - ({}^{\alpha}b)c - {}^{(\alpha^{b})}c\right) = -\left(\alpha^{bc} - (\alpha^{b})^{c}\right), \tag{13}$$

$$A\left(a^{\beta\gamma} - (a^{\beta})^{\gamma}\right) = -\left({}^{a}({}^{\beta}\gamma) - {}^{a}({}^{\beta}\gamma) - {}^{(a\beta)}\gamma\right), \tag{14}$$

$$A\left({}^{\alpha}({}^{\beta}c) - {}^{\alpha\beta}c\right) = -\left(\alpha(\beta^{c}) + \alpha^{\beta_{c}} - (\alpha\beta)^{c}\right), \tag{15}$$

$$A\left(a({}^{\beta}c) + a^{(\beta^{c})} - (a^{\beta})c - {}^{(a\beta)}c\right) = -\left({}^{a}(\beta^{c}) - ({}^{a}\beta)^{c}\right), \tag{16}$$

$$A\left({}^{\alpha}(b^{\gamma}) - ({}^{\alpha}b)^{\gamma}\right) = -\left(\alpha({}^{b}\gamma) + \alpha^{(b^{\gamma})} - (\alpha^{b})\gamma - {}^{(\alpha_{b})}\gamma,\right. \tag{17}$$

$$A\left({}^{\alpha}b + {}^{\alpha}c - {}^{\alpha}(b+c)\right) = -(\alpha^{b} + \alpha^{c} - \alpha^{b+c}), \tag{18}$$

$$A\left(a^{\beta} + a^{\gamma} - a^{\beta+\gamma}\right) = -\left({}^{a}\beta + {}^{a}\gamma - {}^{a}(\beta+\gamma)\right), \tag{19}$$

$$A\left(a^{\gamma} + b^{\gamma} - (a+b)^{\gamma}\right) = -({}^{a}\gamma + {}^{b}\gamma - {}^{a+b}\gamma), \tag{20}$$

$$A\left({}^{\alpha}c + {}^{\beta}c - {}^{\alpha+\beta}c\right) = -\left(\alpha^{c} + \beta^{c} - (\alpha+\beta)^{c}\right). \tag{21}$$

Diese Ringe $R = \mathbf{R}/C$ *sind (bis auf Isomorphie) die sämtlichen Lösungen des aufgestellten Problems. In dem Ring* R *bilden die Elemente* $\overline{(a, o)}$, *bzw.* $\overline{(0, \alpha)}$ *je einen Unterring* R_1, R_2, *für die* $R^+ = R_1^+ + R_2^+$, $R_1 \simeq S_1$, $R_2 \simeq S_1$ *gelten. Der Durchschnitt* $R_1 \cap R_2$ *ist isomorph mit* S_1^* *(und mit* S_2^**), und zwar sind die Elemente* $\overline{(a_*, o)}$ *die sämtlichen verschiedenen Elemente dieses Durchschnitts.*

Den Beweis des Satzes bekommen wir in mehreren Schritten. Aus der Definition der Relation (7) folgt unmittelbar, daß diese Relation eine Äquivalenzrelation ist. Jetzt werden wir notwendige und hinreichende Bedingungen dafür angeben, daß die Relation (7) eine Kongruenz in $\mathbf{R}$ ist. Man kann annehmen, daß zwei beliebige äquivalente Elemente in $\mathbf{R}$ die Form

$$(a, \alpha), \qquad (a + a_*,\ \alpha - Aa_*)$$

haben. Diese Elemente sind kongruent dann und nur dann, wenn aus

$$(a, \alpha) = (a + a_*, \alpha - A\alpha_*)$$

für beliebige $(r, \varrho) \in \mathbf{R}$ die Relationen

$$(a, \alpha) + (r, \varrho) \equiv (a + a_*, \alpha - Aa_*) + (r, \varrho),$$

$$(r, \varrho)(a, \alpha) = (r, \varrho)(a + a_*, \alpha - Aa_*),$$

$$(a, \alpha)(r, \varrho) = (a + a_*, \alpha - Aa_*)(r, \varrho)$$

folgen. Die erste ist trivial. Aus den zwei letzten erhalten wir die folgenden notwendigen und hinreichenden Bedingungen:

$$(22) \quad A\left(ra_* + r^{\alpha - Aa_*} + {}^{\varrho}(a + a_*) - r^{\alpha} - {}^{\varrho}a\right) = -(\varrho^{a + a_*} + {}^{r}(\alpha - Aa_*) - \varrho(Aa_*) - {}^{r}\alpha - \varrho^{a}),$$

$$(23) \quad A\left(a_* r + (a + a_*)^{\varrho} + {}^{\alpha - Aa_*}r - a^{\varrho} - {}^{\alpha}r\right) = -((\alpha - Aa_*)^{r} + {}^{a - a_*}\varrho - (Aa_*)\varrho - \alpha^{r} - {}^{a}\varrho).$$

Nachdem setzen wir die Bedingungen (22), (23), d.h. die Existenz die Faktorstruktur $\mathbf{R}/C$ voraus. Jetzt wollen wir die Bedingungen aufstellen, damit $\mathbf{R}/C$ ein Ring ist. Es ist offenbar nach der Definition der Addition in R/C, daß R/C einen Modul bildet. Deshalb ist $\mathbf{R}/C$ ein Ring dann und nur dann, wenn die Multiplikation in $\mathbf{R}/C$ assoziativ und distributiv ist.

Aus der linksseitigen Distributivität

$$(a, \alpha)\left((b, \beta) + (c, \gamma)\right) \equiv (a, \alpha)(b, \beta) + (a, \alpha)(a, \gamma)$$

ergibt sich die Gleichung

$$(24) \quad A\left(a^{\beta} + a^{\gamma} + {}^{\alpha}b + {}^{\alpha}c - a^{\beta + \gamma} - {}^{\alpha}(b + c)\right) = -(\alpha^{b} + {}^{a}\beta + \alpha^{c} + {}^{a}\gamma(\beta + \gamma) - \alpha^{b + c}).$$

Wenn wir in (24) $a=0$, $\beta=\gamma=o$, bzw. $b=c=0$, $\gamma=0$ setzen, so entstehen (18), bzw. (19). Umgekehrt, (24) ergibt sich aus (18) und (19) durch Addition.

Ähnlicherweise bekommen wir die Gleichungen (20), (21) aus der rechtsseitigen Distributivität.

Wegen der Definition der Addition gilt $(a, \alpha) = (a, o)+(0, \alpha)$, deshalb ist es genügend die Assoziativität der Multiplikation für die Elemente von der Form (a, o) und $(0, \alpha)$ zu beweisen. Es kommen die folgenden Produkte

$$\begin{array}{ll} (a,o)(b,o)(c,o), & (0,\alpha)(0,\beta)(0,\gamma) \\ (a,o)(b,o)(0,\gamma), & (0,\alpha)(b,o)(c,o) \\ (a,o)(0,\beta)(0,\gamma), & (0,\alpha)(0,\beta)(c,o) \\ (a,o)(0,\beta)(c,o), & (0,\alpha)(b,o)(0,\gamma) \end{array}$$

in Betracht. Für die zwei ersten Produkte ist die Assoziativität trivial; die folgenden sind mit den Bedingungen (12)—(17) äquivalent.

Bisher haben wir gezeigt, daß $\mathbf{R}/C$ dann und nur dann (existiert und) ein Ring ist, wenn (12)—(23) gelten.

Jetzt werden wir zeigen, daß die Bedingungen (22), (23) unter Berücksichtigung von (12)—(21) mit der Gleichungen (8)—(11) äquivalent sind.

Aus (22) folgen für $a=r=0$, bzw. für $\alpha=\varrho=o$ die Gleichungen (8) und

$$(25) \qquad A(ra_* + r^{-Aa_*}) = -^r(-Aa_*).$$

Hieraus folgt die Gleichung (9) durch die Ersetzungen $-Aa_*=\alpha_*$, und $r=a$. Umgekehrt wird es gezeigt, daß aus (8), (9) die Gleichung (22) folgt. Nach (18) gilt nämlich

$$A\left({}^{\varrho}a + {}^{\varrho}a_* - {}^{\varrho}(a+a_*)\right) = -(\varrho^a + \varrho^{a_*} - \varrho^{a+a_*}),$$

und daraus ergibt sich

$$(25) \qquad A\left({}^{\varrho}(a+a_*) - {}^{\varrho}a\right) = A({}^{\varrho}a_*) + (\varrho^a + \varrho^{a_*} - \varrho^{a+a_*}).$$

Wenn man hier statt $A(\varrho a_*)$ nach (8) $\varrho(Aa_*)+\varrho^a$ setzt, dann bekommt man:

$$(26) \qquad A\left({}^{\varrho}(a+a_*) - {}^{\varrho}a\right) = -\left(\varrho^{a+a_*} - \varrho^a - \varrho(Aa_*)\right).$$

Aus (14) ergibt sich (25). Durch Addition von (25) und (26) entsteht (22). Ähnlicherweise kann die Äquivalenz der Bedingungen (29) und (10), (11) gezeigt werden. Damit haben wir beweisen, daß $\mathbf{R}/C$ dann und nur dann ein Ring ist, wenn (8) bis (21) gelten.

Betrachten wir nunmehr den Ring $\mathbf{R}/C$. Es ist klar, daß die Elemente $\overline{(a, o)}$, bzw. $\overline{(0, \alpha)}$ je einen Unterring R_1, bzw. R_2 von R/C bilden und

$$\mathbf{R}/C = R_1^+ + R_2^+, \quad R_1 \simeq S_1, \quad R_2 \simeq S_2$$

gelten. Der Durchschnitt D von R_1 und R_2 besteht aus denjenigen Elementen $\overline{(a, 0)}$, die sich auch als $\overline{(0, \alpha)}$ schreiben lassen. Da nach (7) gilt $\overline{(a, o)}=\overline{(0, \alpha)}$ dann und nur dann, wenn $Aa=\alpha$ gilt, so folgt hieraus, daß D aus den Elementen $\overline{(a^*, o)}$ besteht und diese Elemente auch schon verschieden sind, ferner die Isomorphie $D \simeq S_1^*(\simeq S_2^*)$ gilt. Das bedeutet, daß jeder Ring $\mathbf{R}/C$ eine Lösung unseres Problems gibt.

Wir haben noch zu zeigen, daß umgekehrt jede Lösung des aufgestellten Erweiterungsproblems (bis auf Isomorphie) unter den im Satz angegeben Ringen $\mathbf{R}/C$ vorkommt. Es kann aufgenommen werden, daß R eine Lösung ist, d.h.

$$R^+ = R_1^+ + R_2^+, \quad S_1 \simeq R_1, \quad S_2 \simeq R_2$$

gelten, wobei R_1, R_2 Unterringe von R sind. Jedes Element in R hat die Form $a+\alpha$ $(a \in R_1; a \in R_2)$; $a+\alpha = b+\beta$ dann und nur dann, wenn $b-a=-(\beta-\alpha) \in \in R_1 \cap R_2$. Hat ein Element in R die Form $a\alpha$, bzw. αa, so läßt es sich auch in der Form

$$a\alpha = a^\alpha + {}^a\alpha, \text{ bzw. } \alpha a = {}^\alpha a + \alpha^a$$

schreiben, wobei $a^\alpha, {}^\alpha a \in R_1$ und ${}^a\alpha, \alpha^a \in R_2$. Hiernach haben wir:

$$(a+\alpha)+(b+\beta) = (a+b)+(\alpha+\beta),$$

$$(a+\alpha)(b+\beta) = ab+a\beta+\alpha b+\alpha\beta = (ab+a^\beta+{}^\alpha b)+(\alpha^b+{}^a\beta+\alpha\beta).$$

Ist insbesondere $0\alpha=\alpha 0=0$ und $0a=a0=0$, so wird aufgenommen, daß (4), (5) gelten. Wegen (3)—(5) ist R eine Struktur (mit zwei Verknüpfungen) und wegen (1), (2) und (27), (28) gilt die Homomorphie

$$\mathbf{R} \sim R((a, \alpha) \to a+\alpha).$$

Ist C die zugehörige kompatible Klasseneinteilung von $\mathbf{R}$, so entsteht die Isomorphie

$$\mathbf{R}/C \simeq R.$$

Damit haben wir den Satz bewiesen.

Literaturverzeichnis

[1] L. Rédei und J. Szép, Die Verallgemeinerung der Theorie des Gruppenproduktes von Zappa-Casadio, *Acta Sci. Math.*, **16** (1955), 165—170.
[2] J. Szendrei, Über die Szépschen Ringerweiterungen, *Acta Sci. Math.*, **21** (1960), 166—172.
[3] J. Szép, Über eine neue Erweiterung von Ringen. I, *Acta Sci. Math.*, **19** (1958), 51—62.
[4] J. Szép, Über eine neue Erweiterung von Ringen. II, *Acta Sci. Math.*, **20** (1959), 202—214.

(Eingegangen am 15. Juli 1972)

On an identity of Shih-Chieh Chu

By LAJOS TAKÁCS in Cleveland (Ohio, U.S.A.)*)

Dedicated to Professor Béla Szőkefalvi-Nagy on the occasion of his sixtieth birthday July 29, 1973

1. On a formula of Shih-Chieh Chu. In a Chinese treatise of 1303 SHIH-CHIEH CHU [7] described the arithmetic triangle as an ancient invention for determining the terms in the expansion of $(a+b)^n$ where $n=1, 2, \dots$. In his treatise Shih-Chieh Chu obtained several remarkable relations for binomial coefficients, but gave no proofs. One of the identities, attributed to Shih-Chieh Chu, can be expressed in modern notation as follows: If k and n are nonnegative integers, then

$$\sum_{j=0}^{k} \binom{k}{j}^2 \binom{n+2k-j}{2k} = \binom{n+k}{k}^2. \tag{1}$$

As usual, the binomial coefficient $\binom{x}{j}$ is defined for every real or complex x as follows:

$$\binom{x}{j} = \frac{x(x-1)\dots(x-j+1)}{j!} \qquad (j=1,2,\dots),$$

$$\binom{x}{0} = 1, \quad \binom{x}{j} = 0 \qquad (j=-1,-2,\dots).$$

Since both sides of (1) are polynomials of degree $2k$ in the variable n, it follows that (1) also holds if n is any real or complex number.

More details about the rich contents of the treatise of SHIH-CHIEH CHU [7] can be found in the books by Y. MIKAMI [22 pp. 89—98, 124] and J. NEEDHAM and L. WANG [24 pp. 41, 46—47, 133—141]. In 1867 JEN-SHOU LI [18] brought the identity (1) to light in the fourth part of his book entitled Mathematical Studies. The problem of how this identity was found and how it can be proved aroused the interest of several mathematicians.

*) This research was supported by a grant from the National Science Foundation.

In 1937 Yung Chang, a scholar of the history of mathematics, mentioned the identity (1) to G. Szekeres who called the attention of P. Turán to it. Both G. Szekeres and P. Turán have found rather complicated proofs for (1). In 1939 their proofs were published in Chinese in an article by YUNG CHANG [6]. In 1954 P. TURÁN [35] published a Hungarian version of his proof.

In his proof G. Szekeres used mathematical induction and some known summation formulas for the binomial coefficients and proved (1) as a particular case of a more general relation. P. Turán's proof is analytical and based on some properties of the Legendre polynomials. By using Hurwitz's formula and Rodrigues's formula for the Legendre polynomials, P. Turán demonstrated that

$$\sum_{j=0}^{k}\binom{k}{j}\binom{k+j}{j}(x-1)^{k-j}=\sum_{j=0}^{k}\binom{k}{j}^2 x^j \tag{2}$$

holds for every x and $k=0, 1, 2, \ldots$, and by calculating the derivative

$$\frac{1}{k!}\frac{d^k}{dz^k}\left(\frac{z^k}{(x+z)^{k+1}}\right)_{z=-1}$$

in two different ways he proved that

$$\sum_{n=0}^{\infty}\binom{n+k}{k}^2 x^n=(1-x)^{-2k-1}\sum_{j=0}^{k}\binom{k}{j}^2 x^j \tag{3}$$

holds for $|x|<1$. If we form the coefficient of x^n in the power series expansion of the right-hand side of (3), then we obtain (1).

It is interesting to mention that if we form the coefficient of z^k in both forms of the polynomial

$$(1+z)^k[1+z+(x-1)z]^k=(1+z)^k(1+xz)^k,$$

then we get (2). Formula (2) is a particular case of a more general result found in 1947 by W. LJUNGGREN [20]. Furthermore, formula (3) is a particular case of Saalschütz's theorem for hypergeometric functions. The general case of (3) was found in 1890 by L. SAALSCHÜTZ [28], [29]. (See also W. N. BAILEY [1 p. 9] and A. ERDÉLYI, W. MAGNUS, F. OBERHETTINGER and F. G. TRICOMI [8 pp. 65—66].) Let us mention that formula (3) was used in 1943 by O. BOTTEMA and S. C. VAN VEEN [4] in their studies of the game of billiard. (See also Problem 10 in Chapter 11 of W. FELLER [9].)

Subsequent to the investigations of G. Szekeres and P. Turán, simple proofs were found for (1) by L. TAKÁCS [34], J. SURÁNYI [31], [32], G. HUSZÁR [12], J. MÁTÉ [21], L. CARLITZ [5], J. SEITZ (see [16]), J. KAUCKÝ [14], [15], [16], [17], T. S. NANJUNDIAH [23] and R. L. GRAHAM and J. RIORDAN [11]. Most of these proofs are elementary and based on some well-known summation formulas for binomial coefficients. L. CARLITZ [5] showed that (1) immediately follows from a summation formula found in 1890 by L. SAALSCHÜTZ [28]. J. SEITZ and J. KAUCKÝ (see [16]) demonstrated that (3) can also be obtained by calculating

$$\frac{1}{k!}\frac{\partial^{2k}}{\partial x^k \partial z^k}\left(\frac{1}{1-xz}\right)_{z=1}$$

in two different ways.

By using a combinatorial approach, in 1955 J. SURÁNYI [31], [32] proved that if k, l and n are nonnegative integers, then we have

$$(4)\qquad \sum_{j=0}^{k}\binom{k}{j}\binom{l}{j}\binom{n+k+l-j}{k+l}=\binom{n+k}{k}\binom{n+l}{l}.$$

If $l=k$, then (4) reduces to (1). Since both sides of (4) are polynomials of degree $k+l$ in the variable n, therefore (4) also holds if n is any real or complex number. This was demonstrated by LO-KENG HUA (see [31]) who proved that the left-hand side and the right-hand side of (4) have the same roots in n.

In 1958 T. S. NANJUNDIAH [23] proved a generalization of (1), which is equivalent to (4), and mentioned that by the same method it is possible to prove the following more general identity:

$$(5)\qquad \sum_{j=0}^{k}\binom{l+x-y}{j}\binom{k+y-x}{k-j}\binom{y+j}{k+l}=\binom{x}{k}\binom{y}{l}$$

which holds for any nonnegative integers k and l and for any real or complex x and y. In fact T. S. Nanjundiah proved (5) only for $x=y$. In the case where x and y are nonnegative integers (5) was proved in 1970 by M. T. L. BIZLEY [2].

In this paper we shall prove that if k and m are nonnegative integers for which $0\leqq k\leqq m$, and l and c are real or complex numbers, then we have

$$(6)\qquad \sum_{j=0}^{k}\binom{l}{j}\binom{m-l}{k-j}\binom{c+j}{m}=\binom{c}{m-k}\binom{c-m+k+l}{k}.$$

Since both sides of (6) are polynomials of degree k in l, therefore it is sufficient to prove (6) for $l=0, 1, \ldots, k$ and this implies that (6) holds for every l.

Formula (6) is a generalization of (1). If in (6) we replace (k, l, m, c) by $(k, k, 2k, n+k)$, and if we reverse the order in the sum, then we obtain (1). If in (6) we replace (k, l, m, c) by $(k, k, k+l, n+l)$, and if we reverse the order in the sum, then we obtain (4). If in (6) we replace (k, l, m, c) by $(k, l+x-y, k+l, y)$, then we obtain (5).

Now for any real or complex x and q let us define the Gaussian binomial coefficients $\begin{bmatrix} x \\ j \end{bmatrix}$ as follows:

$$\begin{bmatrix} x \\ j \end{bmatrix} = \begin{cases} \dfrac{(q^x-1)(q^{x-1}-1)\dots(q^{x-j+1}-1)}{(q^j-1)(q^{j-1}-1)\dots(q-1)} & \text{for} \quad q \neq 1, \\ \dfrac{x(x-1)\dots(x-j+1)}{j!} & \text{for} \quad q = 1, \end{cases}$$

if $j = 1, 2, \dots$; $\begin{bmatrix} x \\ 0 \end{bmatrix} = 1$, and $\begin{bmatrix} x \\ j \end{bmatrix} = 0$ if $j = -1, -2, \dots$.

By using the Gaussian binomial coefficients $\begin{bmatrix} x \\ j \end{bmatrix}$ with an arbitrary q we can generalize (6) in the following way

$$(7) \qquad \sum_{j=0}^{k} \begin{bmatrix} l \\ j \end{bmatrix} \begin{bmatrix} m-l \\ k-j \end{bmatrix} \begin{bmatrix} c+j \\ m \end{bmatrix} q^{(k-j)(l-j)} = \begin{bmatrix} c \\ m-k \end{bmatrix} \begin{bmatrix} c-m+k+l \\ k \end{bmatrix}$$

which holds if k and m are nonnegative integers for which $0 \leqq k \leqq m$ and l and c are real or complex numbers. This follows from a result which was found in 1968 by E. M. Wright [36]. Actually, if we replace (k, l, m, c) by $(s, s-k, r+s-k, r+t-k)$ in (7), then we obtain Wright's formula (1) in his paper. For another proof of (7) we refer to H. W. Gould [10].

If we replace (k, l, m, c) by $(k, k, k+l, n+l)$ in (7), then we obtain a particular case of (7) which was found in 1965 by Gh. Pic [25].

We note that if in (6) and (7) we replace (k, l, m, c) by $(y, x+a-b, x+y, -a-1)$ and if we form the sum with respect to $i = a-j$, then we obtain the identities of R. P. Stanley [30] and H. W. Gould [10].

It is of some interest to mention that in 1965 Gh. Pic [25] generalized (1) by calculating

$$\frac{(z+x)^{\alpha+\beta}}{z^\alpha} \frac{d^k}{dz^k} \left(\frac{z^{k+\alpha}}{(z+x)^{k+\alpha+\beta+1}} \right)_{z=-1}$$

in two different ways. By using Rodrigues's formula for the Jacobi polynomials (see G. Szegő [33 pp. 58—99]) he obtained a generalization of (3). In the particular case where $\alpha = \beta = 0$ his result reduces to (3).

Finally, we note that by using the elements of the calculus of finite differences we can easily deduce a large number of formulas which have some resemblance to (6), but are actually of different types. Such a formula is the identity

$$\sum_{j=0}^{k} \binom{x}{j+r} \binom{j+r}{r} \binom{r+y}{k-j} = \binom{x}{r} \binom{x+y}{k}$$

which holds for every x and every y if k and r are nonnegative integers. We can easily prove this formula by using Newton's expansion. (See G. BOOLE [3 pp. 11—14] and CH. JORDAN [13 pp. 77—79].) More identities of similar types are given by many authors. See, for example, J. RIORDAN [26 pp. 14—15], [27 pp. 14—17], and D. J. LEWIS [19 p. 58].

2. A probabilistic proof for (6). We shall prove the identity (6) by using probabilistic methods. The following proof makes possible various generalizations of (6).

In the following proof we assume that k, l, m are nonnegative integers satisfying the inequalities $0 \leqq k \leqq m$ and $0 \leqq l \leqq m$ and that c is an arbitrary real or complex number.

In proving (6) we shall use the following result from the theory of probability. Suppose that a box contains l white balls and $m-l$ black balls. We draw k $(0 \leqq k \leqq m)$ balls without replacement. Let us suppose that every outcome of this random trial has the same probability, and denote by ν the number of white balls drawn. Then we have

$$\mathbf{P}\{\nu = j\} = \frac{\binom{k}{j}\binom{m-k}{l-j}}{\binom{m}{l}} = \frac{\binom{l}{j}\binom{m-l}{k-j}}{\binom{m}{k}} \tag{8}$$

for $j=0, 1, \ldots, k$, and

$$\mathbf{E}\left\{\binom{\nu}{r}\right\} = \sum_{j=r}^{k} \binom{j}{r} \mathbf{P}\{\nu = j\} = \frac{\binom{k}{r}\binom{l}{r}}{\binom{m}{r}} \tag{9}$$

for $r=0, 1, \ldots, k$.

Now we shall calculate $\binom{m}{k}\mathbf{E}\left\{\binom{c+\nu}{m}\right\}$ in two different ways. First, by (8) we obtain that

$$\binom{m}{k}\mathbf{E}\left\{\binom{c+\nu}{m}\right\} = \binom{m}{k}\sum_{j=0}^{k}\binom{c+j}{m}\mathbf{P}\{\nu = j\} = \sum_{j=0}^{k}\binom{c+j}{m}\binom{l}{j}\binom{m-l}{k-j}. \tag{10}$$

Second, by (9) we obtain that

$$\binom{m}{k}\mathbf{E}\left\{\binom{c+\nu}{m}\right\} = \binom{m}{k}\sum_{r=0}^{m}\binom{c}{m-r}\mathbf{E}\left\{\binom{\nu}{r}\right\} = \tag{11}$$

$$= \sum_{r=0}^{m}\binom{c}{m-k}\binom{c-m+k}{k-r}\binom{l}{r} = \binom{c}{m-k}\binom{c-m+k+l}{k}.$$

Here we used that

$$\binom{m}{k}\binom{c}{m-r}\binom{k}{r} = \binom{m}{r}\binom{c}{m-k}\binom{c-m+k}{k-r}.$$

If we compare (10) and (11), then we obtain (6) which was to be proved.

We note that if we calculate $\binom{m}{k}\mathbf{E}\left\{\binom{c-\nu}{m}\right\}$ in two different ways, then we obtain that

$$\sum_{j=0}^{k}\binom{l}{j}\binom{m-l}{k-j}\binom{c-j}{m} = \binom{c-k}{m-k}\binom{c-l}{k} \tag{12}$$

whenever k, l, m are nonnegative integers for which $0 \leqq k \leqq m$ and $0 \leqq l \leqq m$ and c is an arbitrary real or complex number. We can immediately see that (12) holds also if l is an arbitrary real or complex number.

In (12) the left-hand side can be obtained by (8). If we take into consideration that

$$\mathbf{E}\left\{\binom{c-\nu}{m}\right\} = (-1)^m \mathbf{E}\left\{\binom{m-c-1+\nu}{m}\right\},$$

then the right-hand side of (12) can be obtained by (11).

Of course (12) can also be obtained from (6) by substituting $m-c-1$ for c.

3. A generalization of (6). The previous probabilistic proof makes it possible to generalize (6) in various ways which we shall demonstrate by the following example.

Let us suppose that a box contains m cards each marked by one of the numbers $1, 2, \dots, s$. Denote by l_i the number of cards marked i $(i=1, 2, \dots, s)$. We have $l_1+l_2+\dots+l_s = m$. We draw k $(0 \leqq k \leqq m)$ cards without replacement. Let us suppose that every outcome of this random trial has the same probability and denote by ν_i the number of cards marked i among the k cards drawn. Then we have

$$\mathbf{P}\{\nu_1 = j_1, \quad \nu_2 = j_2, \dots, \nu_s = j_s\} = \frac{\binom{l_1}{j_1}\binom{l_2}{j_2}\cdots\binom{l_s}{j_s}}{\binom{m}{k}} \tag{13}$$

whenever $j_1+j_2+\dots+j_s = k$, and

$$\mathbf{E}\left\{\binom{\nu_1}{r_1}\binom{\nu_2}{r_2}\cdots\binom{\nu_s}{r_s}\right\} = \frac{\binom{l_1}{r_1}\binom{l_2}{r_2}\cdots\binom{l_s}{r_s}\binom{m-r}{k-r}}{\binom{m}{k}} \tag{14}$$

where $r = r_1+r_2+\dots+r_s$.

Let $c_1, c_2, \ldots, c_s$ be arbitrary real or complex numbers, and $1 \leqq t \leqq s$. By (13) we obtain that

$$\binom{m}{k} \mathbf{E}\left\{\binom{c_1+v_1}{m} \cdots \binom{c_t+v_t}{m}\right\} = \sum_{j_1+\cdots+j_s=k} \binom{c_1+j_1}{m} \cdots \binom{c_t+j_t}{m} \binom{l_1}{j_1} \cdots \binom{l_s}{j_s} =$$

$$\sum_{j_1+\cdots+j_t \leqq k} \binom{c_1+j_1}{m} \cdots \binom{c_t+j_t}{m} \binom{l_1}{j_1} \cdots \binom{l_t}{j_t} \binom{m-l_1-\cdots-l_t}{k-j_1-\cdots-j_t},$$

and by (14) we obtain that

$$\binom{m}{k} \mathbf{E}\left\{\binom{c_1+v_t}{m} \cdots \binom{c_t+v_t}{m}\right\} = \binom{m}{k} \sum_{r_1+\cdots+r_t \leqq k} \binom{c_1}{m-r_1} \cdots \binom{c_t}{m-r_t} \mathbf{E}\left\{\binom{v_1}{r_1} \cdots \binom{v_t}{r_t}\right\} =$$

$$= \sum_{r_1+\cdots+r_t \leqq k} \binom{c_1}{m-r_1} \cdots \binom{c_t}{m-r_t} \binom{l_1}{r_1} \cdots \binom{l_t}{r_t} \binom{m-r_1-\cdots-r_t}{k-r_1-\cdots-r_t}.$$

By comparing the above two formulas we obtain that if $k, l_1, \ldots, l_t$ are non-negative integers for which $k \leqq m$ and $l_1 + \cdots + l_t \leqq m$, then

$$\sum_{j_1+\cdots+j_t \leqq k} \binom{l_1}{j_1} \cdots \binom{l_t}{j_t} \binom{m-l_1-\cdots-l_t}{k-j_1-\cdots-j_t} \binom{c_1+j_1}{m} \cdots \binom{c_t+j_t}{m} =$$

(15)

$$\sum_{r_1+\cdots+r_t \leqq k} \binom{l_1}{r_1} \cdots \binom{l_t}{r_t} \binom{m-r_1-\cdots-r_t}{k-r_1-\cdots-r_t} \binom{c_1}{m-r_1} \cdots \binom{c_t}{m-r_t}$$

holds for arbitrary real or complex $c_1, c_2, \ldots, c_t$. If $t=1$, then (15) can be reduced to (6).

Another generalization of (6) can be obtained in the following way. Let us suppose that we have s boxes and the ith box ($i=1, 2, \ldots, s$) contains l_i white balls and $m-l_i$ black balls. We draw k_1 balls from the first box, k_2 balls from the second box, and so on, k_s balls from the s-th box without replacement. Denote by v the total number of white balls drawn. If we calculate the expectation $\mathbf{E}\left\{\binom{c+v}{m}\right\}$ in two different ways, then we obtain another generalization of (6).

References

[1] W. N. Bailey, *Generalized Hypergeometric Series,* Cambridge University Press, 1935. (Reprinted by Stechert-Hafner, New York, 1964.)

[2] M. T. L. Bizley, A generalization of Nanjundiah's identity, *Amer. Math. Monthly,* **77** (1970), 863—865.

[3] G. Boole, *A Treatise on the Calculus of Finite Differences.* First edition, 1860. Second edition, 1872. (Second edition reprinted by Dover, New York, 1960.)

[4] O. Bottema and S. C. van Veen, Kansberekeningen bij het biljartspel, *Nieuw Archief voor Wiskunde,* **22** (1943), 15—33 and 123—158.

[5] L. Carlitz, On a problem in the history of Chinese mathematics, *Mat. Lapok,* **6** (1955), 219—220. (Hungarian)

[6] Yung Chang, To chi pi lei su chun (Proofs for summation formulas for binomial coefficients), *Science,* **23** (1939), 647—663. (Chinese)

[7] Shih-Chieh Chu, Ssu Yuan Yü Chien (Precious Mirror of the Four Elements), 1303. (Chinese)

[8] A. Erdélyi, W. Magnus, F. Oberhettinger and F. G. Tricomi, *Higher Transcendental Functions.* Vol. I, McGraw Hill (New York, 1953).

[9] W. Feller, *An Introduction to Probability Theory and its Applications.* Vol. 1., John Wiley (New York, 1950). (Second edition 1957; Third edition 1968.)

[10] H. W. Gould, A new symmetrical combinatorial identity, *J. Combinatorial Theory,* (A) **13** (1972), 278—286.

[11] R. L. Graham and J. Riordan, The solution of a certain recurrence, *Amer. Math. Monthly,* **73** (1966), 604—608.

[12] G. Huszár, On a problem in the history of Chinese mathematics, *Matematikai Lapok,* **6** (1955), 36—38. (Hungarian)

[13] Ch. Jordan, *Calculus of Finite Differences* (Budapest, 1939). (Reprinted by Chelsea, New York, 1947.)

[14] J. Kaucký, Proof of the equation $\sum_{k=0}^{\nu}\binom{2N+\nu-k}{\nu-k}\binom{N}{k}^2=\binom{N+\nu}{\nu}^2$, *Sbornik Vojenské Tech. Akad. A. Zápotockého, Brno,* **5** (1957), 6—8. (Czech)

[15] J. Kaucký, Remarks on a paper of P. Turán, *Mat.-Fyz. Časopis Slovensk. Akad. Vied.,* **12,** (1962), 212—216. (Czech)

[16] J. Kaucký, On a problem in the history of Chinese mathematics, *Mat.-Fyz. Časopis Slovensk. Akad. Vied.,* **13** (1963), 32—40. (Czech)

[17] J. Kaucký, Une nouvelle démonstration élémentaire de la formule combinatoire de Li Jen-Shu, *Mat.-Fyz. Časopis Slovensk. Akad. Vied.,* **14** (1964), 50—53.

[18] Jen-Shou Li, Tsê Ku Hsi Chai Suan Hsüeh *(Tsê Ku Hsi Home Mathematical Studies),* Nanking, 1867. (Chinese)

[19] D. J. Lewis, Diophantine equations: p-adic methods, *Studies in Number Theory,* Editor: W. J. LeVeque. *MAA Studies in Mathematics.,* Vol. 6. Mathematical Association of America, 1969, 25—75.

[20] W. Ljunggren, Et elementaert bevis for en formel av A. C. Dixon, *Norsk Mat. Tidskrift,* **29** (1947), 35—38.

[21] J. Máté, On a problem in the history of Chinese mathematics, *Mat. Lapok,* **7** (1956), 112—113. (Hungarian)

[22] Y. Mikami, *The Development of Mathematics in China and Japan,* Teubner (Leipzig, 1913). (Reprinted by Chelsea, New York, 1961.)

[23] T. S. Nanjundiah, Remark on a note of P. Turán, *Amer. Math. Monthly,* **65** (1958), 354.

[24] J. Needham and L. Wang, *Science and Civilisation in China.* Vol. 3., Mathematics and the Sciences of the Heavens and the Earth (Cambridge University Press, 1959).

[25] Gh. Pic, On a combinatorial formula, *Studia Univ. Babeş—Bolyai,* Ser. Math. Phys., **10,** No. 2, 7—15. (Romanian)

[26] J. Riordan, *An Introduction to Combinatorial Analysis,* John Wiley, New York, 1958.

[27] J. Riordan, *Combinatorial Identities,* John Wiley (New York, 1968).

[28] L. Saalschütz, Eine Summationsformel, *Z. für Math. und Phys.,* **35** (1890) 186—188.

[29] L. SAALSCHÜTZ, Über einen Specialfall der hypergeometrischen Reihe dritter Ordnung, *Z. für Math. und Phys.*, **36** (1891), 278—295 and 321—327.
[30] R. P. STANLEY, *Ordered structures and partitions*, Ph. D. Dissertation (Harvard University, 1970).
[31] J. SURÁNYI, Remarks on a problem in the history of Chinese mathematics, *Mat. Lapok,* **6** (1955), 30—35. (Hungarian)
[32] J. SURÁNYI, On a problem of old Chinese mathematics, *Publ. Math. Debrecen,* **4** (1956), 195—197.
[33] G. SZEGŐ, *Orthogonal Polynomials,* American Mathematical Society Colloquium Publications, Vol. XXIII, American Mathematical Society (Providence, R. I., 1939). (Second ed. 1958. Third ed. 1967.)
[34] L. TAKÁCS, Remark on a paper of P. Turán entitled "On a problem in the history of Chinese mathematics", *Mat. Lapok,* **6** (1955), 27—29. (Hungarian)
[35] P. TURÁN, On a problem in the history of Chinese mathematics, *Mat. Lapok,* **5** (1954), 1—6. (Hungarian)
[36] E. M. WRIGHT, An identity and applications, *Amer. Math. Monthly*, **75** (1968), 711—714.

(Received November 29, 1972)

Bemerkung über die Konvergenz der Orthogonalreihen

Von KÁROLY TANDORI in Szeged

Herrn Professor B. Szőkefalvi-Nagy zum 60. Geburtstag gewidmet

1. A. SZÉP hat das folgende Problem aufgeworfen. Gibt es eine quadratisch-integrierbare orthogonale Entwicklung

$$\sum a_n \varphi_n(x),$$

die fast überall im Orthogonalitätsintervall konvergiert, und eine fast überall divergierende Teilreihe

(2) $$\sum a_{n_k} \varphi_{n_k}(x)$$

besitzt?

In dieser Note werden wir auf dieses Problem eine Antwort geben. Man soll aber gewisse Bemerkungen vorausschicken. Die Behauptung kann offensichtlich nicht für jede Koeffizientenfolge $\{a_n\}$ mit $\Sigma a_n^2 < \infty$ richtig sein. Gilt nämlich die Menchoff—Rademachersche Bedingung

$$\sum a_n^2 \log^2 n < \infty,$$

dann folgt auf Grund des Menchoff—Rademacherschen Satzes ([1], [2]), daß für beliebige Indexfolge $\{n_k\}$ die Reihe (2) fast überall konvergiert. Also kann die Behauptung nur für solche Koeffizientenfolgen $\{a_n\}$ richtig sein, für die (3) nicht erfüllt ist.

Wir werden Folgendes beweisen.

Satz. *Es sei $\{a_n\}$ eine monoton nichtwachsende Folge von positiven Zahlen, für die $\Sigma a_n^2 < \infty$ und*

(3) $$\sum a_n^2 \log^2 n = \infty$$

erfüllt sind. Dann gibt es ein im Intervall (0, 1) *orthonormiertes System* $\{\varphi_n(x)\}$ *und* *eine Indexfolge* $\{n_k\}$ *derart, daß die Reihe* (1) *in* (0, 1) *fast überall konvergiert, und die Teilreihe* (2) *in* (0, 1) *fast überall divergiert.*

In der Arbeit [4] wurde es bewiesen, daß unter den Bedingungen dieses Satzes ein orthonormiertes System $\{\varphi_n(x)\}$ derart existiert, daß die Reihe (1) fast überall divergiert.

Vor dem Beweis dieses Satzes werden wir ein anderes Problem erwähnen. Für eine Folge $\{a_n\}$ setzen wir

$$\|\{a_n\}\| = \sup \sqrt{\int_0^1 \sup_i (a_1\varphi_1(x) + \cdots + a_i\varphi_i(x))^2\,dx},$$

wobei das Supremum für jedes in (0, 1) orthonormierte System $\{\varphi_n(x)\}$ gebildet ist. In [5] wurde es bewiesen, daß aus $\|\{a_n\}\| < \infty$ die Konvergenz der Reihe (1) bei jedem orthonormierten System $\{\varphi_n(x)\}$ fast überall folgt; weiterhin aus $\|\{a_n\}\| = \infty$ folgt, daß ein orthonormiertes System $\{\varphi_n(x)\}$ derart existiert, daß die Reihe (1) fast überall divergiert.

Das erwähnte Problem ist folgendes: Sei $\|\{a_n\}\| = \infty$ und $\Sigma a_n^2 < \infty$. Gibt es dann ein orthonormiertes System $\{\varphi_n(x)\}$ und eine Indexfolge $\{n_k\}$ derart, daß die Reihe (1) fast überall konvergiert, und die Teilreihe (2) fast überall divergiert?

2. Zum Beweis des Satzes werden wir gewisse Hilfssätze benützen.

Hilfssatz I. ([4]) *Es seien* $c(\geqq 1)$ *und* $p(\geqq 2)$ *positive ganze Zahlen. Es kann ein im Intervall* [0, 5] *orthonormiertes System von Treppenfunktionen* $\{f_l(c, p; x)\}$ $(l = 1, \ldots, 2p)$ *mit den folgenden Eigenschaften angegeben werden: zu jedem Punkt* $x \in \left(\frac{2}{c}, \frac{3}{c}\right)$ *gibt es eine von* x *abhängige natürliche Zahl* $m(x)$ *derart, daß die Funktionswerte* $f_l(c, p; x)$ $(l = 1, \ldots, m(x))$ *positiv sind, und*

$$f_1(c, p; x) + \cdots + f_{m(x)}(c, p; x) \geqq A_1 \sqrt{cp} \log p$$

gilt, wo A_1 *eine positive, von* x, c *und* p *unabhängige Zahl ist.*

(In Folgendem bezeichnen $A_2, A_3, \ldots$ positive, von den Parametern unabhängige Konstanten.)

Hilfssatz II. ([2]) *Es seien* d *und* q *positive ganze Zahlen,* $0 < d < q$. *Zu jedem Indexpaar* (i, j) *mit* $1 \leqq i \leqq q$, $1 \leqq j \leqq q$ *und* $|i - j| = d$ *soll eine von Null verschiedene Zahl* $\alpha_{i,j}$ *zugeordnet werden; wir bezeichnen mit* β_d *das Maximum der absoluten Beträge der Zahlen* $\alpha_{i,j}$. *In jedem Intervall* (u, v) *mit*

$$v - u > 2\beta_d\,.$$

können wir dann Treppenfunktionen $\varphi_l(x)$ $(l=1, \ldots, q)$ derart definieren, daß die folgenden Bedingungen erfüllt werden:

$$|\varphi_l(x)| = 1 \qquad (u < x < v;\ l = 1, \ldots, q),$$

$$\int_u^v \varphi_i(x)\varphi_j(x)\,dx = -\alpha_{i,j} \qquad (|i-j| = d,\ 1 \leqq i \leqq q,\ 1 \leqq j \leqq q),$$

$$\int_u^v \varphi_i(x)\varphi_j(x)\,dx = 0 \qquad (i \neq j,\ |i-j| \neq d,\ 1 \leqq i \leqq q,\ 1 \leqq j \leqq q).$$

3. Beweis des Satzes. Es sei $N_m = 2(2+\cdots+2^m)$ $(m=1, 2, \ldots)$. Durch Induktion werden wir ein im Intervall $(0, 1)$ orthonormiertes System von Treppenfunktionen $\varphi_n(x)$ $(n=1, 2, \ldots)$ und zwei Folgen von einfachen Mengen[1]) $E_m(\subseteqq(0, 1))$, $F_m(\subseteqq(0, 1))$ $(m=1, 2, \ldots)$ derart definieren, daß die folgenden Bedingungen erfüllt sind:

Die Mengen E_m $(m=1, 2, \ldots)$ sind stochastisch unabhängig, und für jedes m $(m=1, 2, \ldots)$ gelten

$$m(E_m) \geqq \frac{1}{20}\min(1, N_{m+1}a^2_{2N_{m+1}}\log^2 2N_{m+1}), \tag{4}$$

$$m(F_m) \leqq \frac{1}{2m^2}. \tag{5}$$

Weiterhin für jedes m $(m=1, 2, \ldots)$ bestehen

$$\max_{N_m < n \leqq N_{m+1}} \left| \sum_{l=N_m+1}^{n} a_{2l}\varphi_{2l}(x) \right| \geqq \frac{\sqrt{5}}{6}A_1 \qquad (x \in E_m), \tag{6}$$

$$\varphi_{2l}(x) = -\varphi_{2l-1}(x) \qquad (x \notin F_m;\ l = N_m+1, \ldots, N_{m+1}). \tag{7}$$

Es sei

$$\varphi_n(x) = r_n(x) \qquad (n=1, \ldots, 8=2N_1),$$

wobei $r_n(x) = \operatorname{sign}\sin 2^n\pi x$ die n-te Rademachersche Funktion bezeichnet. Es sei m_0 eine natürliche Zahl. Nehmen wir an, daß die Funktionen $\varphi_n(x)$ $(n=1, \ldots, 2N_{m_0})$ und die Mengen $E_1, \ldots, E_{m_0-1}, F_1, \ldots, F_{m_0-1}$ derart definiert sind, daß diese Funktionen Treppenfunktionen sind, in $(0, 1)$ ein orthonormiertes System bilden, diese Mengen einfach sind, weiterhin $E_1, \ldots, E_{m_0-1}$ stochastisch unabhängig sind, und (4)—(7) für $m = 1, \ldots, m_0-1$ erfüllt werden.

Wir wenden den Hilfssatz I im Falle

$$c = \left[\frac{1}{N_{m_0+1}a^2_{2N_{m_0+1}}\log^2 2N_{m_0+1}} + 1\right], \quad p = 2^{m_0+1}$$

---

[1]) Eine Menge wird einfach genannt, wenn sie als Vereinigung endlichvieler Intervalle entsteht.

an ([α] bezeichnet den ganzen Teil von α). Auf Grund des Hilfssatzes I und der Monotonität der Folge $\{a_n\}$ gilt

$$\max_{1\leqq n\leqq 2.2^{m_0+1}} \sum_{l=1}^{n} a_{2N_{m_0}+2l} f_l(c,p;x) \geqq \tag{8}$$

$$\geqq A_1 \frac{1}{\sqrt{N_{m_0+1}}\, a_{2N_{m_0+1}} \log 2N_{m_0+1}} \sqrt{m_0+1}\, a_{2N_{m_0+1}} \sqrt{2^{m_0+1}} \geqq \frac{A_1}{6}$$

$$\left(x \in \left(\frac{2}{c}, \frac{3}{c}\right) = E'_{m_0}\right).$$

Offensichtlich gilt

$$m(E'_{m_0}) \geqq \tfrac{1}{2} \min(1, N_{m_0+1} a^2_{2N_{m_0+1}} \log^2 2N_{m_0+1}). \tag{9}$$

Betrachten wir die Treppenfunktionen

$$\psi_{2l}(x) = \begin{cases} f_l(c,p;x) & (0<x<5), \\ 0 & \text{sonst} \end{cases} \qquad \psi_{2l-1}(x) = \begin{cases} -f_l(c,p;x) & (0<x<5), \\ 0 & \text{sonst} \end{cases}$$

$(l = 1, \ldots, 2.2^{m_0+1})$. Wir setzen

$$\alpha_{i,j} = \int_0^5 \psi_i(x)\psi_j(x)\,dx \qquad (1 \leqq i,j \leqq 4.2^{m_0+1}).$$

Dann gelten

$$\alpha_{i,i} = 1 \quad (1 \leqq i \leqq 4.2^{m_0+1}), \qquad \alpha_{i,j} = 1 \quad (|i-j| = 1,\ 1 \leqq i,j \leqq 4.2^{m_0+1}),$$

$$\alpha_{i,j} = 0 \quad (i \neq j,\ |i-j| \neq 1,\ 1 \leqq i,\ j \leqq 4.2^{m_0+1}).$$

Durch Anwendung des Hilfssatzes II kann man im Intervall (5, 8) Treppenfunktionen $\bar{\varphi}_1(x), \ldots, \bar{\varphi}_{4.2^{m_0+1}}(x)$ derart definieren, daß

$$\int_5^8 \bar{\varphi}_i(x)\bar{\varphi}_j(x)\,dx = -\alpha_{i,j} \quad (|i-j| = 1,\ 1 \leqq i,\ j \leqq 4.2^{m_0+1}),$$

$$\int_5^8 \bar{\varphi}_i(x)\bar{\varphi}_j(x)\,dx = 0 \quad (i \neq j,\ |i-j| \neq 1,\ 1 \leqq i,\ j \leqq 4.2^{m_0+1}),$$

$$\int_5^8 \bar{\varphi}_i^2(x)\,dx = 3 \qquad (i = 1, \ldots, 4.2^{m_0+1})$$

bestehen.

Wir bilden die Funktionen

$$\bar{\psi}_l(x) = \begin{cases} \sqrt{\dfrac{5}{2-1/m_0^2}}\, \psi_l\left(\dfrac{5}{1-1/2m_0^2}\right) & \left(0 < x < 1 - \dfrac{1}{2m_0^2}\right), \\ m_0\, \bar{\varphi}_l\left(3.2m_0^2\left(x - \left(1 - \dfrac{1}{2m_0^2}\right)\right) + 5\right) & \left(1 - \dfrac{1}{2m_0^2} < x < 1\right) \end{cases}$$

($l=1, \ldots, 4.2^{m_0+1}$). Nach den Obigen bilden die Treppenfunktionen $\bar{\psi}_l(x)$ ($l=1, \ldots, 4.2^{m_0+1}$) ein orthonormiertes System im Intervall (0, 1). Weiterhin auf Grund der Definition der Funktionen $\bar{\psi}_l(x)$ ($l=1, \ldots, 4.2^{m_0+1}$) und (8), (9) folgt

$$(10) \qquad \max_{1 \leqq n \leqq 2^{m_0+1}} \sum_{l=1}^{n} a_{2N_{m_0}+2l} \bar{\psi}_{2l}(c, p; x) \geqq \frac{\sqrt{5}}{6} A_1$$

$$\left(x \in \left(\frac{2}{c} \frac{1-1/2m_0^2}{5}, \frac{3}{c} \frac{1-1/2m_0^2}{5}\right) = E''_{m_0}\right),$$

$$(11) \qquad m(E''_{m_0}) \geqq \frac{1}{20} \min(1, N_{m_0+1} a^2_{2N_{m_0+1}} \log^2 2N_{m_0+1}),$$

$$(12) \quad \bar{\psi}_{2l}(x) = -\bar{\psi}_{2l-1}(x) \qquad \left(x \in \left(1-\frac{1}{2m_0^2}, 1\right) = F'_{m_0};\ l = 1, \ldots, 4.2^{m_0+1}\right),$$

$$(13) \qquad m(F'_{m_0}) = \frac{1}{2m_0^2}.$$

Für ein endliches Intervall $I=[a, b]$ und für eine in (0, 1) definierte Funktion $f(x)$ setzen wir

$$f(I; x) = \begin{cases} f\left(\dfrac{x-a}{b-a}\right) & (a < x < b), \\ 0 & \text{sonst}, \end{cases}$$

weiterhin für eine Menge H bezeichnen wir mit $H(I)$ diejenige Menge, die aus H durch die Transformation $y=(b-a)x+a$ entsteht.

Da die Funktionen $\varphi_n(x)$ ($n=1, \ldots, 2N_{m_0}$) Treppenfunktionen sind, und die Mengen $E_1, \ldots, E_{m_0-1}$, $F_1, \ldots, F_{m_0-1}$ einfach ,weiterhin die Mengen $E_1, \ldots, E_{m_0-1}$ stochastisch unabhängig sind, können wir das Intervall (0, 1) in endlichviele paarweise disjunkte Intervalle I_r ($1 \leqq r \leqq \varrho$) derart zerlegen, daß jede Funktion $\varphi_n(x)$ ($1 \leqq n \leqq 2N_{m_0}$) in jedem Intervall I_r ($1 \leqq r \leqq \varrho$) konstant ist, und die Mengen E_m ($1 \leqq m \leqq m_0-1$), F_m ($1 \leqq m \leqq m_0-1$) die Vereinigung gewisser I_r sind. Die zwei Hälften von I_r bezeichnen wir mit I'_r, und I''_r ($1 \leqq r \leqq \varrho$).

Dann setzen wir

$$\varphi_{n+2N_{m_0}}(x) = \sum_{r=1}^{\varrho} \bar{\psi}_n(I'_r; x) - \sum_{r=1}^{\varrho} \bar{\psi}_n(I''_r; x) \qquad (n=1, \ldots, 4.2^{m_0+1}),$$

$$E_{m_0} = \bigcup_{r=1}^{\varrho} (E''_{m_0}(I'_r) \cup E''_{m_0}(I''_r)), \qquad F_{m_0} = \bigcup_{r=1}^{\varrho} (F'_{m_0}(I'_r) \cup F'_{m_0}(I''_r)).$$

Nach dem Obigen und nach (10)—(13) ist es offensichtlich, daß die Treppenfunktionen $\varphi_n(x)$ ($n=1, \ldots, 2N_{m_0+1}$) in (0, 1) ein orthonormiertes System bilden, die Mengen $E_m, F_m (\subseteqq (0, 1))$ ($m=1, \ldots, m_0$) einfach und, die Mengen E_m ($m=1, \ldots, m_0$) stochastisch unabhängig sind, weiterhin (3)—(7) für $m=1, \ldots, m_0$ bestehen. Das erwähnte Funktionensystem $\{\varphi_n(x)\}$ und die Mengenfolgen $\{E_m\}$, $\{F_m\}$ mit den erwähnten Eigenschaften erhalten wir durch Induktion.

Wegen der Monotonität der Folge $\{a_n\}$ ergibt sich

$$\sum a_n^2 \log^2 n \leqq A_2 \sum a_{2n}^2 \log^2 2n \leqq A_3 \sum_{m=1}^{\infty} N_{m+1} a_{2N_{m+1}}^2 \log^2 2N_{m+1}.$$

Daraus und aus (3) folgt

$$\sum \min (1, N_{m+1} a_{2N_{m+1}}^2 \log^2 2N_{m+1}) = \infty. \tag{14}$$

Es sei $E = \overline{\lim\limits_{m\to\infty}}\, E_m$. Da die Mengen E_m $(m=1, 2, \ldots)$ stochastisch unabhängig sind, und (4) für jede natürliche Zahl m besteht, auf Grund von (14), durch Anwendung des zweiten Borel—Cantellischen Lemmas erhalten wir

$$m(E) = 1. \tag{15}$$

Gilt aber $x \in E$, dann besteht (6) für unendlich viele Indizes m, und so divergiert die Reihe

$$\sum a_{2n} \varphi_{2n}(x). \tag{16}$$

Aus (15) folgt, daß die Reihe (16) fast überall divergiert.

Es sei $F = \overline{\lim\limits_{m\to\infty}}\, F_m$. Da (5) für jedes m erfüllt wird, durch Anwendung der ersten Borel—Cantellischen Lemmas erhalten wir:

$$m(F) = 0. \tag{17}$$

Da auch (7) für jedes m erfüllt ist, erhalten wir daß im Falle $x \notin F$ eine von x abhängige natürliche Zahl $m_0 = m_0(x)$ derart existiert, daß

$$s_{2N}(x) - s_{2N_{m_0}}(x) = \sum_{n=2N_{m_0}+1}^{2N} a_n \varphi_n(x) = \sum_{l=N_{m_0}+1}^{N} (a_{2l-1} - a_{2l}) \varphi_{2l}(x), \tag{18}$$

$$s_{2N+1}(x) - s_{2N_{m_0}}(x) = \sum_{l=N_{m_0}+1}^{N} (a_{2l-1} - a_{2l}) \varphi_{2l}(x) + a_{2N+1} \varphi_{2N+1}(x)$$

$(N = N_{m_0}+1, \ldots)$ bestehen.

Wegen der Monotonität der Folge $\{a_n\}$ und wegen der Normierung der Funktionen erhalten wir

$$\sum_{l=1}^{\infty} (a_{2l-1} - a_{2l}) \int_0^1 |\varphi_{2l}(x)|\, dx \leqq \sum_{l=1}^{\infty} (a_{2l-1} - a_{2l}) \sqrt{\int_0^1 \varphi_{2l}^2(x)\, dx} \leqq$$

$$\leqq \sum_{n=2}^{\infty} (a_{n-1} - a_n) = a_1 < \infty,$$

woraus durch Anwendung des B. Levischen Satzes folgt, daß die Reihe

$$\sum_{l+1}^{\infty} (a_{2l-1} - a_{2l}) \varphi_{2l}(x) \tag{19}$$

fast überall konvergiert. Es sei G die Menge derjenigen Punkte x, in denen die Reihe (19) konvergiert. Dann ist also

$$m(G) = 1. \tag{20}$$

Wegen $\Sigma a_n^2 < \infty$ und $\int_0^1 \varphi_n^2(x)dx = 1$ $(n = 1, 2, \ldots)$ erhalten wir

$$\lim_{n\to\infty} a_n \varphi_n(x) = 0 \tag{21}$$

fast überall. Es sei H die Menge derjenigen Punkte x, in denen (21) besteht. Dann ist also

$$m(H) = 1. \tag{22}$$

Es sei $\Omega = ((0, 1)\setminus F) \cap G \cap H$. Nach dem Obigen, auf Grund von (18) und (21) erhalten wir, daß im Falle $x \in \Omega$, die Reihe (1) konvergiert. Weiterhin aus (17), (20) und (22) folgt

$$m(\Omega) = 1.$$

Damit haben wir bewiesen, daß die Reihe (1) fast überall konvergiert.

Wir haben also die Divergenz der Reihe (2) im Falle $n_k = 2k$ $(k = 1, 2, \ldots)$ bewiesen.

Schriftenverzeichnis

[1] D. E. MENCHOFF, Sur les séries de fonctions orthogonales (Première partie), *Fundamenta Math.*, **4** (1923), 82—105.

[2] D. E. MENCHOFF, Sur les séries de fonctions orthogonales bornées dans leur ensemble, *Recueil math. Moscou*, **3** (43) (1938), 103—120.

[3] H. RADEMACHER, Einige Sätze über Reihen von allgemeinen Orthogonalfunktionen, *Math. Annalen*, **87** (1922), 112—138.

[4] K. TANDORI, Über die orthogonalen Funktionen. I, *Acta Sci. Math.*, **18** (1957), 57—130.

[5] K. TANDORI, Über die Konvergenz der Orthogonalreihen. II, *Acta Sci. Math.*, **25** (1964), 219—232.

(Eingegangen am 18. April 1972)

A necessary and sufficient condition of optimality for Markovian control problems

By D. VERMES in Szeged

To Professor Béla Sz.-Nagy on his 60th birthday

1. Introduction

The vital importance of the Markov property of processes in control theory was first pointed out by R. BELLMAN in [1]. As a consequence of this perception he was able to give a necessary and sufficient condition for the optimality of control for discrete-time Markov decision processes. For continuous-time processes there are no general results known. In the deterministic case the Bellman equation could only be proved to be a sufficient condition of optimality (cf. [1], [2]). On the consideration of stochastic problems there arise additional difficulties from the absence of a sufficiently general definition of the controlled Markov process. FLEMING [6], [7], MANDL [12], WONHAM [15] have given conditions for the optimality of the control of diffusion processes, governed by stochastic differential equations, GRIGELIONIS and SHIRYAEV [8] have investigated the properties of the optimal expense function of processes for which the value of the control parameter has been allowed to change only in fixed "switching times", while KUSHNER [11] has regarded families of Markov processes with the strategy space as an index set and has given a sufficient condition of optimality for them.

The aim of the present paper is to prove a necessary and sufficient optimality condition for the control of general discrete or continuous-time Markov processes, by employing the functional analytic theory of Markov processes, which was developed by HILLE, YOSIDA, FELLER, DYNKIN, and others in the last 25 years.

First we give a sufficiently general formulation of the optimal stochastic control problem. Controlled Markov processes are defined similarly to [11], [13] as families of processes. Unless otherwise stated, our considerations hold both for discrete- and continuous-time processes, i.e., both the set N of the non-negative integers and the set R^+ of the non-negative reals are allowed to be the time axis T. In our investigations we shall necessarily consider time dependent controls and so time-inhomogeneous (non-autonomous) processes. The theory of one parameter semigroups of operators — which will be extensively used throughout the paper — is only adequate to describe time-homogeneous processes. But in [13] it was proved, that to every inhomogeneous process on the state space E', there exists an equivalent homogeneous process on the product state space $E = T \times E'$. This way we do not restrict the generality by assuming that our processes are homogeneous and their state space is of the form $T \times E'$. Studying Markov processes we shall refer to the monographs of DYNKIN [4], [5]. For control strategies we allow measurable mappings of the state space into the control region, that is, we consider problems with Markov feed-back control policies. While in the deterministic case the effectivity of open loop and feed-back strategies are the same, for stochastic problems one is obliged to consider closed loop controls (cf. [3]). In the present paper we call a control strategy optimal, if it minimizes the expected loss of the integral form for an arbitrary initial state.

The main result of this paper, proved in the third chapter is a necessary and sufficient condition for the existence of optimal control strategies. The theorem is formulated in the form of a generalized boundary value problem and also presents a characterisation of the optimal strategy. It can be regarded as a comman generalization of the results of papers [1], [6], [7], [8], [11], [12], [13] and [15]. Two corollaries of the main theorem are given in the fourth chapter. They specialize the main theorem for problems having some local control dependence, and this allows a simplification of the optimality conditions. The recursive equations of the dynamic programming and the optimality condition of FLEMING for controlled diffusion processes are also derived as examples for the application of the optimality theorem.

For the computation of an optimal control strategy, or for the verification of its existence, by the optimality theorem given below one has to solve a generalized boundary value problem. Considerations about the existence of its solution can be found in the papers of KRYLOV [9], [10]. The results of the present paper, especially Corollary 2, can be used with success to prove the optimality of a strategy given in advance, e.g., by means of a necessary condition of optimality, and they serve to an extent as substitutions for the existence theorems of optimal control.

The author would like to thank Prof. P. H. MÜLLER, Prof. H. LANGER, Dr. sc. V. NOLLAU from the Technische Universität Dresden and Prof. K. TANDORI from the University Szeged for their valuable suggestions during the course of this work.

2. Optimal Markovian control problems

First we make some further assumptions about the structure of the time axis, the state space and the control region, which will enable us to apply the results of [4], [5] to our problems. The time axis T will be considered as a measurable topological space with its usual topology $\mathscr{C}_T$ and the σ-field of its Borel sets $\mathscr{T}$. The state space E is defined as the topological product of the time axis and some topological measurable space $(E', \mathscr{C}', \mathscr{E}')$, together with the product σ-field $\mathscr{E} = \mathscr{T} \otimes \mathscr{E}'$. Further on we assume that all open sets and also all one-point-sets of E are measurable. We shall call the measurable space $(D, \mathscr{D})$ with measurable one-point-sets the control region, while $\mathbf{B}(E, \mathscr{E})$ denotes the Banach space of all bounded measurable real valued functions on $(E, \mathscr{E})$, with the usual supremum norm.

Suppose we are given an open subset $G \subset E$, a class $\mathscr{U}$ of measurable mappings from $(G, \mathscr{C} \cap G, \mathscr{E} \cap G)$ into $(D, \mathscr{D})$, and a family of (homogeneous) right-continuous strong Markov processes $\{\Xi^U; U \in \mathscr{U}\}$ with $\Xi^U = (\xi^U, \zeta^U, \mathscr{M}_t^U, \mathbf{P}_x^U)$ on the state space $(E, \mathscr{C}, \mathscr{E})$, stopped at the first exit from G. In this paper we call $\{\Xi^U; U \in \mathscr{U}\}$ a controlled Markov process with the target set $E \setminus G$ and the class $\mathscr{U}$ of admissible control strategies (or policies) if the following conditions are satisfied:

a) If $U_1 \in \mathscr{U}$, $U_2 \in \mathscr{U}$ and if, I is an arbitrary (finite or infinite) subinterval of T, and

$$U(s, x') := \begin{cases} U_1(s, x') & \text{for} \quad s \in I, \ x \in E \\ U_2(s, x') & \text{for} \quad s \notin I, \ x \in E \end{cases}$$

then $U \in \mathscr{U}$.

b) For the transition function of Ξ^U we have

$$P^U(t, (s, x'), \ (T \setminus \{s+t\}) \times E') = 0 \tag{2.1}$$

for any $U \in \mathscr{U}$, $x = (s, x') \in E$, $t \in T$ ($\{v\}$ is the set containing the only point v).

c) The first exit time τ^U of the process Ξ^U from G is a Markov time for Ξ^U.

d) If I denotes an arbitrary interval in T then $U_1(x) = U_2(x)$ implies

$$\mathbf{P}_x^{U_1}(\xi_t^{U_1} \in \Gamma) = \mathbf{P}_x^{U_2}(\xi_t^{U_2} \in \Gamma) \tag{2.2}$$

for all $x \in I \times E'$, $\Gamma \subset I \times E'$, $\Gamma \in \mathscr{E}$, $t \in T$.

e) The domains $\mathbf{D}(A^U)$ of the weak infinitesimal Operators A^U of the processes Ξ^U stopped at τ^U coincide for all $U \in \mathscr{U}$. (We denote this common domain by $\mathbf{D}$.)

Let there be given a controlled Markov process $\{\Xi^U; U \in \mathscr{U}\}$ such that, for all $x \in E$, $U \in \mathscr{U}$ we have $\mathbf{E}_x^U < K$ for some $K - 0$ ($\mathbf{E}_x^U$ denotes the expectation w.r. to the measure $\mathbf{P}_x^U$), and non-negative functions $p \in \mathbf{B}(E \setminus G, \mathscr{E} \cap (E \setminus G))$, $q \in \mathbf{B}(G \times D, (\mathscr{E} \cap G) \otimes \mathscr{D})$ (we shall also write $q^U(x) := q(x, U(x))$). For all $x \in E$ we define the performance functionals on $\mathscr{U}$ by

$$J_x(U) := \mathbf{E}_x^U \left\{ p(\xi_{\tau^U}^U) + \int_0^{\tau^U} q(\xi_t^U, U(\xi_t^U))\, dt \right\} \tag{2.3}$$

Our aim is to find a strategy $U^* \in \mathscr{U}$, such that

$$J_x(U^*) \leqq J_x(U)$$

for all $x \in E$, $U \in \mathscr{U}$. A strategy with this property is said to be optimal.

Remarks. For the definitions and propositions cited here cf. DYNKIN [5]. We call $\Xi=(\xi, \zeta, \mathcal{M}_t, \mathbf{P}_x)$ a homogeneous Markov process on the state space $(E, \mathscr{C}, \mathscr{E})$, if:

a) ζ is a function defined on a sample space Ω with values in $T\cup\{+\infty\}$.
b) ξ is a partial map from $T\times\Omega$ into $(E, \mathscr{C}, \mathscr{E})$, and $\xi(t, \omega)=\xi_t(\omega)$ is defined for all $\omega\in\Omega$, $t\in T\cap(0, \zeta(\omega))$. For a fixed $\omega_0\in\Omega$ the function $\xi_t(\omega_0)$ is called a trajectory of process Ξ.
c) $\mathcal{M}_t$ is a σ-field on $\Omega_t:=\{\omega:\zeta(\omega)>t\}$ $(t\in T)$.
d) $\mathbf{P}_x$ is for all $x\in E$ a functional defined on a σ-field $\mathcal{M}$ of subsets of Ω with $\mathcal{M}\supset\bigcup_{t\in T}\mathcal{M}_t$.

And for these elements the following conditions are satisfied:

(1) If $s\leqq t$ and $A\in\mathcal{M}_s$ then $A\cap\{\omega:\zeta(\omega)>t\}\in\mathcal{M}_t$.
(2) $\{\xi_t\in\Gamma\}:=\{\omega:\xi_t(\omega)\in\Gamma\}\in\mathcal{M}_t$ $(t\in T, \Gamma\in\mathscr{E})$.
(3) $\mathbf{P}_x$ is a probability measure on $\mathcal{M}$ for all $x\in E$.
(4) For all $\Gamma\in\mathscr{E}$, $t\in T$,

$$P(t, x, \Gamma):=\mathbf{P}_x(\xi_t\in\Gamma)$$

is an $\mathscr{E}$-measurable function of x. P is called the transition function of the process Ξ.
(5) $P(0, x, E\setminus\{x\})=0$ for all $x\in E$.
(6) For all $s, t\in T$, $\Gamma\in\mathscr{E}$ we have

$$\mathbf{P}_x(\xi_{t+s}\in\Gamma|\mathcal{M}_t) = P(s, \xi_t, \Gamma)$$

$\mathbf{P}_x$ a.e. on Ω_t.
(7) For all $t\in T$, $\omega\in\Omega_t$ there exists an $\omega_t\in\Omega$ such that

$$\zeta(\omega_t) = \zeta(\omega)-t \quad\text{and}\quad \xi_s(\omega_t) = \xi_{s+t}(\omega) \quad\text{for}\quad 0\leqq s<\zeta(\omega)-t.$$

A Markov process is said to be right-continuous if all of its trajectories are right-continuous. Every right-continuous Markov process is measurable, that is,

$$\{(s, \omega): s\leqq t,\ \omega\in\Omega_t,\ \xi_s(\omega)\in\Gamma\}\in(\mathscr{T}\cap[0, t])\otimes\mathcal{M}_t \tag{2.4}$$

for all $t\in T$, $\Gamma\in\mathscr{E}$, where $\mathscr{T}\cap[0, t]$ denotes the restriction of $\mathscr{T}$ to interval $[0, t]$.

A mapping τ of Ω into $T\cup\{+\infty\}$ is called a Markov time for $\Xi=(\xi, \zeta, \mathcal{M}_t, \mathbf{P}_x)$ if

$$\tau(\omega)\leqq\zeta(\omega)$$

and for all $t\in T$

$$\{\omega:\tau(\omega)>t\}\in\mathcal{M}_t\cap\mathcal{N}$$

Where $\mathscr{N}$ denotes the σ-field generated by all sets of the form $\{\omega:\xi_t(\omega)\in\Gamma\}$ for $\Gamma\in\mathscr{E}$, $t\in T$. We set $A\in\mathscr{M}_\tau$, if $A\subset\Omega_\tau:=\{\omega:\tau(\omega)<+\infty\}$ and for any $t\in T$

$$A\cap\{\omega:\tau(\omega)\leqq t\}\in\mathscr{M}_t$$

By this definition $\mathscr{M}_\tau$ is a σ-field on Ω_τ. ([15]. 3.16.)

A measurable Markov process is said to be strongly Markovian if for an arbitrary Markov time τ and for all $t\in T$, $x\in E$, $\Gamma\in\mathscr{E}$

$$\mathbf{P}_x(\xi_{\tau+t}\in\Gamma|\mathscr{M}_\tau) = P(t,\xi_\tau,\Gamma) \tag{2.5}$$

holds $\mathbf{P}_x$ a.e. on Ω_τ ([5]. 3.18).

The function $\tau=\tau(\omega)$ defined by

$$\tau(\omega):=\sup\{t:\{\xi_s(\omega):s\leqq t\}\subset G\}$$

is called the first exit time from the set G ([5], 4.1). Conditions for the first exit time to be Markov are given by DYNKIN in [5], Chapter 4.

Let $\Xi=(\xi,\zeta,\mathscr{M}_t,\mathbf{P}_x)$ be a right-continuous strong Markov process and τ the Markov time of the first exit from the set G. If we set

$$\tilde{\zeta}(\omega):=\begin{cases}\zeta(\omega) & \text{for}\quad \tau(\omega)=\zeta(\omega)\\ +\infty & \text{for}\quad \tau(\omega)<\zeta(\omega)\end{cases}$$

$$\tilde{\xi}_t(\omega):=\xi_{\min[t,\tau(\omega)]}(\omega)\qquad (0\leqq t<\tilde{\zeta}(\omega))$$

$$\tilde{\mathscr{M}}_t:=\{A\in\mathscr{M}:A\subset\{\tilde{\zeta}>t\}\quad\text{and}\quad A\cap\{\tau>t\}\in\mathscr{M}_t\}$$

then the process $\tilde{\Xi}=(\tilde{\xi},\tilde{\zeta},\tilde{\mathscr{M}}_t,\mathbf{P}_x)$ is also a right-continuous strong Markov process, and is said to arise of Ξ by stopping it at τ ([5], 10.4).

A sequence of functions $f_n\in\mathbf{B}(E,\mathscr{E})$ is said to tend weakly to $f\in\mathbf{B}(E,\mathscr{E})$ if for every signed measure φ of bounded variation, defined on the σ-field $\mathscr{E}$

$$\int f_n d\varphi\to\int f d\varphi$$

holds as $n\to\infty$. In $\mathbf{B}(E,\mathscr{E})$ the weak convergence of f_n to f is equivalent to

(i) $f_n(x)\to f(x)$ for all $x\in E$ as $n\to\infty$, and
(ii) there exists a $K>0$ such that $\|f_n\|\leqq K$ for all $n\in N$.

With every Markov process $\Xi=(\xi,\zeta,\mathscr{M}_t,\mathbf{P}_x)$ we can associate a semigroup of linear operators $\{T_t\}_{t\in T}$ on the Banach space $\mathbf{B}(E,\mathscr{E})$ defined by

$$T_t f(x):=\mathbf{E}_x f(\xi_t) \tag{2.6}$$

where $\mathbf{E}_x$ means the expectation w.r. to the measure $\mathbf{P}_x$. A function $f\in\mathbf{B}(E,\mathscr{E})$ is said to be weakly continuous (w.r. to the semigroup T_t) if

$$w\text{-}\lim_{t\downarrow 0} T_t f=f$$

If T is the set of all non-negative reals, the weak infinitesimal operator of $\{T_t\}_{t\in T}$ is defined by

$$Af := w\text{-}\lim_{t\downarrow 0} \frac{T_t f - f}{t}$$

for all $f\in \mathbf{B}(E, \mathscr{E})$ such that the right-hand side tends weakly to a weakly continuous function. If T equals the set of all non-negative integers, we define A for all $f\in \mathbf{B}(E, \mathscr{E})$ by

$$Af := T_1 f - f$$

and call it infinitesimal operator of T_t as well.

For every controlled Markov process $\{(\xi^U, \zeta^U, \mathscr{M}_t^U, \mathbf{P}_x^U); U\in\mathscr{U}\}$, we can construct by tranformation of the sample space a controlled Markov process $\{\hat\xi, \hat\zeta, \hat{\mathscr{M}}_t, \mathbf{P}_x^U); U\in\mathscr{U}\}$ such that for every $U\in\mathscr{U}$ the processes $\hat\Xi^U$ and Ξ^U are equivalent (Dynkin [4]). (E.g. we set $\hat\Xi^U$ for the canonical process of Ξ^U, cf. [4] Lemma 2. 3). By this we may omit index U of ξ, ζ, $\mathscr{M}_t$ without loss of generality. The U-independence of the exit time τ is a consequence of its definition and the fact that ξ does not depend on U.

3. A necessary and sufficient condition of optimality

Theorem. *Let G, $\{\Xi^U; U\in\mathscr{U}\}$, q and p given as in the second section and let q^U be weakly continuous w.r. to T_t^U $(U\in\mathscr{U})$. Then there exists an optimal strategy $U^*\in\mathscr{U}$ if and only if the boundary value problem*

$$\min_{U\in\mathscr{U}} (A^U f + q^U)(x) = 0 \qquad \text{for} \quad x\in G, \tag{3. 1}$$

$$f(x) = p(x) \quad \text{for} \quad x\in E\setminus G \tag{3. 2}$$

possesses a solution $f\in\mathbf{B}(E, \mathscr{E})$. In this case the minimum occurs in (3. 1) *for* U^*.

Remarks. (3. 1) means in detail that for any element x of G

$$A^{U^*} f(x) + q^{U^*}(x) = 0, \tag{3. 3}$$

$$A^U f(x) + q^U(x) \geqq 0 \quad (U\in\mathscr{U}). \tag{3. 4}$$

A^U denotes the weak infinitesimal operator of the processes arising from Ξ^U by stopping at the first exit from G.

If q^U is continuous then right-continuity of Ξ^U implies weak continuity of q^U. ([5]. Lemma 2. 2.)

Proof. Sufficiency: If $f^* \in \mathbf{D}(A^U)$, and τ is a Markov time for process Ξ^U, then by [5], Theorem 5. 1 we have

$$f^*(x) = \mathbf{E}_x^U f^*(\xi_\tau) - \mathbf{E}_x^U \int_0^\tau A^U f^*(\xi_t)\,dt \tag{3. 5}$$

Let f^* be the required solution of the boundary value problem (3. 1)—(3. 2). Since $f^* \in \mathbf{D} = \mathbf{D}(A^U)$ $(U \in \mathscr{U})$, and the first exit time τ from G is a Markov time for all Ξ^U $(U \in \mathscr{U})$, (3. 5) holds for every $U \in \mathscr{U}$. Let us observe that the right continuity of the processes Ξ^U and the fact that G is open imply $\xi_\tau \in E \setminus G$. By virtue of this and of equations (3. 5), (3. 2), (3. 3) we have

$$f^*(x) = \mathbf{E}_x^{U^*}\Big\{f^*(\xi_\tau) - \int_0^\tau A^{U^*} f^*(\xi_t)\,dt\Big\} = \mathbf{E}_x^{U^*}\Big\{p(\xi_\tau) + \int_0^\tau q^{U^*}(\xi_t)\,dt\Big\}. \tag{3. 6}$$

Analogously but with the aid of (3. 4) instead of (3. 3) we get for all $U \in \mathscr{U}$

$$f^*(x) = \mathbf{E}_x^U f^*(\xi_\tau) - \int_0^\tau A^U f^*(\xi_t)\,dt \leqq \mathbf{E}_x^U \Big\{p(\xi_\tau) + \int_0^\tau q^U(\xi_t)\,dt\Big\}. \tag{3. 7}$$

On account of (2. 3) the relations (3. 6) and (3. 7) imply for arbitrary $U \in \mathscr{U}$, $x \in E$:

$$J_x(U^*) = \mathbf{E}_x^{U^*}\Big\{p(\xi_\tau) + \int_0^\tau q^{U^*}(\xi_t)\,dt\Big\} \leqq \mathbf{E}_x^U\Big\{p(\xi_\tau) + \int_0^\tau q^U(\xi_t)\,dt\Big\} = J_x(U).$$

But the last equation implies that U^* is optimal. Hence the sufficiency of our assumption is proved.

Necessity: Let U^* denote the optimal control strategy, and let us introduce f^* by $f^*(x) := J_x(U^*)$. By the boundedness of p, q and $\mathbf{E}_x^U \tau$ we find that $f^* \in \mathbf{B}(E, \mathscr{E})$.

First we shall prove equation (3. 3) and that $f^* \in \mathbf{D}$. Eq. (2. 3), the definition of the stopped processes and that of f^* show that

$$T_h^{U^*} f^*(x) = \mathbf{E}_x^{U^*} \chi_{\{\tau > h\}} f^*(\xi_h) + \mathbf{E}_x^{U^*} \chi_{\{\tau \leqq h\}} f^*(\xi_\tau) = \tag{3. 8}$$

$$= \mathbf{E}_x^{U^*} \chi_{\{\tau > h\}} \mathbf{E}_{\xi_h}^{U^*} p(\xi_\tau) + \mathbf{E}_x^{U^*} \chi_{\{\tau > h\}} \mathbf{E}_{\xi_h}^{U^*} \int_0^\tau q^{U^*}(\xi_t)\,dt + \mathbf{E}_x^{U^*} \chi_{\{\tau \leqq h\}} p(\xi_\tau)$$

(χ_A denotes the characteristic function of the set A). For the first term of the right-hand side of (3. 8) we obtain by applying [5] Theorem (3. 1) that

$$\mathbf{E}_x^{U^*} \chi_{\{\tau > h\}} \mathbf{E}_{\xi_h}^{U^*} p(\xi_\tau) \doteq \mathbf{E}_x^{U^*} \chi_{\{\tau > h\}} \mathbf{E}_x^{U^*} (\theta_h p(\xi_\tau) | \mathscr{M}_h)$$

where the operator θ_t is defined by

$$\theta_t \eta(\omega) := \eta(\omega_t)$$

(η is any random variable; cf. (7) in the definition of Markov processes.) Taking into account the $\mathcal{M}_h$ measurability of $\chi_{\{\tau>h\}}$, the basic properties of conditional expectations, and the definition of θ_h, we find:

$$(3.9)\qquad \mathbf{E}_x^{U^*}\chi_{\{\tau>h\}}\mathbf{E}_x^{U^*}\big(\theta_h p(\xi_\tau)|\mathcal{M}_h\big) = \mathbf{E}_x^{U^*}\mathbf{E}_x^{U^*}\big(\chi_{\{\tau>h\}}\theta_h p(\xi_\tau)|\mathcal{M}_h\big) =$$

$$= \mathbf{E}_x^{U^*}\chi_{\{\tau>h\}}p\big(\xi_{\tau(\omega_h)}(\omega_h)\big) = \mathbf{E}_x^{U^*}\chi_{\{\tau>h\}}p\big(\xi_{h+\tau(\omega)-h}(\omega)\big) = \mathbf{E}_x^{U^*}\chi_{\{\tau>h\}}p(\xi_\tau).$$

For the second term of (3.8) we find with $q_0 := \chi_G \cdot q^{U*}$ that

$$\mathbf{E}_x^{U^*}\chi_{\{\tau>h\}}\mathbf{E}_{\xi_h}^{U^*}\int_0^\tau q^{U^*}(\xi_t)\,dt = \int_0^\infty \mathbf{E}_x^{U^*}\chi_{\{\tau>h\}}\mathbf{E}_{\xi_h}^{U^*}q_0(\xi_t)\,dt =$$

$$= \int_0^\infty \mathbf{E}_x^{U^*}\mathbf{E}_{\xi_h}^{U^*}q_0(\xi_t)\,dt - \int_0^\infty \mathbf{E}_x^{U^*}\chi_{\{\tau\le h\}}q_0(\xi_\tau)\,dt.$$

The change of the order of integration was allowed by the measurability of the processes (cf. (2.4)) and by Fubini's theorem. Let us observe that $q_0(\xi_\tau)=0$, then by means of definition of T_t^{U*} we obtain:

$$(3.10)\quad \mathbf{E}_x^{U^*}\chi_{\{\tau>h\}}\mathbf{E}_{\xi_h}^{U^*}\int_0^\infty q^{U^*}(\xi_t)\,dt = \int_0^\infty \mathbf{E}_x^{U^*}\mathbf{E}_{\xi_h}^{U^*}q_0(\xi_t)\,dt =$$

$$= \int_0^\infty T_{t+h}^{U^*}q_0(x)\,dt = \int_h^\infty T_t^{U^*}q_0(x)\,dt = \mathbf{E}_x^{U^*}\int_h^\tau q_0^*(\xi_t)\,dt =$$

$$= \mathbf{E}_x^{U^*}\int_0^\tau q^{U^*}(\xi_t)\,dt - \int_0^h T_t^{U^*}q_0(x)\,dt.$$

Substituting (3.9), (3.10) into (3.8) and taking into account of the definition of f^* we get

$$T_h^{U^*}f^*(x) = \mathbf{E}_x^{U^*}\Big\{p(\xi_\tau) + \int_0^\tau q^{U^*}(\xi_t)\,dt\Big\} - \int_0^h T_t^{U^*}q_0(x)\,dt = f^*(x) - \int_0^h T_t^{U^*}q_0(x)\,dt.$$

On account of the definition of A^{U^*} we obtain hence that

$$A^{U^*}f^* = w\text{-}\lim_{h\downarrow 0}\frac{T_h^{U^*}f^* - f^*}{h} = w\text{-}\lim \frac{1}{h}\int_0^h T_t^{U^*}q_0\,dt = -q_0$$

Since q_0 is weakly continuous we get that $f^* \in \mathbf{D}$ so (3.3) is proved.

Now we prove inequality (3.4) indirectly. Let us assume, there exist a strategy $U_0 \in \mathcal{U}$ and state $x_0 \in E$ such that

$$A^{U_0}f^*(x_0) + q^{U_0}(x_0) < 0.$$

From the weak continuity (w.r. to $T_t^{U_0}$) of q^{U_0} and $A^{U_0}f^*$ it follows the existence of a $t_0>0$ such that for all $0\leqq t<t_0$ we have

$$T_t^{U_0}(A^{U_0}f^*+q^{U_0})(x_0)<0.$$

But with the notation $\tau_0(\omega):=\min[t_0,\tau(\omega)]$ we obtain the inequality

$$\mathbf{E}_{x_0}^{U_0}\int_0^{\tau_0}(A^{U_0}f^*+q^{U_0})(\xi_t)\,dt<0. \tag{3.11}$$

Let us denote by U_1 the control strategy

$$U_1(s,x'):=\begin{cases}U^*(s,x') & \text{for} \quad s\geqq s_0+t_0, \quad x'\in E',\\ U_0(s,x') & \text{for} \quad s<s_0+t_0, \quad x'\in E',\end{cases}$$

where s_0 denotes the time-component of x_0, more precisely $x_0=(s_0,x_0')$ with some $x_0'\in E'$. Since $J(U^*)=f^*\in\mathbf{D}$ and τ_0 is a Markov time, we may apply [5] Theorem 5. 1 (cf. also Eq. (3. 5)) to Ξ^{U_1} and we get

$$J_{x_0}(U^*)=f^*(x_0)=\mathbf{E}_{x_0}^{U_1}\Big\{f^*(\xi_{\tau_0})+\int_0^{\tau_0}A^{U_1}f^*(\xi_t)\,dt\Big\} \tag{3.12}$$

In virtue of the definition of $J(U_1)$ we obtain:

$$J_{x_0}(U_1)=\mathbf{E}_{x_0}^{U_1}\Big\{p(\xi_\tau)+\int_0^{\tau}q^{U_1}(\xi_t)\,dt\Big\}= \tag{3.13}$$

$$=\mathbf{E}_{x_0}^{U_1}\Big\{p(\xi_\tau)+\int_{\tau_0}^{\tau}q^{U_1}(\xi_t)\,dt\Big\}+\mathbf{E}_{x_0}^{U_1}\Big\{\int_0^{\tau_0}q^{U_1}(\xi_t)\,dt\Big\}.$$

For the first term of (3. 13) it holds the decomposition

$$\mathbf{E}_{x_0}^{U_1}\Big\{p(\xi_\tau)+\int_{\tau_0}^{\tau}q^{U_1}(\xi_t)\,dt\Big\}=$$

$$=\mathbf{E}_{x_0}^{U_1}\chi_{\{\tau_0=\tau\}}\,p(\xi_\tau)+\mathbf{E}_{x_0}^{U_1}\chi_{\{\tau>\tau_0\}}\,p(\xi_\tau)+\mathbf{E}_{x_0}^{U_1}\chi_{\{\tau>\tau_0\}}\int_{t_0}^{\tau}q(\xi_t)\,dt$$

The second and the third term of the last equation can be transformed analogously to formulae (3. 9), (3. 10) and we obtain

$$\mathbf{E}_{x_0}^{U_1}\Big\{p(\xi_\tau)+\int_{\tau_0}^{\tau}q^{U_1}(\xi_t)\,dt\Big\}=$$

$$=\mathbf{E}_{x_0}^{U_1}\chi_{\{\tau_0=\tau\}}\,p(\xi_\tau)+\mathbf{E}_{x_0}^{U_1}\chi_{\{\tau>\tau_0\}}\,\mathbf{E}_{\xi_{\tau_0}}^{U_1}p(\xi_\tau)+\mathbf{E}_{x_0}^{U_1}\chi_{\{\tau>\tau_0\}}\,\mathbf{E}_{\xi_{\tau_0}}^{U_1}\int_0^{\tau}q^{U_1}(\xi_t)\,dt.$$

By the properties (2. 1), (2. 2) and the definition of U_1 this equals

$$\mathbf{E}_{x_0}^{U_1}\chi_{\{\tau_0=\tau\}}p(\xi_\tau)+\mathbf{E}_{x_0}^{U_1}\chi_{\{\tau_0<\tau\}}\mathbf{E}_{\xi_{\tau_0}}^{U_1}\left\{p(\xi_\tau)+\int_0^\tau q^{U_1}(\xi_t)\,dt\right\}=$$

$$=\mathbf{E}_{x_0}^{U_1}\left[\chi_{\{\tau_0=\tau\}}p(\xi_\tau)+\chi_{\{\tau_0<\tau\}}\mathbf{E}_{\xi_{\tau_0}}^{U^*}\left\{p(\xi_\tau)+\int_0^\tau q^{U^*}(\xi_t)dt\right\}\right]=\mathbf{E}_{x_0}^{U_1}f^*(\xi_{\tau_0}).$$

In this way for (3. 13) we get that

$$J_{x_0}(U_1)=\mathbf{E}_{x_0}^{U_1}\left\{f^*(\xi_{\tau_0})+\int_0^\tau q^{U_1}(\xi_t)dt\right\}. \tag{3. 13'}$$

By subtraction (3. 12) from (3. 13′), and by the definition of U_1 we obtain

$$J_{x_0}(U_1)-J_{x_0}(U^*)=\mathbf{E}_{x_0}^{U_1}\{f^*(\xi_{\tau_0})-f^*(\xi_{\tau_0})\}+\mathbf{E}_{x_0}^{U_1}\int_0^{\tau_0}(A^{U_1}f^*+q^{U_1})(\xi_t)\,dt=$$

$$=\mathbf{E}_{x_0}^{U_0}\int_0^{\tau_0}(A^{U_0}f^*+q^{U_0})(\xi_t)\,dt.$$

But by (3. 11) this means

$$J_{x_0}(U_1)<J_{x_0}(U^*).$$

in contradiction to the supposed optimality of U^*.

The necessity of (3. 2) follows from the fact that $\mathbf{P}_x^U(\tau=0)=1$ for all $U\in\mathscr{U}$, $x\in E\setminus G$. By this our theorem is proved.

Remarks. As we see from the proof, the solution f^* of (3. 1)—(3. 2) has the meaning of expected loss in the state x, if we apply the optimal strategy:

$$f^*(x)=J_x(U^*).$$

From the proof one can see, that we may formulate our theorem in any other topology. E.g. in the strong topology we find: Let q^U be strong continuous (w.r. to T_t^U). Then there exists an optimal control strategy $U^*\in\mathscr{U}$ if and only if the boundary value problem (3. 1)—(3. 2) possesses a solution $f^*\in\mathbf{B}(E,\mathscr{E})$. In this case A^U means the strong infinitesimal operator of P^U.

4. Applications of the main theorem

In this part of our paper let us assume that $\mathscr{U}$ contains all constant strategies $U(x)\equiv d$ $(d\in D)$.

We denote by $\mathbf{D}_L$ the set of all functions $f\in\mathbf{D}$ such that for any x the equality $U_1(x)=U_2(x)$ implies

$$A^{U_1}f(x)=A^{U_2}f(x) \tag{4. 1}$$

i.e. $\mathbf{D}_L:=\{f\in\mathbf{D}: U_1(x)=U_2(x)\Rightarrow A^{U_1}f(x)=A^{U_2}f(x)\}$.

Corollary 1. *Given a controlled Markov process* $\{\Xi^U; U\in\mathcal{U}\}$ *with* $\mathbf{D}=\mathbf{D}_L$, *if* q^U *is weakly continuous w.r. to* T_t^U *then for the existence of an optimal control strategy* $U^*\in\mathcal{U}$ *is necessary and sufficient that there exist an* $f^*\in\mathbf{B}(E,\mathcal{E})$ *and a* $U'\in\mathcal{U}$, *which satisfy the relations*

$$A^{U'(x)}f^*(x)+q(x,U'(x)) = 0 \quad \text{for} \quad x\in G, \tag{4.2}$$

$$A^d f^*(x)+q(x,d) \geqq 0 \quad \text{for} \quad d\in D,\ x\in G, \tag{4.3}$$

$$f^*(x)=p(x) \quad \text{for} \quad x\in E\setminus G \tag{4.4}$$

(A^d *denotes the infinitesimal operator of* Ξ^U *for which* $U\equiv d$).

Proof. Necessity: Let $U^*\in\mathcal{U}$ be optimal. Then (4. 2) follows with $U':=U^*$ from (3. 3), if we observe, that by (4. 1)

$$A^U f(x) = A^{U(x)}f(x) \tag{4.5}$$

holds for all $U\in\mathcal{U}, f\in\mathbf{D}=\mathbf{D}_L$.

Let us suppose the existence of an $x_0\in E$, $d_0\in D$ such that

$$A^{d_0}f^*(x)+q^{d_0}(x_0) < A^{U'}f^*(x_0)+q^{U'}(x_0) = 0$$

where $f^*(x)=J_x(U')$. But since $U\equiv d_0$ is an admissible strategy, according to the theorem the last inequality contradicts the optimality of U'.

Sufficiency: For an arbitrary $x\in E$ there exists a $U'(x)\in D$ such that

$$A^d f^*(x)+q(x,d) \geqq A^{U'(x)}f^*(x)+q(x,U'(x)) = 0$$

Put $U^*(x):=U'(x)\in\mathcal{U}$. By (4. 5) the equation (3. 3) holds true for U^*. Let $U\in\mathcal{U}$ arbitrary, then (3. 4) follows from (4. 3), and Theorem implies the statement of Corollary 1.

Corollary 2. *Let* $U^*\in\mathcal{U}$, $f^*:=J(U^*)\in\mathbf{D}_L$ *and let* $T_t^U q^U$ *be weakly continuous. Then* U^* *is optimal if and only if for* U^* *and* f^* *relations* (4. 2)—(4. 4) *hold true (with* $U^*\equiv U'$*).*

Proof. Analogous to the proof of Corollary 1.

Remarks: We can restate Corollary 1 in the following form: There exists an optimal control strategy if and only if there exists an $f^*\in\mathbf{B}(E,\mathcal{E})$ such that for all $x\in G$ we have

$$\min_{d\in D}[A^d f^*(x)+q(x,d)] = (A^{U^*(x)}f^*)(x)+q(x,U^*(x)) = 0 \tag{4.6}$$

and for all $x\in E\setminus G$ we have $f(x) = p(x)$, furthermore the strategy U^* for which the minimum occurs in (4. 6) belongs to $\mathcal{U}$.

Worth of noting is that this way Corollary 1 states the optimal policy U^* to be independent on the class of admissible control strategies. This implies that exactly those classes have optimal policies which contain the strategy U^* obtained by minimization of equation (4. 6) (assuming that such a minimization exists).

From practical point of view the main advantage of Corollary 1 over Theorem is that the minimization in it has to be carried out only over the control region, the cardinality of which is generally smaller than that of the class of the admissible strategies.

Corollary 2 is useful if we want to decide about a given strategy U^* (which has been determined earlier e.g. by the application of a necessary condition of optimality) whether it is optimal or not. In this case we have to prove the truth of (4. 1) only for $f^*=J(U^*)$, and have the advantage of minimizing over the control region.

Finally, for the illustration of the theorems above, we give two examples which show how to derive the results of Bellman and Fleming from those of this paper.

Discrete-time dynamic programming:

Let the time be discrete. $T=N$, let the sets D, E' be finite and denote by $\mathcal{U}$ the set of all functions from $T\times E'$ in D. Furthermore for all $U\in\mathcal{U}$ let Ξ^U be a Markov chain on the state space $E= T\times E'$ with

$$P^d(i,(k,x'),\{i\}\times E) = \chi_{\{k+i\}}(j)$$

for all $k, i, j\in T$, $x'\in E'$, $d\in D$ (cf. (1. 1)). For G we choose a bounded subset of $T\times E'$. From Corollary 1 it follows that a control U^* is optimal iff there is an f^* such that

$$(5.5)\quad f^*(k,x') = \min_{U(k,x')\in D} [E^{U(k,x')}_{(k,x')} f^*(k+1,\xi'_{k+1})+q(k,x',U(k,x')) \quad \text{for} \quad (k,x')\in G,$$

$$(5.6)\qquad f^*(k,x') = p(k,x') \qquad \text{for} \quad (k,x')\in E\setminus G$$

In (5. 5)—(5. 6) we recognize the well-known recursive equations of the stochastic dynamic programming [1].

Controlled diffusions (cf. FLEMING [6], [7], MANDL [12]).

Let $T=R^+$, $E'\subset R^{n-1}$, $D\subset R^m$ and let $\mathcal{U}$ the class of all measurable bounded functions from E into D. Let the functions b', σ', p, q be continuously differentiable and bounded together with their first order partial derivatives, and let $a':=\frac{1}{2}\sigma'\sigma'^{\mathrm{T}}$ (where σ'^{T} is the matrix transposed of σ') such that the eigenvalues of $a'(x)$ are bounded from below by some $c>0$ for all $x'\in E'$. We introduce the notations

$$b = \begin{pmatrix} 1 \\ b' \end{pmatrix}, \quad \sigma = \begin{pmatrix} 0 & 0 \\ 0 & \sigma' \end{pmatrix}, \quad a = \tfrac{1}{2}\sigma\sigma^T, \quad w_t = \begin{pmatrix} 0 \\ w'_t \end{pmatrix};$$

where w'_t is a Brownian motion process on E'. Then the Ito integral equation

$$\xi_t = x + \int_0^t b(\xi_t, U(\xi_t))\,dt + \int_0^t \sigma(\xi_t)\,dw_t$$

determines a Markov process for all $x \in E$, $U \in \mathcal{U}$ with the infinitesimal generator:

$$A^U f(x) = \sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial^2 f(x)}{\partial x_i \partial x_j} + \langle a(x, U(x));\, f_x(x)\rangle$$

where f_x denotes the gradient of f and $\langle ; \rangle$ the inner product in R^n. From this we see that $\mathbf{D} = \mathbf{D}_L$ holds, and by Corollary 1 it follows:

Necessary and sufficient for the optimality of U^* is the existence of an $f^* \in \mathbf{B}(E, \mathscr{E})$ satisfying

$$\sum_{i,j=1}^{n} a_{ij}(x) \frac{\partial f^*}{\partial x_i \partial x_j} + \min_{d \in D} [\langle a(x,d);\, f_x^*(x) + q(x,d)] \quad \text{for} \quad x \in G$$

$$f(x) = p(x) \quad \text{for} \quad x \in E \setminus G.$$

References

[1] R. Bellman, *Dynamic Programming* (Princeton, 1957).

[2] В. Г. Болтянский, Достаточные условия оптимальности и обоснование метода динамического программирования, *Изв. АН. СССР., сер. матем.*, **28** (1964), 481—514.

[3] S. E. Dreyfus, Introduction to stochastic optimization and control, *Stochastic Optimization and Control,* Proc. Adv. Seminar, University of Wiscosin, 1967 (New York, 1968).

[4] E. B. Dynkin, *Die Grundlagen der Theorie der Markoffschen Processe* (Berlin, 1961).

[5] E. B. Dynkin, *Markov Processes* (Berlin, 1965).

[6] W. H. Fleming, Duality and a priori estimates in Markovian optimization problems, *J. Math. Anal. Appl.*, **16** (1961), 254—279. Erratum *ibid.*, **19** (1966), 204.

[7] W. H. Fleming, Optimal continuous-parameter stochastic control, *SIAM Review*, **11** (1969), 470—509.

[8] Б. И. Григелионис—А. Н. Ширяев, Об управяемых марковских процессах и задаче Стефана, *Пробл. перед. Инф.*, **4** (1968), 60—72.

[9] Н. В. Крылов, Об одном классе нелинейных уравнений в пространстве измеримых функций, *Матем. сб.*, **80 (122)** (1969), 253—265.

[10] Н. В. Крылов, Управление марковскими процессами и пространства *W*, *Изв. АН. СССР, сер. матеw.*, **35** (1971), 224—255.

[11] H. J. Kushner, *Stochastic Stability and Control* (New York, 1967).

[12] P. Mandl, *Analytical Treatment of One-dimensional Markov Processes* (Praha, Berlin, 1968).

[13] V. Nollau and D. Vermes, Zur optimalen Steuerung inhomogener Markoffscher Prozesse *Math. Operationsforsch. Statist.* (to appear).

[14] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes* (New York, 1962).

[15] W. M. Wonham, Random differential equations in control theory, *Probabilistic Methods ni Applied Mathematics,* Vol. II (New York, 1970).

(Received September 28, 1972)