ACTA UNIVERSITATIS SZEGEDIENSIS

ACTA SCIENTIARUM MATHEMATICARUM

ADIUVANTIBUS

L. KALMÁR, L. RÉDEI ET K. TANDORI

REDIGIT

B. SZ.-NAGY

TOMUS XXIX

FASC. 1—2

SZEGED, 1968

INSTITUTUM BOLYAIANUM UNIVERSITATIS SZEGEDIENSIS

A JÓZSEF ATTILA TUDOMÁNYEGYETEM KÖZLEMÉNYEI

ACTA SCIENTIARUM MATHEMATICARUM

KALMÁR LÁSZLÓ, RÉDEI LÁSZLÓ ÉS TANDORI KÁROLY
KÖZREMŰKÖDÉSÉVEL
SZERKESZTI

SZŐKEFALVI-NAGY BÉLA

29. KÖTET
1.—2. FÜZET

SZEGED, 1968. JÚNIUS

JÓZSEF ATTILA TUDOMÁNYEGYETEM BOLYAI-INTÉZETE

INDEX — TARTALOM

TOMUS XXIX — 1968 — 29. KÖTET

BIBLIOGRAPHIE

Commutants de certains opérateurs

Par BÉLA SZ.-NAGY à Szeged et CIPRIAN FOIAŞ à Bucarest

1. Introduction et préliminaires

1. Soit T une contraction d'un espace de Hilbert, de classe C_{00}, c'est-à-dire telle que T^n et T^{*n} convergent fortement vers O lorsque $n \to \infty$. Supposons de plus que T a ses indices de défaut égaux à 1, c'est-à-dire que $I-T^*T$ et $I-TT^*$ sont de rang 1.[1])

Dans des conditions équivalentes à celles énumérées, D. SARASON [2] a démontré que *les opérateurs qui permutent à T sont ceux qui peuvent être représentés sous la forme $X=b(T)$ où b est une fonction analytique bornée dans le disque unité ouvert D du plan des nombres complexes* et que de plus *b peut être choisie de façon que le supremum de sa valeur absolue dans D soit égal à la norme de X*. Un résultat analogue porte sur les commutants d'une somme orthogonale de répliques de la même contraction T, mais au lieu de la fonction scalaire b on aura alors une fonction matricielle. SARASON indique des applications intéressantes de ses résultats à certains problèmes d'interpolation.

L'un des buts de la présente Note est de généraliser ces résultats aux contractions de classe C_{00}, d'ailleurs arbitraires, et même à des couples T, T' de telles contractions, en déterminant alors les opérateurs X tels que $TX=XT'$.

Notre résultat principal, dont les autres dérivent, concerne les opérateurs X qui vérifient l'équation

$$TX=XS \tag{1. 1}$$

où T est une contraction de classe C_{00} et S est une translation unilatérale, de multiplicité donnée ω.

Nous abordons ce problème en représentant T et S par leurs modèles fonctionnels et déterminons alors la forme fonctionnelle correspondante des solutions X de (1. 1).

[1]) Pour une contraction $T \in C_{00}$ les rangs de ces deux opérateurs sont toujours égaux (finis ou infinis).

Tous les espaces de Hilbert que nous allons considérer seront supposés séparables, mais ils peuvent être aussi de dimension finie.

2. Introduisons quelques notions et notations qui nous seront nécessaires.

C désigne le cercle unité et D le disque unité ouvert dans le plan des nombres complexes; on désigne par z le point variable sur C et par λ le point variable dans D; m est la mesure de Lebesgue normée sur C. Pour un espace de Hilbert $\mathfrak{F}$ quelconque, $L^p(\mathfrak{F})$ $(1 \leqq p \leqq \infty)$ est l'espace des fonctions v sur C, à valeurs $v(z) \in \mathfrak{F}$, mesurables et telles que

$$\|v\|_p = \begin{cases} \left[\int \|v(z)\|_{\mathfrak{F}}^p \, dm(z)\right]^{1/p} < \infty \text{ si } 1 \leqq p < \infty, \\ \sup \|v(z)\|_{\mathfrak{F}} < \infty \quad \text{(supremum essentiel)} \quad \text{si } p = \infty; \end{cases}$$

on ne distingue pas deux fonctions comme éléments de $L^p(\mathfrak{F})$ si elles diffèrent seulement dans un ensemble de mesure 0. Les fonctions $v \in L^p(\mathfrak{F})$ pour lesquelles $\int z^n v(z) dm(z) = 0$ $(n = 1, 2, \cdots)$, forment la classe de Hardy $H^p(\mathfrak{F})$; c'est un sous-espace de $L^p(\mathfrak{F})$. Chaque fonction $v \in H^p(\mathfrak{F})$ admet un prolongement naturel analytique $v(\lambda)$ dans D. L'espace $L^2(\mathfrak{F})$ est hilbertien; normes et produits scalaires dans $L^2(\mathfrak{F})$ seront désignés par $\|\cdot\|$ et $(\cdot,\cdot)$ sans ajouter l'indice 2.

$\mathfrak{F}$ et $\mathfrak{G}$ étant deux espaces de Hilbert, on désignera par $L^\infty(\mathfrak{F}, \mathfrak{G})$ l'espace des fonctions V sur C, à valeurs $V(z)$ opérateurs de $\mathfrak{F}$ dans $\mathfrak{G}$, fonctions mesurables (faiblement et alors aussi fortement) et telles que

$$\|V\|_\infty = \sup \|V(z)\| < \infty \quad \text{(supremum essentiel)},$$

$\|V(z)\|$ désignant ici la norme de $V(z)$ comme opérateur de $\mathfrak{F}$ dans $\mathfrak{G}$. On ne distingue pas deux fonctions comme éléments de $L^\infty(\mathfrak{F}, \mathfrak{G})$ si elles coïncident p. p. Les fonctions V pour lesquelles $\|V\|_\infty \leqq 1$, seront appelées contractives. La classe de Hardy correspondante sera désignée par $H^\infty(\mathfrak{F}, \mathfrak{G})$, les fonctions de cette classe admettent des prolongements analytiques $V(\lambda)$ dans D. [2]) Les fonctions $V \in H^\infty(\mathfrak{F}, \mathfrak{G})$ telles que $\|V\|_\infty \leqq 1$ s'appellent analytiques contractives; lorsque de plus $\|V(0)f\|_{\mathfrak{G}} < \|f\|_{\mathfrak{F}}$ pour tout $f \in \mathfrak{F}$ $(f \neq 0)$, V s'appelle analytique contractive *pure*. Chaque fonction $V \in L^\infty(\mathfrak{F}, \mathfrak{G})$ sera considérée aussi comme un opérateur de l'espace $L^2(\mathfrak{F})$ dans l'espace $L^2(\mathfrak{G})$, notamment celui défini par

$$(Vv)(z) = V(z)\, v(z), \quad v \in L^2(\mathfrak{F}).$$

La norme de V comme élément de $L^\infty(\mathfrak{F}, \mathfrak{G})$ est égale alors à sa norme comme opérateur: $\|V\|_\infty = \|V\|$. Lorsque $V \in H^\infty(\mathfrak{F}, \mathfrak{G})$, on a

$$VH^2(\mathfrak{F}) \subset H^2(\mathfrak{G}).$$

[2]) D'après la notation de [A], chap. V, il s'agit donc d'une fonction opératorielle analytique bornée $\{\mathfrak{F}, \mathfrak{G}, V(\lambda)\}$.

La fonction $V \in H^\infty(\mathfrak{F}, \mathfrak{G})$ est appelée *extérieure* lorsque

$$\overline{V H^2(\mathfrak{F})} = H^2(\mathfrak{G}) \quad (\text{adhérence dans } L^2(\mathfrak{G}))$$

et *intérieure* lorsque V comme opérateur de $L^2(\mathfrak{F})$ dans $L^2(\mathfrak{G})$ est isométrique, ou, ce qui revient au même, lorsque $V(z)$ est un opérateur isométrique de $\mathfrak{F}$ dans $\mathfrak{G}$, presque partout (p. p.). V s'appelle $*$*-extérieure* ou $*$*-intérieure* lorsque la fonction associée $V^\sim \in L^\infty(\mathfrak{G}, \mathfrak{F})$, définie par $V^\sim(z) = V(\bar{z})^*$, est extérieure ou intérieure, selon les cas. Pour que V soit *intérieure des deux côtés* (c'est-à-dire à la fois intérieure et $*$-intérieure) il faut et il suffit que l'opérateur V soit unitaire de $L^2(\mathfrak{F})$ dans $L^2(\mathfrak{G})$, ou, d'une manière équivalente, que les valeurs $V(z)$ soient des opérateurs unitaires de $\mathfrak{F}$ dans $\mathfrak{G}$, p. p. [3])

Lorsque $\mathfrak{F}$ et $\mathfrak{G}$ sont de dimension 1, les espaces $L^p(\mathfrak{F})$, $H^p(\mathfrak{F})$, $L^\infty(\mathfrak{F}, \mathfrak{G})$, $H^\infty(\mathfrak{F}, \mathfrak{G})$ se réduisent d'une manière évidente aux espaces scalaires L^p, H^p, L^∞, H^∞, selon les cas.

3. Cela étant, le modèle fonctionnel d'une contraction $T \in C_{00}$ s'obtient de la manière suivante. [4]) On prend une fonction $\Theta \in H^\infty(\mathfrak{F}, \mathfrak{F})$ contractive pure, intérieure des deux côtés, $\mathfrak{F}$ étant un espace de Hilbert quelconque ($1 \leqq \dim \mathfrak{F} \leqq \aleph_0$). On construit l'espace

$$H = H^2(\mathfrak{F}) \ominus \Theta H^2(\mathfrak{F}) \tag{1.2}$$

qui est un sous-espace de $H^2(\mathfrak{F})$ et par conséquent de $L^2(\mathfrak{F})$ (notons que Θ est un opérateur unitaire dans $L^2(\mathfrak{F})$). On définit dans H l'opérateur T par

$$Tu = P_H(zu) \quad (u \in H), \tag{1.3}$$

où P_H est la projection orthogonale de $L^2(\mathfrak{F})$ sur H. L'opérateur adjoint T^* sera alors donné par

$$T^*u = v \quad \text{où} \quad v(z) = \frac{1}{z}[u(z) - u(0)] \qquad (u \in H). \tag{1.3*}$$

Les opérateurs T ainsi obtenus sont des contractions de classe C_{00}, et toute contraction de classe C_{00} d'un espace de Hilbert ($\neq \{0\}$) s'obtient de cette façon, à équivalence unitaire près (notamment si l'on prend pour Θ la „fonction caractéristique" de la contraction en question).

Pour toute contraction T d'un espace de Hilbert, de classe C_{00} (et plus généralement pour toute contraction complètement non-unitaire) et pour toute fonction $\varphi(\lambda) = \sum_0^\infty c_k \lambda^k \in H^\infty$ on peut définir l'opérateur $\varphi(T)$ par

$$\varphi(T) = \lim_{r \to 1-0} \sum_0^\infty c_k r^k T^k \tag{1.4}$$

[3]) Pour ces notions cf. [A], chap. V.

[4]) Cf. [A], chap. VI.

où la série converge en norme et la limite existe au sens de la convergence forte des opérateurs. [5]) Dans le cas particulier que nous avons en vue, il dérive de (1. 2—3) la représentation

(1. 5) $$\varphi(T)u = P_H(\varphi u) \qquad (u \in H).$$

Remarquons aussi la relation

(1. 6) $$TP_H v = P_H(zv) \quad \text{pour} \quad v \in H^2(\mathfrak{F}).$$

En effet, (1. 3) entraîne

$$TP_H v = P_H(z \cdot P_H v) = P_H(zv) - P_H(z \cdot (I - P_H)v)$$

et il ne reste qu'à montrer que le dernier terme est 0. Or, cela résulte de ce que, en vertu de la définition (1. 2) de H,

$$P_H(z \cdot (I - P_H)v) \in P_H(z \cdot \Theta H^2(\mathfrak{F})) \subset P_H(\Theta H^2(\mathfrak{F})) = \{0\}.$$

D'ailleurs, la multiplication par z dans un espace $L^2(\mathfrak{G})$, ou dans un espace $H^2(\mathfrak{G})$, est le modèle fonctionnel d'une translation bilatérale, ou unilatérale, selon les cas, de multiplicité égale à dim $\mathfrak{G}$. L'opérateur S figurant dans (1. 1) sera donc représenté par son modèle

(1. 7) $$Sw = z \cdot w \quad \text{où} \quad w \in H^2(\mathfrak{G}), \quad \dim \mathfrak{G} = \omega.$$

2. Les théorèmes

Notre résultat principal dans cette Note sera le suivant:

Théorème 1. *Soient la contraction T de classe C_{00} et la translation unilatérale S de multiplicité ω $(1 \leqq \omega \leqq \aleph_0)$ représentées par leurs modèles fonctionnels* (1. 2—3) *et* (1. 7). *Les opérateurs X de $H^2(\mathfrak{G})$ dans H vérifiant l'équation*

(2. 1) $$TX = XS$$

sont alors précisément ceux qui peuvent être représentés sous la forme

(2. 2) $$Xw = P_H Bw, \quad w \in H^2(\mathfrak{G}),$$

moyennant une fonction $B \in H^\infty(\mathfrak{G}, \mathfrak{F})$, de plus cette fonction peut être choisie telle que

(2. 3) $$\|B\|_\infty = \|X\|.$$

La démonstration de ce théorème fera l'objet des n^os^ 4—6. Ici nous voulons en déduire le suivant

[5]) Cf. [A], chap. III.

Théorème 2. *Soient T et T' deux contractions de classe C_{00}, représentées par leurs modèles fonctionnels de type* (1. 2—3), *avec des fonctions $\Theta \in H^\infty(\mathfrak{F}, \mathfrak{F})$ et $\Theta' \in H^\infty(\mathfrak{F}', \mathfrak{F}')$ (contractives pures, intérieures des deux côtés), selon les cas. Les opérateurs Y de H' dans H vérifiant la condition*

$$TY = YT' \tag{2. 4}$$

sont précisément ceux qui peuvent être représentés sous la forme

$$Yu = P_H Bu \qquad (u \in H'), \tag{2. 5}$$

moyennant une fonction $B \in H^\infty(\mathfrak{F}', \mathfrak{F})$ telle que

$$B\Theta' H^2(\mathfrak{F}') \subset \Theta H^2(\mathfrak{F}), \tag{2. 6}$$

de plus cette fonction peut toujours être choisie telle que

$$\|B\|_\infty = \|Y\|. \tag{2. 7}$$

Démonstration. On déduit de (2. 4) et de la relation (1. 6), appliquée à T' au lieu de T, que

$$TYP_{H'}u = YT'P_{H'}u = YP_{H'}(zu) \quad \text{pour} \quad u \in H^2(\mathfrak{F}').$$

Ainsi, en définissant l'opérateur S dans $H^2(\mathfrak{F}')$ et l'opérateur X de $H^2(\mathfrak{F}')$ dans H par

$$Su = zu \quad \text{et} \quad Xu = YP_{H'}u \qquad (u \in H^2(\mathfrak{F}')), \tag{2. 8}$$

on aboutit à l'équation $TX = XS$. En vertu du théorème 1 il existe alors une fonction $B \in H^\infty(\mathfrak{F}', \mathfrak{F})$ telle que

$$Xu = P_H Bu \quad \text{pour} \quad u \in H^2(\mathfrak{F}'), \quad \text{et} \quad \|B\|_\infty = \|X\|; \tag{2. 9}$$

vu que (2. 8) entraîne les relations $X|H' = Y$ et $\|Y\| = \|X\|$, il résulte de (2. 9) que

$$Y = P_H B|H' \quad \text{et} \quad \|Y\| = \|B\|_\infty.$$

Par conséquent notre fonction B vérifie (2. 5) et (2. 7). Elle vérifie aussi (2. 6) parce que de (2. 8—9) il dérive $P_H B(I - P_{H'})u = X(I - P_{H'})u = YP_{H'}(I - P_{H'})u = 0$ pour $u \in H^2(\mathfrak{F}')$; en vertu de la définition (1. 2) de H et celle analogue pour H' on en déduit que $P_H B\Theta' H^2(\mathfrak{F}') = \{0\}$, $B\Theta' H^2(\mathfrak{F}') \subset \Theta H^2(\mathfrak{F})$.

Inversement, une fonction $B' \in H^\infty(\mathfrak{F}', \mathfrak{F})$ quelconque, vérifiant la relation (2. 6) (avec B' au lieu de B) engendre, moyennant la formule $Y'u = P_H B'u$ $(u \in H')$, analogue à (2. 5), un opérateur Y' de H' dans H tel que $TY' = Y'T'$. En effet, on a alors

$$P_H B'(I - P_{H'}) H^2(\mathfrak{F}') \subset P_H B'\Theta' H^2(\mathfrak{F}') \subset P_H \Theta H^2(\mathfrak{F}) = \{0\},$$

d'où, faisant usage aussi de (1. 6), on obtient pour $u \in H'$:

$$TY'u = TP_H B'u = P_H(z \cdot B'u) = P_H B'(zu) =$$
$$= P_H B' P_{H'}(zu) + P_H B'(I - P_{H'})(zu) = P_H B' T'u = Y'T'u.$$

Cela achève la démonstration du théorème 2, comme une conséquence du théorème 1.

Remarque. Lorsque B vérifie (2. 6), la fonction B^+ définie par

$$B^+(z) = \Theta(z)^* B(z) \Theta'(z) \tag{2.10}$$

vérifie $B^+ H^2(\mathfrak{F}') \subset H^2(\mathfrak{F})$. On en déduit que B^+ est analytique, donc

$$B^+ \in H^\infty(\mathfrak{F}', \mathfrak{F}). \tag{2.11}$$

Inversement, (2. 11) entraîne évidemment (2. 6). Ainsi, les deux conditions sont équivalentes.

Un cas particulier où ces conditions sont vérifiées pour toute fonction $B \in H^\infty(\mathfrak{F}', \mathfrak{F})$, est celui où $\Theta'(z) = \varphi(z) I_{\mathfrak{F}'}$, φ étant une fonction intérieure scalaire non constante, et Θ admet φ comme un multiple, c'est-à-dire qu'il existe une fonction $\Omega \in H^\infty(\mathfrak{F}, \mathfrak{F})$ telle que $\Theta(z)\Omega(z) = \varphi(z) I_{\mathfrak{F}}$. C'est le cas tout particulièrement, si $\Theta(z) = \varphi(z) I_{\mathfrak{F}}$. Si l'on a de plus $\mathfrak{F} = \mathfrak{F}'$, c'est-à-dire que $\Theta = \Theta' = \varphi I_{\mathfrak{F}}$, on aboutit ainsi essentiellement au théorème 3 de Sarason [2], et si $\mathfrak{F} = \mathfrak{F}' = E^1$, à son théorème 1. Dans ce dernier cas la fonction B se réduit à une fonction scalaire $b \in H^\infty$ et l'équation (2. 5) prend la forme

$$Yu = P_H(bu) = b(T)u \qquad (u \in H)$$

où l'on a fait usage aussi de (1. 5). Ainsi, $Y = b(T)$, $\|b\|_\infty = \|Y\|$: le théorème de Sarason en sa forme citée au commencement du n° 1.

D'ailleurs, les méthodes adoptées dans [2] sont différentes des nôtres. Un point commun des deux raisonnements est un lemme de factorisation qu'on va établir au n° 3.

3. Un lemme de factorisation

Il s'agit du suivant

Lemme. *Toute fonction $\Omega \in H^\infty(\mathfrak{L}, \mathfrak{N})$ admet une factorisation*

$$\Omega(z) = \Omega_2(z)\, \Omega_1(z) \tag{3.1}$$

en produit de deux fonctions, $\Omega_1 \in H^\infty(\mathfrak{L}, \mathfrak{M})$ et $\Omega_2 \in H^\infty(\mathfrak{M}, \mathfrak{N})$, dont la première est extérieure, et qui sont telles que

$$\Omega_1(z)^* \Omega_1(z) = [\Omega(z)^* \Omega(z)]^{\frac{1}{2}} \tag{3.2}$$

et

$$\Omega_2(z)^* \Omega_2(z) = \Omega_1(z)\, \Omega_1(z)^*. \tag{3.3}$$

Ce résultat, généralisant des résultats antérieurs de HELSON—LOWDENSLAGER et de DEVINATZ, se trouve (même pour des fonctions non bornées, mais de carré sommable) dans SARASON [2] (théorème 4). Nous en allons donner une démonstration simple, basée sur un critère général de factorisation dû à LOWDENSLAGER [1], *cf.* aussi [A], proposition V. 4. 2.

Démonstration. Il ne restreint évidemment pas la généralité de supposer que la fonction Ω est contractive. Posons

$$N(z) = [\Omega(z)^* \Omega(z)]^{\frac{1}{2}} \quad \text{et} \quad N_1(z) = N(z)^{\frac{1}{2}};$$

ce sont des fonctions contractives à valeurs opérateurs autoadjoints positifs dans $\mathfrak{L}$. Comme $N - N^2 = N(I-N) \geqq O$, on a pour $v \in L^2(\mathfrak{L})$:

$$||N_1 v_2|| = (N_1^2 v, v) = (Nv, v) \geqq (N^2 v, v) = (\Omega^* \Omega v, v) = ||\Omega v||^2,$$

d'où il résulte en particulier qu'il existe une contraction

$$Z: \overline{N_1 \cdot H^2(\mathfrak{L})} \to H^2(\mathfrak{N})$$

pour laquelle

$$Z(N_1 z) = \Omega u \qquad (u \in H^2(\mathfrak{L})).$$

Comme les opérateurs N_1 et Ω permutent à la multiplication par des fonctions scalaires, il en est de même pour Z et par conséquent

$$Z \bigcap_{n \geqq 0} z^n \overline{N_1 \cdot H^2(\mathfrak{L})} \subset \bigcap_{n \geqq 0} z^n \overline{\Omega \cdot H^2(\mathfrak{L})} \subset \bigcap_{n \geqq 0} z^n H^2(\mathfrak{N}) = \{0\}.$$

De ce résultat on pourra conclure

(3. 4) $$\bigcap_{n \geqq 0} z^n \overline{N_1 \cdot H^2(\mathfrak{L})} = \{0\}$$

dès qu'on aura montré que le seul élément w de $\overline{N_1 \cdot H^2(\mathfrak{L})}$ pour lequel $Zw=0$, est $w=0$. Or, soit $w = \lim v_n$ où $v_n = N_1 u_n$, $u_n \in H^2(\mathfrak{L})$. Alors, $0 = Zw = \lim Zv_n = \lim \Omega u_n$ entraîne

$$\|Nu_n\| = \|\Omega u_n\| \to 0, \quad Nu_n \to 0, \quad N_1 v_n = N_1 N_1 u_n = Nu_n \to 0.$$

Ainsi, $N_1 w = \lim N_1 v_n = 0$. Comme $w \in \overline{N_1 \cdot H^2(\mathfrak{L})}$, cela entraîne $w=0$.

La relation (3. 4) que nous venons d'établir, entraîne, d'après le critère cité, qu'il existe une fonction $\Omega_1 \in H^\infty(\mathfrak{L}, \mathfrak{M})$, même extérieure, telle que

(3. 5) $$\Omega_1(z)^* \Omega_1(z) = N_1(z)^2;$$

$\mathfrak{M}$ est un certain espace de Hilbert. C'est l'équation (3. 2).

Vu que $N_1(z)^2 = N(z) \geqq N(z)^2 = \Omega(z)^* \Omega(z)$, (3. 5) entraîne

(3. 6) $$\Omega_1(z)^* \Omega_1(z) \geqq \Omega(z)^* \Omega(z).$$

Vu que Ω_1 est une fonction extérieure, il s'ensuit de (3. 6) qu'il existe une fonction $\Omega_2 \in H^\infty(\mathfrak{M}, \mathfrak{N})$ telle que la factorisation (3. 1) subsiste; *cf.* [A], proposition V. 4. 1. En réunissant ces résultats on obtient

$$[\Omega_1(z)^* \Omega_1(z)]^2 = N_1(z)^4 = N(z)^2 = \Omega(z)^* \Omega(z) = (\Omega_2(z)\Omega_1(z))^*(\Omega_2(z)\Omega_1(z))$$

et par conséquent

$$\Omega_1(z)^* \Delta(z) \Omega_1(z) = 0, \quad \text{où} \quad \Delta(z) = \Omega_1(z)\Omega_1(z)^* - \Omega_2(z)^*\Omega_2(z).$$

Puisque $\overline{\Omega_1(z)\mathfrak{L}} = \mathfrak{M}$ (*cf.* [A], proposition V. 2. 4 (b)) et que par conséquent $\Omega_1(z)^*$ est inversible, on conclut que $\Delta(z) = 0$; tout cela p. p. Cela établit la relation (3. 3) et achève la démonstration.

4. Une inégalité

1. Nous commençons la démonstration du théorème 1. Soient donc T et S donnés par leurs modèles fonctionnels (1. 2—3) et (1. 7) où $\mathfrak{F}$ et $\mathfrak{G}$ sont deux espaces de Hilbert ($\neq \{0\}$), $\dim \mathfrak{G} = \omega$, et où Θ est une fonction $\in H^\infty(\mathfrak{F}, \mathfrak{F})$, contractive pure, intérieure des deux côtés.

Il ne restreint pas la généralité de choisir $\mathfrak{G} = E^\omega$, l'espace des vecteurs $x = [x_k]_1^\omega$ [6]) à composantes nombres complexes, avec

$$\|x\|_{E^\omega} = \left[\sum_1^\omega |x_k|^2\right]^{\frac{1}{2}} < \infty.$$

On désignera par $e^{(k)}$ le vecteur dont la composante de rang k est 1 et les autres sont 0.

Auprès de l'espace $H = H^2(\mathfrak{F}) \ominus \Theta H^2(\mathfrak{F})$, dans lequel T est défini, on envisagera aussi le sous-espace suivant de $L^2(\mathfrak{F})$:

(4. 1) $$G = L^2(\mathfrak{F}) \ominus \Theta H^2(\mathfrak{F}) = [L^2(\mathfrak{F}) \ominus H^2(\mathfrak{F})] \oplus H.$$

Pour toute fonction scalaire $\varphi \in H^\infty$ on a $\varphi \Theta H^2(\mathfrak{F}) = \Theta \varphi H^2(\mathfrak{F}) \subset \Theta H^2(\mathfrak{F})$ et par conséquent

(4. 2) $$\bar{\varphi} G \subset G. \text{ [7])}$$

Nous définissons dans $L^2(\mathfrak{F})$ des opérateurs Φ et Ψ de la manière suivante:

(4. 3) $$(\Phi v)(z) = \bar{z} \cdot \Theta(z) v(\bar{z}), \quad (\Psi v)(z) = \bar{z} \cdot \Theta^\sim(z) v(\bar{z}).$$

---

[6]) Dans le cas $\omega = \aleph_0$ convenons d'entendre par $k = 1, \ldots, \omega$ la suite des nombres naturels.

[7]) Pour une fonction scalaire ψ nous définissons les fonctions $\psi^\sim$ et $\bar{\psi}$ par $\psi^\sim(z) = \overline{\psi(\bar{z})}$ et $\bar{\psi}(z) = \overline{\psi(z)}$. On a alors $\bar{\psi}^\sim(z) = \overline{\psi^\sim(z)} = \psi(\bar{z})$. Pour $\psi \in H^\infty$ on a aussi $\psi^\sim \in H^\infty$.

Comme Θ est intérieure des deux côtés, ses valeurs sont des opérateurs unitaires dans $\mathfrak{F}$, p. p., donc on a

(4. 4) $$\|(\Phi v)(z)\|_{\mathfrak{F}} = \|v(\bar{z})\|_{\mathfrak{F}} = \|(\Psi v)(z)\|_{\mathfrak{F}} \quad \text{p.p.}$$

Or il est manifeste que pour toute fonction $a \in L^1$, l'intégrale de $a(\bar{z})$ sur C est égale à celle de $a(z)$, donc il dérive de (4. 4) par l'intégration des carrés que $\|\Phi v\| = \|v\| = \|\Psi v\|$: les opérateurs Φ et Ψ sont donc isométriques. De plus on a $(\Phi\Psi v)(z) = \bar{z}\Theta(z)\cdot z\Theta^{\sim}(\bar{z})\, v(z) = \bar{z}z\Theta(z)\Theta(z)^* v(z) = v(z)$ et une relation analogue pour $\Psi\Phi$, d'où il résulte que Φ et Ψ sont des opérateurs unitaires dans $L^2(\mathfrak{F})$ et que $\Phi = \Psi^{-1}$.

Il est manifeste que si $v(z)$ parcourt les fonctions dans $H^2(\mathfrak{F})$, $\bar{z}v(\bar{z})$ parcourt les fonctions dans $L^2(\mathfrak{F}) \ominus H^2(\mathfrak{F})$ et $\bar{z}\Theta(z)\, v(\bar{z})$ parcourt les fonctions dans $\Theta[L^2(\mathfrak{F}) \ominus H^2(\mathfrak{F})] = L^2(\mathfrak{F}) \ominus \Theta H^2(\mathfrak{F}) = G$. Cela prouve que

(4. 5) $$\Phi H^2(\mathfrak{F}) = G, \quad \Psi G = H^2(\mathfrak{F}).$$

Notons encore que pour toute fonction scalaire $\varphi \in L^\infty$ on a $\varphi(z)\cdot(\Psi v)(z) = \varphi(z)\ \bar{z}\ \Theta(z)\, v(\bar{z}) = \bar{z}\ \Theta(z)\ \bar{\varphi}^{\sim}(\bar{z})\, v(\bar{z}) = \big(\Psi(\bar{\varphi}^{\sim} v)\big)(z)$, donc

(4. 6) $$\varphi \cdot \Psi v = \Psi(\bar{\varphi}^{\sim} v) \qquad (v \in L^2(\mathfrak{F}), \quad \varphi \in L^\infty).$$

2. Soit X un opérateur (de $H^2(E^\omega)$ dans H) vérifiant l'équation $TX = XS$. On a alors aussi $T^n X = XS^n$ pour $n = 0, 1, \cdots$ et par conséquent $\varphi(T)\cdot X = X\cdot\varphi(S)$ pour $\varphi \in H^\infty$. Or, $\varphi(T)$ s'exprime aussi par (1. 5), et on déduit de (1. 7) que $\varphi(S)$ est la multiplication par la fonction φ. Donc nous avons

(4. 7) $$P_H(\varphi \cdot Xu) = X(\varphi u) \quad \text{pour} \quad u \in H^2(E^\omega) \quad \text{et} \quad \varphi \in H^\infty.$$

Désignons

(4. 8) $$a_k = Xe^{(k)} \qquad (k = 1, \ldots, \omega),$$

où on considère $e^{(k)}$ comme la fonction constante ayant cette valeur. (4. 7) entraîne

(4. 9) $$P_H(\varphi a_k) = X(\varphi e^{(k)}).$$

Soit $g \in G$ où G est le sous-espace de $L^2(\mathfrak{F})$ défini par (4. 1) et soit $\varphi \in H^\infty$. En appliquant (4. 9) à $\varphi^{\sim} \in H^\infty$, on obtient

$$(g, X(\varphi^{\sim} e^{(k)})) = (g, P_H(\varphi^{\sim} a_k)) = (P_H g, \varphi^{\sim} a_k) = (g, \varphi^{\sim} a_k),$$

la dernière équation dérivant de ce que, en vertu de la seconde des représentations (4. 1) de G, la projection de g à H est égale à sa projection à $H^2(\mathfrak{F})$. Comme $\overline{\varphi^{\sim}} = \bar{\varphi}^{\sim}$, notre résultat peut être écrit aussi sous la forme

(4. 10) $$(g, X(\varphi^{\sim} e^{(k)})) = (\bar{\varphi}^{\sim} g, a_k).$$

Envisageons alors une somme de la forme

$$\sum_{k=1}^{r}\Big(\sum_{j=1}^{s}\bar{\varphi}_{\tilde{j}\tilde{k}}g_j',\, a_k\Big)$$

où r et s sont des nombres naturels, r ne dépassant pas ω, et où

$$\varphi_{jk}\in H^\infty \quad \text{et} \quad g_j'\in G.$$

En vertu de (4. 10) cette somme est égale à

$$\sum_{j=1}^{s}\Big(g_j',\, X\sum_{k=1}^{r}\varphi_{\tilde{j}\tilde{k}}e^{(k)}\Big)$$

et sa valeur absolue est par conséquent majorée par

$$\|X\|\cdot\sum_{j=1}^{s}\Big[\|g_j'\|\cdot\Big\|\sum_{k=1}^{r}\varphi_{\tilde{j}\tilde{k}}e^{(k)}\Big\|\Big].$$

Ce résultat peut être formulé aussi dans la forme suivante: pour $g_k\in G$ $(k=1,\cdots,r)$ quelconques,

$$\Big|\sum_{k=1}^{r}(g_k, a_k)\Big| \leqq \|X\|\cdot\inf\sum_{j=1}^{s}\Big[\|g_j'\|\cdot\Big\|\sum_{k=1}^{r}\varphi_{\tilde{j}\tilde{k}}e^{(k)}\Big\|\Big] \tag{4.11}$$

où g_j', φ_{jk} varient sous les conditions

$$g_j'\in G,\quad \varphi_{jk}\in H^\infty,\quad \sum_{j=1}^{s}\bar{\varphi}_{\tilde{j}\tilde{k}}g_j' = g_k \qquad (k=1,\cdots,r). \tag{4.12}$$

Soit Ψ l'opérateur unitaire dans $L^2(\mathfrak{F})$ introduit ci-dessus, et posons

$$\alpha_k=\Psi a_k \qquad (k=1,\cdots,\omega). \tag{4.13}$$

Puisque Ψ applique G sur $H^2(\mathfrak{F})$ (*cf.* (4. 5)), on déduit de (4. 11—12) et de (4. 6) que pour $\gamma_k\in H^2(\mathfrak{F})$ $(k=1,\cdots,r)$ quelconques on a

$$\Big|\sum_{k=1}^{r}(\gamma_k, \alpha_k)\Big| \leqq \|X\|\cdot\inf\sum_{j=1}^{s}\Big[\|\gamma_j'\|\cdot\Big\|\sum_{k=1}^{r}\varphi_{\tilde{j}\tilde{k}}e^{(k)}\Big\|\Big] \tag{4.14}$$

où γ_j' et φ_{jk} varient sous les conditions

$$\gamma_j'\in H^2(\mathfrak{F}),\quad \varphi_{jk}\in H^\infty,\quad \sum_{j=1}^{s}\varphi_{jk}\gamma_j' = \gamma_k \qquad (k=1,\cdots,r). \tag{4.15}$$

Envisageons le cas particulier où les fonctions γ_k sont *bornées.* Elles engendrent alors une fonction $\gamma\in H^\infty(E^r, \mathfrak{F})$ par la définition

$$\gamma(z)\,x = \sum_{k=1}^{r}\gamma_k(z)x_k, \quad x=[x_k]_1^r\in E^r. \tag{4.16}$$

D'après le lemme du n°. 3 il existe des fonctions $\Omega_1 \in H^\infty(E^r, \mathfrak{M})$ et $\Omega_2 \in H^\infty(\mathfrak{M}, \mathfrak{F})$, avec un espace de Hilbert intermédiaire $\mathfrak{M}$, telles que Ω_1 est extérieure et les relations suivantes sont vérifiées:

$$\gamma(z) = \Omega_2(z)\Omega_1(z), \tag{4. 17}$$

$$\Omega_1(z)^*\Omega_1(z) = (\gamma(z)^*\gamma(z))^{\frac{1}{2}}, \tag{4. 18}$$

$$\Omega_2(z)^*\Omega_2(z) = \Omega_1(z)\Omega_1(z)^*. \tag{4. 19}$$

Le fait que Ω_1 est extérieure entraîne $\dim \mathfrak{M} \leqq \dim E^r = r$; *cf.* [A], corollaire à la proposition V. 2. 4. Ainsi, on peut supposer $\mathfrak{M} = E^s$, $0 \leqq s \leqq r$. Laissant à part le cas banal $\gamma = 0$ (c'est-à-dire où $\gamma_k = 0$ pour tous les k) on a même $1 \leqq s \leqq r$. Les fonctions

$$\Omega_1 \in H^\infty(E^r, E^s), \quad \Omega_2 \in H^\infty(E^s, \mathfrak{F})$$

peuvent être représentées sous les formes

$$\Omega_1(z)x = \Big[\sum_{k=1}^{r} \varphi_{jk}(z)x_k\Big]_{j=1}^{s} \quad \text{où} \quad x = [x_k]_1^r \in E^r, \tag{4. 20}$$

$$\Omega_2(z)y = \sum_{j=1}^{s} \gamma_j'(z)y_j \quad \text{où} \quad y = [y_j]_1^s \in E^s, \tag{4. 21}$$

moyennant des fonctions $\varphi_{jk} \in H^\infty$ $(j = 1, \cdots, s;\ k = 1, \cdots, r)$ et $\gamma_j' \in H^\infty(\mathfrak{F})$ $(j = 1, \cdots, s)$. En vertu de (4. 17) on en déduit pour la fonction $\gamma \in H^\infty(E^r, \mathfrak{F})$ la représentation

$$\gamma(z)x = \sum_{j=1}^{s} \gamma_j'(z) \sum_{k=1}^{r} \varphi_{jk}(z)x_k \quad \text{où} \quad x \in E^r. \tag{4. 22}$$

En comparant (4. 16) à (4. 22) il résulte que

$$\gamma_k = \sum_{j=1}^{s} \gamma_j' \varphi_{jk} \qquad (k = 1, \cdots, r).$$

Ainsi, les fonctions obtenues γ_j' et φ_{jk} vérifient les relations (4. 15).

Evaluons la double somme correspondante, figurant au second membre de (4. 14). A cette fin, désignons par $e^{(h)}$ les vecteurs de base de E^r, ainsi que les vecteurs de base de E^s ($h = 1, \cdots, r$ et $h = 1, \cdots, s$, selon les cas). En vertu de (4. 21) on a $\Omega_2(z)e^{(j)} = \gamma_j'(z)$; vu aussi (4. 19) et (4.20) on en déduit

$$\|\gamma_j'(z)\|_{\mathfrak{F}}^2 = \|\Omega_2(z)e^{(j)}\|_{\mathfrak{F}}^2 = (\Omega_2(z)^*\Omega_2(z)e^{(j)}, e^{(j)})_{E^s} =$$

$$= (\Omega_1(z)\Omega_1(z)^*e^{(j)}, e^{(j)})_{E^s} = \sum_{k=1}^{r} |\varphi_{jk}(z)|^2.$$

Par intégration sur C il en résulte

$$\|\gamma_j'\|^2 = \sum_{k=1}^{r} \|\varphi_{jk}\|^2.$$

D'autre part, il est manifeste que

$$\|\sum_{k=1}^{r} \varphi_{\tilde{j}k} e^{(k)}\|^2 = \sum_{k=1}^{r} \|\varphi_{\tilde{j}k}\|^2 = \sum_{k=1}^{r} \|\varphi_{jk}\|^2.$$

Faisant usage aussi de (4. 18) on obtient donc

$$\sum_{j=1}^{s} [\|\gamma_j'\| \cdot \|\sum_{k=1}^{r} \varphi_{\tilde{j}k} e^{(k)}\|] = \sum_{j=1}^{s} [\sum_{k=1}^{r} \|\varphi_{jk}\|^2] = \sum_{k=1}^{r} [\sum_{j=1}^{s} \|\varphi_{jk}\|^2] =$$

$$= \sum_{k=1}^{r} \int (\Omega_1(z)^* \Omega_1(z) e^{(k)}, e^{(k)})_{E^r}\, dm(z) = \int \operatorname{tr} [\Omega_1(z)^* \Omega_1(z)]\, dm(z) =$$

$$= \int \operatorname{tr} [\gamma(z)^* \gamma(z)]^{\frac{1}{2}}\, dm(z),$$

„tr" désignant la trace de l'opérateur.

En utilisant la notation $|\tau| = (\tau^*\tau)^{\frac{1}{2}}$ pour un opérateur τ (d'un espace de Hilbert en soi-même ou dans un autre) nous pouvons résumer nos résultats dans *l'inégalité*

$$\left| \sum_{k=1}^{r} (\gamma_k, \alpha_k) \right| \leqq \|X\| \cdot \int \operatorname{tr} |\gamma(z)| \cdot dm(z), \tag{4. 23}$$

valable pour les $\alpha_k = \Psi a_k$ *et pour des* $\gamma_k \in H^\infty(\mathfrak{F})$ *quelconques, en nombre fini* $r \leqq \omega$. [8])

Ajoutons que si l'on complète la suite $[\gamma_k]_1^r$ à une suite $[\gamma_k]_1^\omega$ par des fonctions $\gamma_k = 0$ $(k > r)$ il y correspondra, par la même formule (4. 16), une fonction $\gamma \in H^\infty(E^\omega, \mathfrak{F})$. La trace de $|\gamma(z)|$ ne change pas lors de ce prolongement.

5. Application du théorème de Hahn—Banach

1. L'inégalité que nous venons d'obtenir impose d'introduire l'espace linéaire normé L suivant. Ses éléments sont les fonctions $\gamma \in L^\infty(E^\omega, \mathfrak{F})$ définies par

$$\gamma(z)x = \sum_{1}^{\omega} \gamma_k(z) x_k, \quad x = [x_k]_1^\omega \in E^\omega, \tag{5. 1}$$

où $\gamma_k \in L^\infty(\mathfrak{F})$ $(k = 1, \cdots, \omega)$, $\gamma_k = 0$ pour $k > r$, r étant un nombre fini $\leqq \omega$, dépendant de γ. Les opérations linéaires étant celles évidentes, on définit la norme par

$$\|\gamma\| = \int \operatorname{tr} |\gamma(z)| \cdot dm(z). \tag{5. 2}$$

[8]) Le cas $\gamma = 0$, mis à part dans la démonstration, s'y range d'une manière évidente.

Cette intégrale existe. En effet, $|\gamma(z)| = (\gamma(z)^* \gamma(z))^{\frac{1}{2}}$ est un opérateur autoadjoint de rang $\leqq r$, de borne indépendante de z et fonction (faiblement et alors aussi fortement) mesurable de z; par conséquent tr $|\gamma(z)|$ est une fonction mesurable bornée de z. La propriété $\|c\gamma\| = |c|\,\|\gamma\|$ est manifeste. La propriété $\|\gamma + \gamma'\| \leqq \|\gamma\| + \|\gamma'\|$ résulte de ce qu'on a même $\mathrm{tr}|\gamma(z) + \gamma'(z)| \leqq \mathrm{tr}|\gamma(z)| + \mathrm{tr}|\gamma'(z)|$ par points, et cela en vertu de l'inégalité

$$\mathrm{tr}\,|\tau + \tau'| \leqq \mathrm{tr}\,|\tau| + \mathrm{tr}\,|\tau'|,$$

valable pour des opérateurs τ, τ' (d'un espace de Hilbert dans soi-même ou dans un autre), de la „classe de trace" („trace class"), classe qui comprend en particulier les opérateurs de rang fini. [9])

On déduit de (5. 1) que l'opérateur $\gamma(z)^*$ (de $\mathfrak{F}$ dans E^ω) est défini par

$$\gamma(z)^* f = [(f, \gamma_k(z))_{\mathfrak{F}}]_{k=1}^{\omega} \qquad (f \in \mathfrak{F})$$

et par conséquent on a

$$(5.\,3) \qquad \gamma(z)^* \gamma(z) x = \left[\sum_{j=1}^{\omega} x_j (\gamma_j(z), \gamma_k(z))_{\mathfrak{F}}\right]_{k=1}^{\omega} \qquad (x \in E^\omega),$$

c'est-à-dire la matrice de $\gamma(z)^*\gamma(z)$ par rapport à la base $\{e^{(k)}\}$ a ses éléments

$$(5.\,4) \qquad m_{kj}(z) = (\gamma_j(z), \gamma_k(z))_{\mathfrak{F}} \qquad (j, k = 1, \cdots, \omega).$$

2. Envisageons le sous-ensemble $\boldsymbol{H}$ de $\boldsymbol{L}$, évidemment linéaire, constitué des $\gamma \in \boldsymbol{L}$ tels que $\gamma_k \in H^\infty(\mathfrak{F})$ pour tous les k, et la fonctionnelle linéaire L dans $\boldsymbol{H}$ définie par

$$(5.\,5) \qquad L(\gamma) = \sum_{k=1}^{\omega} (\gamma_k, \alpha_k) \qquad (\gamma \in \boldsymbol{H}).$$

D'après (4. 23) et la remarque y ajoutée, L vérifie l'inégalité

$$(5.\,6) \qquad |L(\gamma)| \leqq \|X\| \cdot \|\gamma\|.$$

En vertu du théorème de Hahn—Banach il existe donc un prolongement de L à tout l'espace $\boldsymbol{L}$ de façon qu'elle reste une fonctionnelle linéaire vérifiant (5. 6) pour tous les $\gamma \in \boldsymbol{L}$.

Pour chaque nombre naturel $r \leqq \omega$ définissons la fonctionnelle L_r sur l'espace $L^\infty(\mathfrak{F})$ par

$$(5.\,7) \qquad L_r(v) = L(ve^{(r)}) \qquad (v \in L^\infty(\mathfrak{F})).$$

Pour $\gamma = ve^{(r)}$ la matrice (5. 4) a tous ses éléments 0 sauf l'élément $m_{rr}(z)$ qui est égal à $\|v(z)\|_{\mathfrak{F}}^2$. Il en résulte que tr $|\gamma(z)| = \|v(z)\|_{\mathfrak{F}}$ et par conséquent

$$\|ve^{(r)}\| = \int \|v(z)\|_{\mathfrak{F}}\, dm(z) = \|v\|_1,$$

[9]) Cf. [3], lemma 5.14.

d'où

(5. 8) $$|L_r(v)| \leqq \|X\| \cdot \|v\|_1 \qquad (v \in L^\infty(F)).$$

Le dual de l'espace $L^1(\mathfrak{F})$ étant $L^\infty(\mathfrak{F})$, il s'ensuit de (5. 8) qu'il existe des fonctions $\beta_r \in L^\infty(\mathfrak{F})$ $(r=1, \cdots, \ )$ telles que

(5. 9) $$L_r(v) = \int (v(z), \beta_r(z))_{\mathfrak{F}}\, dm(z).$$

Comme v et β_r sont compris à fortiori dans $L^2(\mathfrak{F})$, (5. 9) s'écrit aussi sous la forme

(5. 10) $$L_r(v) = (v, \beta_r) \quad \text{(produit scalaire dans } L^2(\mathfrak{F})).$$

Dans le cas particulier où $v \in H^\infty(\mathfrak{F})$, on a $ve^{(r)} \in \boldsymbol{H}$ et il résulte de (5. 5), (5. 7) et (5. 10) que

(5. 11) $$(v, \beta_r) = (v, \alpha_r) \qquad (r = 1, \cdots, \omega).$$

Puisque les fonctions bornées sont partout denses dans $H^2(\mathfrak{F})$ dans la métrique de ce dernier espace (il n'y a qu'à prendre les sommes partielles de la série de Fourier) on conclut que (5. 11) *subsiste pour toutes les fonctions* $v \in H^2(\mathfrak{F})$.

3. Envisageons un élément $\gamma \in \boldsymbol{L}$ quelconque. On peut l'écrire sous la forme

$$\gamma = \sum_1^\omega \gamma_k e^{(k)} \quad \text{(il n'y a qu'un nombre fini de termes } \neq 0);$$

vu (5. 7) et (5. 10) il en dérive

(5. 12) $$L(\gamma) = \sum_1^\omega L(\gamma_k e^{(k)}) = \sum_1^\omega L_k(\gamma_k) = \sum_1^\omega (\gamma_k, \beta_k).$$

Soit en particulier γ la fonction aux composantes

(5. 13) $$\gamma_k(z) = \bar{x}_k \cdot \delta(z) \cdot f \qquad (k = 1, \cdots, \omega)$$

où $f \in \mathfrak{F}$, $\delta \in L^\infty$ à valeurs $\delta(z) \geqq 0$, et $x = [x_k]_1^\omega \in E^\omega$ tel que $x_k = 0$ pour tous les k qui dépassent un nombre fini r (dépendant de x). En vertu de (5. 12) on a alors

(5. 14) $$L(\gamma) = \sum_1^\omega \bar{x}_k \int \delta(z) (f, \beta_k(z))_{\mathfrak{F}}\, dm(z)$$

et la matrice correspondante (5. 4) a ses éléments

$$m_{kj}(z) = \bar{x}_j x_k \delta(z)^2 \|f\|_{\mathfrak{F}}^2.$$

Or la matrice hermitienne $[\bar{x}_j x_k]$ $(j, k = 1, \cdots, r)$ a son polynôme caractéristique

$$P(\xi) = \xi^r - \xi^{r-1} \sum_1^r \|x_j\|^2$$

parce que tous ses mineurs principaux d'ordre $\geqq 2$ s'annulent. Si $x \neq 0$, la seule valeur propre $\neq 0$ de cette matrice est donc égale à $\|x\|_{E^\omega}^2$ et par conséquent la racine carrée positive de cette matrice a la seule valeur propre $\neq 0$ égale à $\|x\|_{E^\omega}$, qui est alors égale à la trace de cette racine carrée. Ce résultat subsiste aussi dans le cas $x=0$. On en déduit que pour la fonction γ aux composantes (5. 13) on a

$$\operatorname{tr} |\gamma(z)| = \delta(z) \cdot \|f\|_{\mathfrak{F}} \|x\|_{E^\omega},$$

d'où

$$\|\gamma\| = \int \delta(z)\, dm(z) \cdot \|f\|_{\mathfrak{F}} \cdot \|x\|_{E^\omega}. \tag{5. 15}$$

En combinant (5. 6), (5. 14) et (5. 15) on obtient

$$\left| \sum_1^r \bar{x}_k \int \delta(z) (f, \beta_k(z))_{\mathfrak{F}}\, dm(z) \right| \leqq \|X\| \cdot \int \delta(z)\, dm(z) \cdot \|f\|_{\mathfrak{F}} \|x\|_{E^\omega}. \tag{5. 16}$$

Choisissons δ en particulier de la manière suivante:

$$\delta(z) = \frac{2\pi}{h} \quad \text{lorsque} \quad t \leqq \arg z < t+h \pmod{2\pi}$$

et $\delta(z)=0$ ailleurs. L'intégrale au second membre de (5. 16) est alors égale à 1 et, lorsque $h\to 0$, l'intégrale dans le premier membre tend vers

$$(f, \beta_k(\zeta))_{\mathfrak{F}} \quad (\zeta = e^{it})$$

en presque tous les points $\zeta \in C$. On a donc l'inégalité

$$\left| \sum_1^r \bar{x}_k\, (f, \beta_k(\zeta))_{\mathfrak{F}} \right| \leqq \|X\| \cdot \|f\|_{\mathfrak{F}} \cdot \|x\|_{E^\omega} \tag{5. 17}$$

en presque tous les points de C, l'ensemble exceptionnel pouvant dépendre de f. Grâce à la séparabilité de $\mathfrak{F}$ il existe alors un sous-ensemble ε de C, de mesure 0, tel que (5. 17) est vérifiée pour $z \notin \varepsilon$ quels que soient $f \in \mathfrak{F}$ et $x \in E^\omega$. Pour $\zeta \notin \varepsilon$ (5. 17) entraîne (en choisissant $f = \sum_1^r x_k \beta_k(\zeta)$):

$$\left\| \sum_1^r \beta_k(\zeta) x_k \right\|_{\mathfrak{F}} \leqq \|X\| \cdot \|x\|_{E^\omega} \quad \text{pour tout} \quad x \in E^\omega. \tag{5. 18}$$

6. Conclusion de la démonstration du théorème 1

Les fonctions β_r obtenues dans le n° 5 étant comprises dans $L^2(\mathfrak{F})$ on y peut appliquer l'opérateur Φ du n° 3; soient

$$b_r = \Phi \beta_r \qquad (r = 1, \cdots, \omega).$$

Vu que $\Phi\alpha_r = \Phi\Psi a_r = a_r$, *cf.* (4. 13), on déduit de (5. 11) que

$$(g, b_r) = (g, a_r) \qquad (r = 1, \cdots, \omega) \tag{6. 1}$$

pour tout $g \in \Phi H^2(\mathfrak{F}) = G$; *cf.* (4. 5). Les éléments g du sous-espace $L^2(\mathfrak{F}) \ominus H^2(\mathfrak{F})$ de G sont orthogonaux aux a_r (parce que $a_r \in H \subset H^2(\mathfrak{F})$); (6. 1) entraîne alors qu'ils sont orthogonaux aussi aux b_r, donc

$$b_r \in H^2(\mathfrak{F}) \qquad (r = 1, \cdots, \omega). \tag{6. 2}$$

D'autre part, envisagée pour les éléments g du sous-espace H de G, la relation (6. 1) entraîne que

$$a_r = P_H b_r \qquad (r = 1, \cdots, \omega). \tag{6. 3}$$

Appliquons la première des relations (4. 4) à $v = \sum_1^r x_k \beta_k$. Grâce à (5. 18) on obtient

$$\left\| \sum_1^r b_k(z) x_k \right\|_{\mathfrak{F}} \leqq \|X\| \cdot \|x\|_{E^\omega} \tag{6. 4}$$

pour tout $x \in E^\omega$ et pour tout $z \in C$ sauf peut-être les points z d'un ensemble $\bar{\varepsilon}$ de mesure 0. Il s'ensuit que si $z \notin \bar{\varepsilon}$, $\sum_1^\omega b_k(z) x_k$ converge (dans $\mathfrak{F}$) pour tout $x \in E^\omega$ et que

$$\left\| \sum_1^\omega b_k(z) x_k \right\|_{\mathfrak{F}} \leqq \|X\| \cdot \|x\|_{E^\omega}. \tag{6. 4*}$$

Définissons la fonction B, à valeurs $B(z)$ opérateurs de E^ω dans $H^2(\mathfrak{F})$, par

$$B(z)x = \sum_1^\omega b_k(z) x_k \qquad (x \in E^\omega). \tag{6. 5}$$

En vertu de (6. 4*) on a p.p.

$$\|B(z)x\|_{\mathfrak{F}} \leqq \|X\| \cdot \|x\|_{E^\omega},$$

donc $B \in H^\infty(E^\omega, \mathfrak{F})$ et

$$\|B\|_\infty \leqq \|X\|. \tag{6. 6}$$

Pour $\varphi \in H^\infty$ on a

$$P_H(B \cdot \varphi e^{(k)}) = P_H(\varphi \cdot B e^{(k)}) = P_H(\varphi \cdot P_H B e^{(k)}) \tag{6. 7}$$

parce que

$$P_H(\varphi \cdot (I - P_H) B e^{(k)}) \in P_H(\varphi \cdot \Theta H^2(\mathfrak{F})) \subset P_H \Theta H^2(\mathfrak{F}) = \{0\}.$$

Vu que $B(z)e^{(k)} = b_k(z)$, *cf.* (6. 5), on déduit de (6. 7), (6. 3) et (4. 9):

$$P_H B(\varphi e^{(k)}) = P_H(\varphi \cdot P_H b_k) = P_H(\varphi \cdot a_k) = X(\varphi e^{(k)}). \tag{6.8}$$

Soit $u \in H^\infty(E^\omega)$. On peut l'écrire sous la forme

$$u = \sum_1^\omega u_k e^{(k)} \qquad (u_k \in H^\infty)$$

(convergence dans la métrique de $H^2(E^\omega)$), donc, grâce à (6. 8),

$$Xu = \sum_1^\omega X(u_k e^{(k)}) = \sum_1^\omega P_H B(u_k e^{(k)}) = P_H Bu.$$

Toute fonction $u \in H^2(E^\omega)$ étant la limite, dans la métrique de $H^2(E^\omega)$, d'une suite de fonctions bornées $u^{(n)} \in H^\infty(E^\omega)$, l'équation

$$Xu = P_H Bu \tag{6.9}$$

s'étend par continuité à toutes les fonctions $u \in H^2(E^\omega)$.

Inversement, toute fonction $B' \in H^\infty(E^\omega, \mathfrak{F})$ engendre, moyennant la formule (6. 9), un opérateur X' de $H^2(E^\omega)$ dans H vérifiant (1. 1) et tel que

$$\|X'\| \leqq \|B'\|_\infty. \tag{6.10}$$

En effet, la linéarité de X' étant manifeste, il n'y a qu'à observer que pour $u \in H^2(E^\omega)$

$$\|X'u\| = \|P_H B'u\| \leqq \|B'u\| \leqq \|B'\|_\infty \|u\|$$

et que, en vertu de (1. 6—7),

$$TX'u = TP_H B'u = P_H(z \cdot B'u) = P_H B'(zu) = X'(zu) = X'Su.$$

En choisissant pour B' la fonction B que nous avons construite plus haut, (6. 6) et (6. 10) entraînent qu'il y a égalité dans (6. 6).

Cela achève la démonstration du théorème 1 et finit notre étude.

Ouvrages cités

[A] B. Sz.-Nagy et C. Foiaş, *Analyse harmonique des opérateurs de l'espace de Hilbert* (Budapest, 1967).

[1] D. B. Lowdenslager, On factoring matrix valued functions, *Annals of Math.*, **78** (1963), 450—454.

[2] D. Sarason, Generalized interpolation in H^∞, *Trans. Amer. Math. Soc.*, **127** (1967), 179—203.

[3] R. Schatten, *A theory of cross-spaces* (Annals of Math. Studies, vol. 26, Princeton, 1950).

(Reçu le 20. octobre 1967)

Operators of the form C^*C in indefinite inner product spaces

By J. BOGNÁR and A. KRÁMLI in Budapest

To professor M. G. Krein on the occasion of his 60th anniversary

1. Introduction

Let A denote a continuous self-adjoint operator in a J-space H (for definitions see Sections 2 and 3 below). In the present paper we give a necessary and sufficient condition (Theorem 1) for the existence of a continuous linear operator C in H such that

(1) $$A = C^*C.$$

In the special case when the space H is of type H_k we obtain (in Theorems 2 and 3) the solution of a problem proposed in [1]. Partial answers to other questions contained in [1] are to be found in the communications [2]—[5].

We mention that Theorem 2 is equivalent to an early result of POTAPOV ([6], Chapter 2, Theorem 2).

In J-spaces whose positive and negative components are of equal infinite dimension it turns out that the representation (1) is always possible (Theorem 4).

In some J-spaces property (1) is known [1] to be less restrictive than the existence of a self-adjoint square root:

(2) $$A = B^2 \qquad (B^* = B).$$

Therefore we do not hope that our conditions would have a significance similar to that of positivity in Hilbert space. However, they are so simple comparatively to the criteria for (2), contained in [2] that it seems desirable to use the factorization (1) instead of (2) as far as possible.

Lemmas 1—3 are known; Lemmas 4—7 slightly generalize some results of GINZBURG, JOHVIDOV and WITTSTOCK. It should be noted that Lemmas 5 and 6 show the invariant character of some of the notions applied, but actually they are not used in the following.

2. Basic facts concerning J-spaces

We consider a complex vector space H and a hermitian form $(\cdot, \cdot)$ defined on $H \times H$. The corresponding quadratic form is not assumed to be positive definite. We shall say that (x, y) is the *inner product* of the elements $x, y \in H$.

Two elements, x and y are *orthogonal* to each other if $(x, y)=0$. Two sets $F, G \subset H$ are said to be orthogonal if any element of F is orthogonal to any element of G.

An element $x \in H$ is called *positive* if $(x, x)>0$, *neutral* if $(x, x)=0$, and *negative* if $(x, x)<0$. A subspace (linear manifold) $L \subset H$ is said to be positive (neutral, negative) if all its elements except 0 are positive (neutral, negative).

The positive (negative) subspace L is *intrinsically complete* if it is complete with respect to the *intrinsic norm*

$$|x|_L=|(x, x)|^{\frac{1}{2}} \qquad (x \in L). \tag{3}$$

In the following we assume that H is a *J-space* i. e. H is an orthogonal direct sum

$$H=H^+ \oplus H^- \tag{4}$$

of an intrinsically complete positive subspace H^+ and an intrinsically complete negative subspace H^-.

In the special case $\dim H^- =k<\infty$ we say that H is a *space of type* H_k. Spaces with $\dim H^+<\infty$ have essentially the same properties.

In a J-space H we put

$$[x, y] = (x^+, y^+)-(x^-, y^-) \qquad (x, y \in H) \tag{5}$$

where $x=x^+ +x^-$, $y=y^+ +y^-$ denote the decompositions of x and y corresponding to (4).

It is evident that $[x, x]>0$ if $x \neq 0$. Therefore the hermitian form (5) may be called the *definite inner product* belonging to the decomposition (4). By definition, the J-space is a Hilbert space with respect to this definite inner product. The functional

$$\|x\| =[x, x]^{\frac{1}{2}} \qquad (x \in H) \tag{6}$$

is called the *norm* belonging to (4).

The norm (6) defines a topology in H. In the following the words "*closed*", "*continuous*", etc. will always refer to this topology.

Lemma 1. *For any $x, y \in H$ we have* $|(x, y)| \leqq \|x\| \, \|y\|$.

Proof. We shall use the same notations as in (5). By the orthogonality of H^+ and H^- we have

$$(x^+, y^-) = (x^-, y^+) = 0.$$

On the other hand, it follows from (5) that

$$[x^+, x^-] = [y^+, y^-] = 0.$$

Thus we can write

$$|(x, y)| \leqq |(x^+, y^+)| + |(x^- y^-)| = |[x^+, y^+]| + |[x^-, y^-]| \leqq \|x^+\| \|y^+\| + \|x^-\| \|y^-\| \leqq$$
$$\leqq (\|x^+\|^2 + \|x^-\|^2)^{\frac{1}{2}} (\|y^+\|^2 + \|y^-\|^2)^{\frac{1}{2}} = \|x\| \|y\|.$$

Let T be a continuous (everywhere defined) linear operator in the J-space H By virtue of Lemma 1 (Tx, y) is a continuous linear form in x and the Riesz representation theorem assures the existence of an element $y_* \in H$ such that

$$(Tx, y) = [x, y_*] \qquad (x \in H).$$

Setting

$$T^* y = y_*^+ - y_*^- \qquad (y \in H)$$

where $y_* = y_*^+ + y_*^-$ $(y_*^+ \in H^+, y_*^- \in H^-)$ we obtain

$$(Tx, y) = (x, T^* y) \qquad (x, y \in H).$$

One verifies easily that T^* is a single-valued continuous linear operator. We call T^* the *adjoint* of T. The operator T is said to be *self-adjoint* provided $T^* = T$.

For a continuous self-adjoint operator A one can define the *A-inner product* by

(7) $$(x, y)_A = (Ax, y) \qquad (x, y \in H).$$

The form $(\cdot, \cdot)_A$ is hermitian and continuous (Lemma 1). In the special case $A = I$ it turns into the original inner product $(\cdot, \cdot)$.

Using the A-inner product the notions of *A-orthogonality*, *A-positivity*, *intrinsic A-completeness* etc. can be introduced in the same way as orthogonality, positivity, intrinsic completeness have been defined with the help of the original inner product. The *intrinsic A-norm* on an A-positive or A-negative subspace L has the form

(8) $$|x|_{A, L} = |(x, x)_A|^{\frac{1}{2}} \qquad (x \in L).$$

An *A-fundamental decomposition* is a representation of H as the A-orthogonal direct sum of an A-neutral subspace H_A^0, an A-positive subspace H_A^+ and an A-negative subspace H_A^-:

(9) $$H = H_A^0 \dot{+} H_A^+ \dot{+} H_A^-.$$

In the case $A = I$ we speak of a *fundamental decomposition*. E. g. the decomposition (4) appearing in the definition of a J-space is a fundamental one.

An A-fundamental decomposition (9) is *regular* if $H_A^+ \dot{+} H_A^-$ is closed.

Lemma 2. (See e.g. [7], § 3, section 2.) *Let A be a continuous self-adjoint operator in the J-space H. Then H admits at least one regular A-fundamental decomposition.*

Lemma 3. *The A-neutral component H_A^0 of an arbitrary A-fundamental decomposition* (9) *consists of all elements in H that are A-orthogonal to H:*

(10) $$H_A^0 = \{x : (x, y)_A = 0 \text{ for every } y \in H\}.$$

Proof. If $x \in H_A^0$ then x is A-orthogonal to H_A^+ and H_A^- by the definition of the A-fundamental decomposition. Furthermore, $(\cdot\,,\cdot)_A$ is a semi-definite form on H_A^0, hence the Schwarz inequality $|(x, y)_A|^2 \leqq (x, x)_A(y, y)_A$ $(x, y \in H_A^0)$ is valid. It follows that x is A-orthogonal to H_A^0.

If, conversely, the element

(11) $$x = x_A^0 + x_A^+ + x_A^- \qquad (x_A^0 \in H_A^0,\ \ x_A^+ \in H_A^+,\ \ x_A^- \in H_A^-)$$

is A-orthogonal to H then $(x, x_A^+)_A = (x_A^+, x_A^+)_A = 0$ and $(x, x_A^-)_A = (x_A^-, x_A^-)_A = 0$ But H_A^+ is A-positive and H_A^- is A-negative. Therefore x_A^+ and x_A^- must be 0.

As a corollary we obtain that every fundamental decomposition of a J-space is of the form (4).

Lemma 4. *Each component of a regular A-fundamental decomposition* (9) *is closed.*

Proof. H_A^0 is closed by Lemma 3 and the continuity of the A-inner product (cf. Lemma 1).

Denote by $\overline{H}_A^+$ the relative closure of H_A^+ in $H_A^+ \dotplus H_A^-$. If $\overline{H}_A^+ \neq H_A^+$ then $\overline{H}_A^+$ has a non-trivial intersection with the A-negative subspace H_A^-. But this is impossible, since it follows from the continuity of the A-inner product that $\overline{H}_A^+$ is A-non-negative. Hence H_A^+ is closed in $H_A^+ \dotplus H_A^-$ and, the decomposition (9) being regular, in H as well.

For H_A^- the argument is similar.

3. Invariant properties. Intrinsic A-dimension

We consider an A-fundamental decomposition (9) and define

(12) $$[x, y]_A = (x_A^+, y_A^+)_A - (x_A^-, y_A^-)_A \qquad (x, y \in H).$$

Here

$$x = x_A^0 + x_A^+ + x_A^-, \quad y = y_A^0 + y_A^+ + y_A^-$$

are the decompositions corresponding to (9).

It is clear that $[x, x]_A \geqq 0$ for every $x \in H$. We call $[\cdot\,,\cdot]_A$ the *semi-definite A-inner product* belonging to (9). The corresponding *A-semi-norm* is

(13) $$\|x\|_A = [x, x]_A^{\frac{1}{2}} \qquad (x \in H).$$

Lemma 5. *The A-semi-norms belonging to any two A-fundamental decompositions are topologically equivalent.*

Proof. Let H_A^0 be defined by (10), and put

$$H_A^{(1)} = \{x : [x, y] = 0 \text{ for every } y \in H_A^0\}.$$

Then $H_A^{(1)}$ is a closed subspace and we have

(14) $$H = H_A^0 \dotplus H_A^{(1)}.$$

We consider an arbitrary A-fundamental decomposition (9), and set

(15) $$Vx = x_A^0 + x_A^{(1)} \qquad (x \in H)$$

where x_A^0 is defined by (11), and $x_A^{(1)}$ is the component of x in $H_A^{(1)}$ corresponding to the decomposition (14):

(16) $$x = x_A^{(0)} + x_A^{(1)} \qquad (x_A^{(0)} \in H_A^0, x_A^{(1)} \in H_A^{(1)}).$$

It is evident that V is a one-to-one linear mapping of H onto H which leaves the elements of H_A^0 fixed, and carries $H_A^+ \dotplus H_A^-$ into $H_A^{(1)}$. Moreover, according to (15), (16) and (10) the identity

(17) $$(Vx, Vy)_A = (x, y)_A \qquad (x, y \in H)$$

holds. Introducing the notations

(18) $$VH_A^+ = H_A^{(+)}, \quad VH_A^- = H_A^{(-)}$$

we obtain an A-fundamental decomposition

(19) $$H = H_A^0 \dotplus H_A^{(+)} \dotplus H_A^{(-)}$$

where

(20) $$H_A^{(+)} \dotplus H_A^{(-)} = H_A^{(1)}.$$

Let

(21) $$Vx_A^+ = x_A^{(+)}, \quad Vx_A^- = x_A^{(-)}.$$

Then in virtue of (18) we have $x_A^{(+)} \in H_A^{(+)}$, $x_A^{(-)} \in H_A^{(-)}$ and the relations (11), (21) imply that

(22) $$Vx = x_A^0 + x_A^{(+)} + x_A^{(-)}.$$

Comparing (22) with (15) we obtain:

(23) $$x_A^{(+)} + x_A^{(-)} = x_A^{(1)}.$$

The equalities (23) and (16) yield:

(24) $$x = x_A^{(0)} + x_A^{(+)} + x_A^{(-)} \qquad (x_A^{(0)} \in H_A^0, x_A^{(+)} \in H_A^{(+)}, x_A^{(-)} \in H_A^{(-)}).$$

The A-semi-norm belonging to (19) is

$$\|x\|_A^{(\ )} = \left((x_A^{(+)}, x_A^{(+)})_A - (x_A^{(-)}, x_A^{(-)})_A\right)^{\frac{1}{2}} \qquad (x \in H). \tag{25}$$

Taking the analogous definition (13), (12) of $\|x\|_A$ and the relations (21), (17) into account we see that

$$\|x\|_A = \|x\|_A^{(\ )} \qquad (x \in H).$$

But, according to (23), (24) and (25),

$$\|x\|_A^{(\ )} = \|x_A^{(1)}\|_A^{(\ )} \qquad (x \in H).$$

Hence

$$\|x\|_A = \|x_A^{(1)}\|_A^{(\ )} \qquad (x \in H).$$

Consequently, it is sufficient to show that for any two A-fundamental decompositions, which are of the form (19) and satisfy (20), the corresponding A-semi-norms (25) are topologically equivalent on the closed subspace $H_A^{(1)}$. As, by virtue of (10) and (14), the A-inner product is non-degenerate on $H_A^{(1)}$, i.e. for $x \in H_A^{(1)}$ $(x \neq 0)$ there is an element $y \in H_A^{(1)}$ such that $(x, y)_A \neq 0$, the statement of our lemma follows from WITTSTOCK's theorem ([8], Theorem 15; cf. also [9]).

In the special case $A = I$ Lemma 5 (or WITTSTOCK's theorem itself) asserts that the topology of H does not depend on the choice of the fundamental decomposition (4).

Lemma 6. *If the component $H_A^{(+)}$ $(H_A^{(-)})$ of an A-fundamental decomposition* (19) *is intrinsically A-complete then the respective component H_A^+ (H_A^-) of any other A-fundamental decomposition* (9) *is also intrinsically A-complete.*

Proof. We denote by P the projection operator belonging to the subspace H_A^+ and the decomposition (9), i. e.

$$Px = x_A^+ \qquad (x \in H)$$

where x_A^+ is defined by (11).

According to (12) and (13) we have

$$\|x_A^+\|_A^2 + \|x_A^-\|_A^2 = \|x\|_A^2 \qquad (x \in H). \tag{26}$$

On the other hand, $\|x_A^+\|_A^2 - \|x_A^-\|_A^2 = (x, x)_A \geqq 0$ for $x \in H_A^{(+)}$, so that

$$0 \leqq \|x_A^-\|_A^2 \leqq \|x_A^+\|_A^2 \qquad (x \in H_A^{(+)}). \tag{27}$$

Using (27) we obtain from (26)

$$\|x_A^+\|_A \leqq \|x\|_A \leqq \sqrt{2}\,\|x_A^+\|_A \qquad (x \in H_A^{(+)}). \tag{28}$$

Since $\|\,.\,\|_A$ is a norm on both of the A-positive subspaces $H_A^{(+)}$, H_A^+, the relations (28) and the definition of P imply that with respect to $\|\,.\,\|_A$ the operator

P induces a topological isomorphism (a linear, one-to-one, bicontinuous mapping) between $H_A^{(+)}$ and the subspace $PH_A^{(+)} \subset H_A^+$ (cf. [7]).

If $H_A^{(+)}$ is intrinsically A-complete then it is complete in the norm

$$|x|_{A,\,H_A^{(+)}} = \|x\|_A^{(\ )} \qquad (x \in H_A^{(+)})$$

(see (8) and (25)) and, as a consequence of Lemma 5, in the norm $\|\,.\,\|_A$. Therefore the image $PH_A^{(+)}$ is also complete in the norm

$$\|x\|_A = |x|_{A,\,H_A^+} = |x|_{A,\,PH_A^{(+)}} \quad (x \in PH_A^{(+)}).$$

In other words, the subspace $PH_A^{(+)}$ is intrinsically A-complete.

We shall show that $PH_A^{(+)} = H_A^+$. Assuming the contrary, the intrinsic A-completeness of $PH_A^{(+)}$ would imply the existence of an element $x_0 \in H_A^+$ $(x_0 \neq 0)$ which is A-orthogonal to $PH_A^{(+)}$. Then x_0 is A-orthogonal to $H_A^{(+)}$, so that the span of x_0 and $H_A^{(+)}$ is an A-positive extension of $H_A^{(+)}$. But this is impossible because, in virtue of (19), any subspace properly containing $H_A^{(+)}$ has a non-trivial intersection with the A-non-positive subspace $H_A^0 \dotplus H_A^{(-)}$.

For an intrinsically A-complete $H_A^{(-)}$ the proof is similar.

In the special case $A = I$ we obtain that the components of any fundamental decomposition of a J-space are intrinsically complete.

Consider an A-positive or A-negative subspace $L \subset H$. The dimension of the completion of L with respect to the intrinsic A-norm (8) will be called the *intrinsic A-dimension* of L. It is equal to the minimal power of those systems in L which are complete in L with respect to (8). The equivalence of the two definitions follows essentially by the same argument as the separability of the subsets of a separable metric space (see [10], Section 33).

Ju. L. Šmul'jan called our attention to the fact that for a closed A-positive or A-negative subspace L the intrinsic A-dimension coincides with the usual Hilbert dimension. This can be seen as follows.

L is a Hilbert space with respect to the definite inner product (5). Let L be A-positive. Then (Ax, y) is a continuous positive form on L, and there exists a continuous positive operator B acting in the Hilbert space L such that

$$(Ax, y) = [Bx, y] \qquad (x, y \in L).$$

Taking the positive square root $B^{\frac{1}{2}}$ we have

$$(Ax, y)) = [B^{\frac{1}{2}}x, B^{\frac{1}{2}}y] \qquad (x, y \in L).$$

Therefore if a system $\{e_\gamma\}_{\gamma \in \Gamma}$ is complete in L with respect to the A-inner product then $\{B^{\frac{1}{2}}e_\gamma\}_{\gamma \in \Gamma}$ is complete in $B^{\frac{1}{2}}L$ with respect to the definite inner product. As $B^{\frac{1}{2}}L$ is dense in L we obtain that the Hilbert dimension of L is not greater than the intrinsic A-dimension of L. The converse inequality is trivial.

Instead of "intrinsic I-dimension" we shall use the term *intrinsic dimension.*

Lemma 7. *Let $L^+(L^-)$ be an A-positive (A-negative) subspace of the J-space H, and let* (9) *denote any A-fundamental decomposition of H. Denote the intrinsic A-dimensions of L^+ and H_A^+ (L^- and H_A^-) by d^+ and k_A^+ (d^- and k_A^-) respectively. Then $d^+ \leqq k_A^+$ ($d^- \leqq k_A^-$). In particular, the cardinal numbers k_A^+, k_A^- do not depend on the choice of the A-fundamental decomposition.*

Proof. In the same way as it has been done in the first half of the preceding proof one can show that, with respect to the A-seminorm $\| \, . \, \|_A$ belonging to the decomposition (9), L^+ is topologically isomorphic to a subspace of H_A^+. Observing that

$$\|x\|_A = |x|_{A, H_A^+} \qquad (x \in H_A^+)$$

we obtain the inequality $d_1^+ \leqq k_A^+$ where d_{1A}^+ stands for the dimension of the completion of L^+ with respect to $\| \, . \, \|_A$, i. e. for the minimal power of systems in L^+ which are complete in L^+ with respect to $\| \, . \, \|_A$.

On the other hand, for $x \in L^+$ we have

$$|x|_{A, L^+}^2 = (x, x)_A = (x_A^+, x_A^+)_A + (x_A^-, x_A^-)_A \leqq (x_A^+, x_A^+)_A - (x_A^-, x_A^-)_A = \|x\|_A^2.$$

Therefore $d^+ \leqq d_1^+$.

We have proved that $d^+ \leqq k_A^+$. The inequality $d^- \leqq k_A^-$ can be verified similarly.

4. The representation $A = C^*C$

Theorem 1. *Consider a continuous self-adjoint operator A in the J-space H. Denote by k^+ and k^- the intrinsic dimension of the positive resp. negative component of a fundamental decomposition, and by k_A^+ and k_A^- the intrinsic A-dimension of the A-positive resp. A-negative component of an A-fundamental decomposition of H. Then A admits a representation* (1) *with a continuous linear operator C if and onlv if*

(29) $$k_A^+ \leqq k^+$$

and

(30) $$k_A^- \leqq k^-.$$

Proof. First we remark that (1) is equivalent to the identity

(31) $$(Ax, y) = (Cx, Cy) \qquad (x, y \in H).$$

Now we assume that A and C satisfy (31). Applying Lemma 2 we choose some A-fundamental decomposition (9) and put $CH_A^+ = R^+$. Then, in virtue of (31), R^+ is a positive subspace, and C is a linear one-to-one mapping of H_A^+ onto R^+.

Applying the relation (31) once again we obtain that the intrinsic dimension r^+ of R^+ is equal to the intrinsic A-dimension k_A^+ of H_A^+. But, according to Lemma 7, $r^+ \leqq k^+$. Therefore (29) is valid. The inequality (30) can be proved in the same way.

Conversely, let the operator A satisfy the relations (29), (30). Consider a regular A-canonical decomposition (9) (Lemma 2) and a fundamental decomposition (4) of the space H.

The completion $\tilde{H}_A^+$ of H_A^+ with respect to the A-inner product is a Hilbert space of dimension k_A^+ (Lemma 7). On the other hand, H^+ is a Hilbert space of dimension k^+ with respect to the original inner product (Lemmas 6 and 7). It follows from (29) that there exists a linear isometric imbedding of $\tilde{H}_A^+$ into H^+. Restricting the imbedding operator to H_A^+ we obtain a linear operator C^+ such that

$$(Ax, y) = (C^+x, C^+y) \qquad (x, y \in H_A^+). \tag{32}$$

Analogously, one can find a linear operator C^-, which maps H_A^- into H^- and has the property

$$(Ax, y) = (C^-x, C^-y) \qquad (x, y \in H_A^-). \tag{33}$$

For an arbitrary element (11) we define

$$Cx = C^+x_A^+ + C^-x_A^- \qquad (x \in H). \tag{34}$$

C is a linear operator of H to itself. Moreover, as a consequence of (32), (33), the orthogonality of the decomposition (4), the A-orthogonality of the decomposition (9), and the A-orthogonality of H_A^0 to H (Lemma 3), C fulfils the relation (31).

It remains to prove that C is continuous. For this purpose we apply the norm (6) which belongs to the fundamental decomposition (4) occuring in the above construction (cf. Lemma 5). We have

$$\|Cx\|^2 = (Cx_A^+, Cx_A^+) - (Cx_A^-, Cx_A^-) = (Ax_A^+, x_A^+) - (Ax_A^-, x_A^-).$$

Therefore, by Lemma 1, one obtains

$$\|Cx\|^2 \leqq \|A\| (\|x_A^+\|^2 + \|x_A^-\|^2). \tag{35}$$

As we are considering a regular A-fundamental decomposition, H_A^0 and $H_A^+ \dotplus H_A^-$ are closed subspaces of the complete space H (Lemma 4). Thus, according to a well-known corollary to BANACH's theorem, $x_A^+ \dotplus x_A^-$ depends continuously on x. On the other hand, H_A^+ and H_A^- are closed subspaces of H (Lemma 4) and, consequently, they are closed subspaces of $H_A^+ \dotplus H_A^-$. A second application of the Banach theorem yields that x_A^+ and x_A^- depend continuously on $x_A^+ + x_A^-$. As a result, x_A^+ and x_A^- are continuous functions of the element x. This fact together with the relation (35) implies the continuity of the operator C.

The theorem is proved.

In the following we consider some consequences of Theorem 1.

Theorem 2. *Let A denote a self-adjoint operator in an n-dimensional space H of type H_k $(0 \leqq k \leqq n < \infty)$. Then A admits a representation* (1) *with a linear C if and only if the following two conditions are fulfilled:*

α) H contains an A-non-positive subspace of dimension k;

β) H does not contain any A-negative subspace of dimension $k+1$.

Proof. With the notations of Theorem 1 we have:

$$k^- = k, \quad k^+ = n - k. \tag{36}$$

Consider an A-fundamental decomposition (9) and put $\dim H_A^0 = k_A^0$. If L is a subspace and $\dim L > k_A^0 + k_A^-$, then L has a non-trivial intersection with the A-positive subspace H_A^+. Analogously, if $\dim L > k_A^-$, then L has a non-trivial intersection with the A-non-negative subspace $H_A^0 \dotplus H_A^+$. Therefore $k_A^0 + k_A^-$ (k_A^-) is equal to the maximal dimension of A-non-positive (resp. A-negative) subspaces.

It follows from the foregoing that the conditions α), β) can be written in the form

$$k_A^0 + k_A^- \geqq k, \tag{37}$$

resp.

$$k_A^- \leqq k. \tag{38}$$

By virtue of (36) the relations (37), (38) are equivalent to the conditions (29), (30) in Theorem 1.

Theorem 3. *Consider an infinite-dimensional space H of type H_k, and a continuous self-adjoint operator A in H. The representation* (1), *where C denotes a continuous linear operator, is possible if and only if A satisfies condition β) of Theorem* 2.

Proof. One of the statements of the preceding proof remains valid for the present situation in the modified form that k_A^- is equal to the maximal dimension of A-negative subspaces, provided that one of these numbers is finite. Consequently, β) is equivalent to (38) or, what is the same, to (30) even now.

On the other hand, the ordinary dimension (i. e. the dimension with respect to a norm (6)) of H is $k^+ + k^- = k^+ + k = k^+$. As the continuity of the A-inner product (see Lemma 1) implies that the intrinsic A-dimension of the component H_A^+ of an A-fundamental decomposition (9) is not greater than the ordinary dimension of H_A^+, in our case the inequality (29) holds for every A.

Now the conclusion of our theorem follows from Theorem 1.

Theorem 4. *If the J-space H has infinite dimension, and the cardinal numbers k^+ and k^- defined in Theorem* 1 *are equal to each other, then any continuous self-adjoint operator A in H admits the representation* (1) *with a continuous linear C.*

Proof. In the present case both of the relations (29), (30) are always satisfied, since the ordinary dimension of H is equal to $k^+ + k^- = k^+ = k^-$, and any ordinary dimension in H, a fortiori (see Lemma 1) any intrinsic A-dimension in H, does not exceed this common value. Therefore our theorem is a consequence of Theorem 1.

References

[1] J. Bognár, Non-negativity properties of operators in spaces with indefinite metric, *Ann. Acad. Sci. Fenn.*, A. I, no. 336/10 (1963).

[2] J. Bognár, О существовании квадратного корня из оператора, самосопряженного относительно индефинитной метрики, *MTA Mat. Kut. Int. Közl.*, **6** (1961), 351—363.

[3] J. Bognár, О некоторых соотношениях между свойствами неотрицательности операторов в пространствах с индефинитной метрикой, *MTA Mat. Kut. Int. Közl.*, **8** (1963), 201—212.

[4] ——— II, *Studia Sci. Math. Hung.*, **1** (1966), 97—102.

[5] ——— III, *Studia Sci. Math. Hung.*, **1** (1966), 419—426.

[6] В. П. Потапов, Мультипликативная структура J-нерастягивающих матриц-функций, *Труды Московского матем. общества*, **4** (1955), 125—236.

[7] Ю. П. Гинзбург—И. С. Иохвидов, Исследования по геометрии бесконечномерных пространств с билинейной метрикой, *Успехи матем. наук*, **17** (1962), No. 4, 3—56.

[8] G. Wittstock, Über Zerlegungsmajoranten indefiniter Metriken, *Math. Zeitschr.*, **91**, (1966) 421—430.

[9] J. Bognár, On decomposition majorants of an indefinite metric, *Math. Zeitschr.*, **101** (1967), 65—67.

[10] F. Riesz—B. Sz.-Nagy, *Leçons d'analyse fonctionnelle* (3rd edition, Paris—Budapest, 1955).

(Received April 20, 1967)

The spectrum of the Cesàro operator

By D. W. BOYD in Edmonton (Canada)*)

Introduction

Suppose that x is a locally integrable function on $R^+ = [0, \infty)$ and that the Cesàro average of x is defined by

(1) $$Px(t) = \frac{1}{t} \int_0^t x(s)\, ds.$$

In [3], BROWN, HALMOS and SHIELDS considered the operator P as a bounded operator from $L^2(R^+)$ to itself and showed that the spectrum in this case is the circle

(2) $$\sigma(P; L^2) = \{\lambda : |\lambda - 1| = 1\}.$$

In this paper, we examine P as an operator in $L^p(R^+)$ when $p \neq 2$ and show that the spectrum in this case is the following set:

(3) $$\sigma(P; L^p) = \{\lambda : \mathrm{Re}\,(1/\lambda) = (p-1)/p\},$$

which, for $p > 1$, is a circle with centre $2(p-1)/p$ and the same radius, and for $p = 1$, is the imaginary axis.

The result can be extended to include certain rearrangement invariant spaces X, in which case the spectrum becomes the following lune:

(4) $$\sigma(P; X) = \{\lambda : 1 - \beta \leqq \mathrm{Re}\,(1/\lambda) \leqq 1 - \alpha\},$$

where α and β are the indices associated with the space X as in [1]. The proof for this will appear elsewhere.

The method of proof is to exhibit integral operators which are proved to be the resolvents of P for $\mathrm{Re}\,(1/\lambda) < (p-1)/p$ and $\mathrm{Re}\,(1/\lambda) > (p-1)/p$, respectively. A short additional argument then shows that the spectrum is indeed given by (3).

*) The author is presently at the California Institute of Technology, Pasadena.

Preliminary Lemmas

Let x be a locally integrable function, and let ζ be a complex number. Define the operators P_ζ and Q_ζ by

(5)
$$P_\zeta x(t) = \int_0^1 s^{-\zeta} x(st)\, ds,$$

whenever

$$\int_0^1 |s^{-\zeta} x(st)|\, ds < \infty \quad \text{a.e.},$$

and

(6)
$$Q_\zeta x(t) = \int_1^\infty s^{-\zeta} x(st)\, ds,$$

whenever

$$\int_1^\infty |s^{-\zeta} x(st)|\, ds < \infty \quad \text{a.e.}$$

We denote the space of bounded linear operators on L^p by $B(L^p)$ and the spectral radius and norm of $T \in B(L^p)$ by $r(T; L^p)$ and $\|T\|_p$, respectively.

Lemma 1. *Let* $1 \leqq p \leqq \infty$, *and the operators* P_ζ *and* Q_ζ *be defined by* (5) *and* (6).

(a) $P_\zeta \in B(L^p)$ *with domain all of* L^p *if and only if*

(7)
$$\operatorname{Re} \zeta < (p-1)/p \quad (=1, \ \textit{if} \ \ p = \infty).$$

In this case,

(8)
$$\|P_\zeta\|_p = r(P_\zeta; L^p) = \left[\frac{p-1}{p} - \operatorname{Re} \zeta\right]^{-1}.$$

(b) $Q_\zeta \in B(L^p)$ *with domain all of* L^p *if and only if*

(9)
$$\operatorname{Re} \zeta > (p-1)/p.$$

In this case,

(10)
$$\|Q_\zeta\|_p = r(Q_\zeta; L^p) = \left[\operatorname{Re} \zeta - \frac{p-1}{p}\right]^{-1}.$$

Proof. The proof that (7) implies that $P_\zeta \in B(L^p)$ and that (9) implies that $Q_\zeta \in B(L^p)$ can be derived from ([4], Th. 318). The other parts are given for real ζ in ([2], Theorem 2 and introductory remarks), and the proofs given there are easily extended to complex ζ.

Lemma 2. *Let* $1 \leqq p \leqq \infty$. *Let* $x \in L^p$ *be such that* $Px \in L^p$.

(a) *If* $P_\zeta \in B(L^p)$, *then* $PP_\zeta x \in L^p$, *and*

(11)
$$\zeta PP_\zeta x = \zeta P_\zeta Px = (P - P_\zeta)x,$$

(b) *If* $Q_\zeta \in B(L^p)$, *then* $PQ_\zeta x \in L^p$, *and*

$$\zeta P Q_\zeta x = \zeta Q_\zeta P x = (P + Q_\zeta)x. \tag{12}$$

Proof. (a) Since $Px \in L^p$, and $P_\zeta \in B(L^p)$, $P_\zeta P x \in L^p$, and

$$\int_0^1 |s^{-\zeta}(Px)(st)|\, ds < \infty, \quad t > 0. \tag{13}$$

We can write (13) as an iterated integral using the definition of P to show that

$$\int_0^1 s^{-\mathrm{Re}\,\zeta - 1}\, ds \int_0^s |x(ut)|\, du < \infty, \quad t > 0, \tag{14}$$

and then apply Fubini's theorem to the following iterated integral

$$\begin{aligned} \zeta P_\zeta P x(t) &= \zeta \int_0^1 s^{-\zeta-1}\, ds \int_0^s x(ut)\, du = \zeta \int_0^1 x(ut)\, du \int_u^1 s^{-\zeta-1}\, ds = \\ &= \int_0^1 (1 - u^{-\zeta}) x(ut)\, du = Px(t) - P_\zeta x(t). \end{aligned} \tag{15}$$

Also, changing variables in another way and using (14) to justify the interchange of order of integration,

$$P_\zeta P x(t) = \int_0^1 s^{-\zeta}\, ds \int_0^1 x(sut)\, du = \int_0^1 du \int_0^1 s^{-\zeta} x(sut)\, du = PP_\zeta x(t), \quad t > 0. \tag{16}$$

This proves (11). (Note that $\mathrm{Re}\,\zeta < 1$ is necessary for $P_\zeta \in B(L^p)$ by Lemma 1, so we have used this fact freely.)

(b) The proof of (12) follows the same pattern as in (a), and we leave the appropriate manipulations to the reader.

The resolvent of P

By Lemma 1, applied to $\zeta = 0$, it is clear that $P \in B(L^p)$ iff $1 < p \leqq \infty$. Of course, this is a well known result of Hardy. In case $p = 1$ we can define P as a closed linear operator with range L^1 and domain $D(P; L^1)$ dense in L^1 by the simple expedient of defining

$$D(P; L^1) = \left\{ x \in L^1 : \int_0^1 dt \int_0^1 |x(st)|\, ds < \infty \right\}. \tag{17}$$

To show $D(P; L^1)$ is dense in L^1, we note that it contains all functions in L^1 vanishing in a neighbourhood of 0. Since convergence in norm in L^1 implies convergence a.e., it is easy to prove that P is closed as an operator $D(P; L^1) \to L^1$.

For $p>1$, we define $D(P; L^p)=L^p$.

The resolvent set of P considered as an operator $D(P; L^p)\to L^p$ will be denoted $\varrho(P; L^p)$ and the spectrum by $\sigma(P; L^p)$.

Theorem 1. *Let λ be a complex number satisfying*

(18) $$\mathrm{Re}\,(1/\lambda)<(p-1)/p \quad or \quad \mathrm{Re}\,(1/\lambda)>(p-1)/p.$$

Then, $\lambda\in\varrho(P; L^p)$, and for each $x\in L^p$,

(19) $$(\lambda-P)^{-1}x=(\lambda^{-1}+\lambda^{-2}P_{1/\lambda})x, \quad \mathrm{Re}\,(1/\lambda)<(p-1)/p,$$

(20) $$(\lambda-P)^{-1}x=(\lambda^{-1}-\lambda^{-2}Q_{1/\lambda})x, \quad \mathrm{Re}\,(1/\lambda)>(p-1)/p.$$

Proof. Let $\zeta=\lambda^{-1}$. And $\mathrm{Re}\,(\zeta)<(p-1)/p$. From Lemma 1, we have $P_\zeta\in B(L^p)$, and from Lemma 2,

(21) $$(\lambda-P)(\zeta+\zeta^2P_\zeta)x=[I-\zeta P+\zeta P_\zeta-\zeta^2PP_\zeta]x=x,$$

and also

(22) $$(\zeta+\zeta^2P_\zeta)(\lambda-P)x=x,$$

for every $x\in D(P; L^p)$. But $D(P; L^p)$ is dense in L^p and hence (21) and (22) are enough to show that $(\lambda-P)$ has the bounded inverse $\zeta+\zeta^2P_\zeta$, for $\mathrm{Re}\,(\zeta)<(p-1)/p$.

Similarly, $(\lambda-P)$ has the bounded inverse given in (20) for $\mathrm{Re}\,(\zeta)>(p-1)/p$.

Theorem 2. *Let λ be a complex number satisfying* $\mathrm{Re}\,(1/\lambda)=(p-1)/p$. *Then $\lambda\in\sigma(P; L^p)$.*

Proof. Let λ_n be a sequence of complex numbers with $\mathrm{Re}\,(1/\lambda_n)<(p-1)/p$, approaching λ. Then by Lemma 1, if $\zeta_n=\lambda_n^{-1}$,

$$\|\zeta_n+\zeta_n^2P_\zeta\|_p\geqq|\zeta_n|^2\|P_{\zeta_n}\|-|\zeta_n|=|\zeta_n|^2\cdot[(p-1)/p-\mathrm{Re}\,\zeta_n]^{-1}-|\zeta_n|\to\infty$$

as $\zeta_n\to\zeta$. Hence $\lambda\in\sigma(P; L^p)$.

References

[1] D. W. Boyd, Spaces between a pair of reflexive Lebesgue spaces, *Proc. Amer. Math. Soc.*, **18** (1967), 215—219.

[2] ———— The spectral radius of averaging operators, *Pac. J. Mat.*, **29** (1967), 79—95.

[3] A. Brown, P. R. Halmos, A. L. Shields, Cesàro operators, *Acta Sci. Math.*, **26** (1965), 125—137.

[4] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities* (Cambridge, 1934).

(Received April 12, 1967)

Convergence of random products of Markov transition functions

By S. R. FOGUEL in Jerusalem (Israel)

1. Introduction

The purpose of this note is to apply a theorem of AMEMIYA and ANDO [1] to study a Markov process $\{x_n\}$ with the phase space (X, Σ) which is not stationary but *there exists a mapping* $n \to r(n)$ *of the set of all non-negative integers onto the finite set* $\{1, \ldots, N\}$ *such that the conditional probabilities*

$$\Pr\{x_{n+1} \in A | x_n = x\} \qquad (x \in X,\ A \in \Sigma,\ n = 0, 1, 2, \ldots)$$

depend (besides on A and X) only on $r(n)$, *i.e.*

$$\Pr\{x_{n+1} \in A | x_n = x\} = P_{r(n)}(x, A).$$

By means of $P_j(x, A)$ $(j = 1, \ldots, N)$ one defines in the Banach space of the finite measures on (X, Σ) operators $\nu \to \nu P_j$ by

$$(\nu P_j)(A) = \int P_j(x, A)\, \nu(dx). \tag{1.1}$$

Thus if μ is a probability measure on (X, Σ) and we choose it to be the distribution of x_0 then the distribution μ_n of x_n is, as well known,

$$\mu_n = \mu\, P_{r(0)} \cdots P_{r(n-1)}.$$

We shall prove the convergence of $\mu_n(A)$ as $n \to \infty$ under suitable conditions.

2. Convergence of random products

Let the mapping $n \to r(n)$ be defined on the set of all non-negative integers and assume the values $1, 2, \ldots, N$. Let $P_j(x, A)$ be subtransition functions on the measure space (X, Σ, λ) with λ being a σ-finite subinvariant measure: $P_j(x, A)$ is a function on $X \times \Sigma$ which is for each $x \in X$ a non-negative measure of total measure $\leqq 1$ and, for each $A \in \Sigma$, a measurable function and

$$\int P_j(x, A)\, \lambda(dx) \leqq \lambda(A). \tag{2.1}$$

It is well known that $P_j(x, A)$ induce contraction operators on $L_2(X, \Sigma, \lambda)$ by

$$(P_j f)(x) = \int P_j(x, dy) f(y) \tag{2.2}$$

From now on P_j will be considered as contraction operators on L_2 only, also every relation between functions will be a. e.

In [1] it is proved that if each P_j satisfies the condition:

$$\|P_j f\| = \|f\| \quad \text{implies } P_j f = f \tag{W'}$$

and each $1 \leqq j \leqq N$ appears infinitely often in the sequence $r(n)$ then the sequences of operators $P_{r(n)} \dots P_{r(1)}$, $P_{r(1)} \dots P_{r(n)}$ both converge weakly to the projection on the intersection $\bigcap_{j=1}^{N} \{f\colon P_j f = f\}$. Let us study property (W′) in our case.

Lemma 1. *Let $P(x, A)$ be a subtransition function on (X, Σ, λ) with λ being a subinvariant measure. Let $K = \{f\colon \|Pf\| = \|f\|\}$. Then K is generated by characteristic functions of sets of finite measure and if $1_A \in K$* [1]) *for some $A \in \Sigma$ then $P(x, A)$ assumes the values zero and one only.*

Proof. The equation $\|Pf\| = \|f\|$ is equivalent to $P^*Pf = f$ since P is a contraction operator. Thus K is a subspace of L_2.

Now if $f \geqq 0$ then $Pf \geqq 0$. Also if $f \geqq 0$ then $P^*f \geqq 0$: otherwise, if $P^*f < 0$ on a set A of positive finite measure, then $0 > \int_A P^*f \, d\lambda = \int P 1_A \cdot f \, d\lambda \geqq 0$. Thus if $f \in K$ and is real, $P^*P|f| \geqq |P^*Pf| = |f|$. Inequality is impossible since $\|P^*P\| \leqq 1$, hence $|f| \in K$. Now if $0 \leqq f \leqq c$ then $0 \leqq Pf \leqq c$, also $P^*f \leqq c$: otherwise, if $P^*f > c$ on a set of positive finite measure A, then

$$c\lambda(A) < \int_A P^*f \, d\lambda = \int f P 1_A \, d\lambda = \int f(x) P(x, A) \lambda(dx) \leqq c \int P(x, A) \lambda(dx) \leqq c\lambda(A).$$

Therefore if $0 \leqq f \in K$ and $c > 0$ then $P^*P(\min(f, c)) \leqq f$ and $P^*P(\min(f, c)) \leqq c$ hence $P^*P(f - \min(f, c)) = f - P^*P(\min(f, c)) \geqq f - \min(f, c)$. Inequality is impossible, hence $f - \min(f, c) \in K$ and therefore $\min(f, c) \in K$.

Thus the conditions of [2], Lemma 1, are satisfied and K is generated by the characteristic functions it contains. Let us conclude the proof by showing that if $1_A \in K$ then $P 1_A = P(x, A)$ is a characteristic function:

Put $f = P 1_A$ then $0 \leqq f \leqq 1$ and $P^*f = 1_A$.

---

[1]) $1_A(x) = \begin{cases} 1 & \text{for } x \in A, \\ 0 & \text{for } x \in X \setminus A. \end{cases}$

Now if $B_\varepsilon=\{x: f(x)>\varepsilon\}$ and C is a set of finite λ measure disjoint to A then $0\leqq\int P^*1_{B_\varepsilon}\cdot 1_C\,d\lambda\leqq\frac{1}{\varepsilon}\int P^*f\cdot 1_C\,d\lambda=0$ or $P^*1_{B_\varepsilon}\leqq 1_A=P^*f$. Thus

$$0\geqq\int P^*(1_{B_\varepsilon}-f)\cdot 1_A\,d\lambda=\int(1_{B_\varepsilon}-f)\cdot f\,d\lambda\to\int(1_B-f)\cdot f\,d\lambda$$

where $B=\{x: f(x)>0\}$. Now $1_B\geqq f$ hence $f=1_B$.

Lemma 2. *Let $P(x, A)$ satisfy the conditions of Lemma* 1. *The corresponding operator P on L_2 satisfies* (W′) *provided:*

(2. 3) $\qquad P(\cdot, A)=1_B$ *and* $\lambda(A)<\infty$ *implies* $1_B=1_A$.

Proof. Condition 2. 3 implies that every characteristic function in K is left invariant by P thus K itself is invariant under P by Lemma 1.

Remark. If $P(\cdot, A)=1_B$ this means that whenever there is a positive probability to move from x to A then the process moves x to A surely. Let us call such a set A a "trap set". Thus condition (2. 3) can be rephrased:

Every "trap set" captures only its own members.

Theorem. *Let $r(n)$ assume the values* 1, ..., N *infinitely often. Let μ be a probability measure which is absolutely continuous with respect to λ and put $d\mu=f\,d\lambda$, $0\leqq f\in L_1(X, \Sigma, \lambda)$. If $P_j(x, A)$ $(1\leqq j\leqq N)$ satisfy* (2. 3) *then for every set A with $\lambda(A)<\infty$*

$$\lim_{n\to\infty}(\mu P_{r(0)}\dots P_{r(n)})(A)=\mu_0(A)$$

where $d\mu_0=f_0\,d\lambda$ and f_0 is the conditional expectation of f on the field $\Sigma'=$ $=\bigcap_{j=1}^{N}\{B: \lambda(B)<\infty, P_j(\cdot, B)=1_B\}$.

Proof. With no loss of generality we may assume that $f\in L_2(X, \Sigma)$. It is well known that the conditional expectation f_0 of f on the field Σ' is equal to the orthogonal projection $P'f$ of f on the subspace L' of $L_2(X, \Sigma, \lambda)$, spanned by Σ'.

By the Amemiya—Ando theorem $P_{r(0)}\dots P_{r(n)}$ and therefore $P^*_{r(n)}\dots P^*_{r(0)}$ converge weakly to the orthogonal projection P' on L'. Thus

$$(\mu P_{r(0)}\dots P_{r(n)})A=\int(P_{r(0)}\dots P_{r(n)}1_A)\,d\mu=$$
$$=\int(P_{r(0)}\dots P_{r(n)}1_A)f\,d\lambda=\int 1_A P^*_{r(n)}\dots P^*_{r(0)}f\,d\lambda$$

tends to $\int_A P'f\,d\lambda=\int_A f_0\,d\lambda=\mu_0(A)$.

Let us consider the case where the only "trap sets", for the subtransition functions $P_j(x, A)$, are the trivial sets. Then if λ is not finite, $K=\{0\}$ and $\lim(\mu P_{r(1)}\dots P_{r(n)})(A)=0$ whenever $\lambda(A)<\infty$. On the other hand if $\lambda(X)<\infty$ then $K=\{\text{constant functions}\}$ and $\lim_{n\to\infty}(\mu P_{r(1)}\dots P_{r(n)})(A)=\lambda(A)\,\mu(X)/\lambda(X)$.

References

[1] I. Amemiya and T. Ando, Convergence of random products of contractions in Hilbert space, *Acta Sci. Math.*, **26** (1965), 239–244.

[2] S. R. Foguel, On order preserving contractions, *Israel J. Math.*, **1** (1963), 54—59.

(Received May 31, 1966)

Minimaxprinzipe für stark gedämpfte Scharen

Von ROLF KÜHNE in Dresden (DDR)*)

O. Einleitung

Es sei $\mathfrak{H}$ ein komplexer [1]) Hilbertraum, versehen mit dem Skalarprodukt (x, y) $(x, y \in \mathfrak{H})$, und es seien A, B, C auf $\mathfrak{H}$ definierte selbstadjungierte beschränkte Operatoren. Wir betrachten die für jede komplexe Zahl λ durch

$$L(\lambda) = \lambda^2 A + \lambda B + C \tag{0. 1}$$

definierte Schar L unter folgender Voraussetzung:

(D) Die Schar L sei *stark gedämpft*, d. h., es gelte

$$(Bx, x)^2 > 4(Ax, x)(Cx, x) \quad \text{für alle} \quad x \in \mathfrak{H},\ x \neq 0.$$

(Zur Terminologie vgl. [1], [2] und [4].)

Die benötigten Voraussetzungen, Definitionen und einfache sich daraus ergebende Folgerungen enthält Abschnitt 2.

Abschnitt 3 beschäftigt sich mit den für L (durch (2. 2), (2. 3)) definierten Funktionalen erster und zweiter Art (sog. verallgemeinerte Rayleigh-Quotienten, s. [1]).

In Abschnitt 4 betrachten wir das Spektrum der Schar L. Mit Hilfe der Funktionale erster und zweiter Art werden für das Folgende wesentliche Teile des Spektrums, das sog. Spektrum erster bzw. zweiter Art, ausgesondert und näher behandelt.

Das eigentliche Ziel der Untersuchungen ist der Inhalt des Abschnittes 5, wo unter gewissen Voraussetzungen an das Spektrum der Schar L Minimaxprizipe für die Bestimmung sog. Eigenwerte erster (bzw. zweiter) Art und zugehöriger Eigenelemente angegeben werden (Sätze 5. 1 und 5. 2). Ergebnisse dieser Art erhielt

---

*) Diese Arbeit ist im wesentlichen Teil einer Dissertation. Herrn Prof. Dr. P. H. MÜLLER, Dresden, bin ich für seine stete Unterstützung, ihm und Herrn Prof. Dr. M. A. NEUMARK, Moskau, außerdem für die Übernahme der Referate zu großem Dank verpflichtet.

[1]) Die folgenden Untersuchungen bleiben mutatis mutandis auch in reellen Räumen gültig. Eine gewisse Ausnahmestellung nehmen dabei reelle Räume der Dimension 2 ein. An entsprechender Stelle wird darauf hingewiesen.

erstmals R. J. DUFFIN [1], [2] unter zusätzlichen Voraussetzungen in endlichdimensionalen Räumen (Näheres s. Ende des Abschnittes 6); die dort benutzten Methoden beruhen aber großenteils auf charakteristischen Eigenschaften endlichdimensionaler Räume und sind deshalb auf den hier betrachteten allgemeinen Fall im wesentlichen nicht übertragbar.

Anwendungen der Minimaxprinzipe auf Scharen speziellen Typs erfolgen in Abschnitt 6 (Sätze 6.1 und 6.2), wobei sich als Spezialfall die Aussagen von [1] bzw. [2] ergeben.

Schließlich sind in Abschnitt 1 einige für unsere Untersuchungen wesentliche Beziehungen zwischen quadratischen Formen, deren Nullkegel nur das Nullelement des Raumes gemeinsam haben, vorangestellt.

Es sei noch erwähnt, daß Teile der vorgelegten Ergebnisse anfangs über Zusammenhänge gewisser Scharen mit speziellen J-selbstadjungierten Operatoren (zur Terminologie s. [4], [5]) gewonnen wurden. Genauer gesagt läßt sich einer Schar L z. B. unter den zusätzlichen Voraussetzungen $A=I$ und $(Bx, x)\geqq 0$, $(Cx, x)\geqq 0$ $(x\in\mathfrak{H})$ in einem geeignet gewählten J-Raum ein J-selbstadjungierter Operator T zuordnen, dessen von 0 verschiedenes Spektrum mit dem von 0 verschiedenen Spektrum der Schar L übereinstimmt (eine solche Linearisierung wurde von M. G. KREIN und H. LANGER in [4] beim Studium der Operatorgleichung $Z^2+BZ+C=0$ betrachtet und ist einer entsprechenden von P. H. MÜLLER ([6], [7]) unter der Voraussetzung $(Cx, x)\leqq 0$ $(x\in\mathfrak{H})$ verwendeten Linearisierung analog; der Verfasser verdankt Herrn Prof. DR. H. LANGER, Dresden, den Hinweis auf solche Zusammenhänge sowie auf die Arbeiten [1] und [2]).

Wie H. LANGER bemerkte, gehört der Operator T genau dann zu der in [5] betrachteten Operatorenklasse, wenn L stark gedämpft ist. Wird zusätzlich gefordert, daß C vollstetig ist, so gelangt man dann unter teilweiser Benutzung der in [5] angegebenen Aussagen und Methoden auch auf diesem Wege der Linearisierung zu Spezialfällen unserer Ergebnisse.

1. Quadratische Formen

Dieser Abschnitt enthält unmittelbare Verallgemeinerungen der Lemmata 1.1 und 1.2 aus [5].

Es sei $\mathfrak{E}$ ein komplexer linearer Raum. Unter einer *hermetischen Bilinearform* auf $\mathfrak{E}$ versteht man eine Abbildung a von $\mathfrak{E}\times\mathfrak{E}$ in die Menge der komplexen Zahlen mit den Eigenschaften

$$a(\lambda_1 x_1+\lambda_2 x_2, y)=\lambda_1 a(x_1, y)+\lambda_2 a(x_2, y)$$
$$a(x, y)=\overline{a(y, x)}$$

für beliebige komplexe λ_1, λ_2 und $x_1, x_2, x, y\in\mathfrak{E}$.

Die durch $a(x, x)$ $(x \in \mathfrak{E})$ auf $\mathfrak{E}$ definierte (reellwertige) Funktion heißt die zu a gehörige *quadratische Form*. Die Menge

$$\mathfrak{a}^0 = \{x \in \mathfrak{E} | a(x, x) = 0\}$$

nennen wir deren *Nullkegel*.

Lemma 1. 1. [2]) *Es seien a und b zwei hermetische Bilinearformen, wobei zwischen den Nullkegeln $\mathfrak{a}^0$ und $\mathfrak{b}^0$ der zugehörigen quadratischen Formen die Bedingung*

$$\mathfrak{a}^0 \cap \mathfrak{b}^0 = \{0\} \tag{1. 1}$$

bestehe. Dann gilt eine der beiden Relationen

$$a(x, x) > 0 \quad \text{für alle} \quad x \in \mathfrak{b}^0 \setminus \{0\} \tag{1. 2}$$

oder

$$a(x, x) < 0 \quad \text{für alle} \quad x \in \mathfrak{b}^0 \setminus \{0\}. \tag{1. 3}$$

Beweis. Angenommen, es gäbe zwei Elemente $x, y \in \mathfrak{b}^0$ mit $a(x, x) < 0$ und $a(y, y) > 0$. Dann ließe sich eine reelle Zahl φ und danach ein reelles $\alpha \neq 0$ so wählen, daß $\operatorname{Re}(e^{i\varphi} b(x, y)) = 0$ und $a(x, x) + 2\alpha \operatorname{Re}(e^{i\varphi} a(x, y)) + \alpha^2 a(y, y) = 0$ gilt. Für $z_0 = x + \alpha e^{i\varphi} y$ folgt daraus $z_0 \in \mathfrak{a}^0 \cap \mathfrak{b}^0$ und wegen $a(x, x) < 0$ und $a(y, y) > 0$ aber auch $z_0 \neq 0$ im Widerspruch zu (1. 1).

Lemma 1. 2. [3]) *a und b seien zwei hermitesche Bilinearformen; dabei sei die quadratische Form $b(x, x)$ $(x \in \mathfrak{E})$ streng indefinit, und es gelte für alle $z \in \mathfrak{b}^0 \setminus \{0\}$ die Ungleichung $a(z, z) > 0$. Dann ist für alle $x \in \mathfrak{E}$ mit $b(x, x) > 0$ und alle $y \in \mathfrak{E}$ mit $b(y, y) < 0$*

$$\frac{a(y, y)}{b(y, y)} < \frac{a(x, x)}{b(x, x)}. \tag{1. 4}$$

Ferner gilt $\mu = \inf_{b(x, x) > 0} \dfrac{a(x, x)}{b(x, x)} > -\infty$, *und es ist*

$$a(z, z) \geqq \mu b(z, z) \qquad (z \in \mathfrak{E}). \tag{1. 5}$$

Beweis. Angenommen, es gäbe zwei Elemente $x_0, y_0 \in \mathfrak{E}$ mit $b(x_0, x_0) = -b(y_0, y_0) = 1$ und $-a(y_0, y_0) \geqq a(x_0, x_0)$. Wählt man dann eine reelle Zahl φ, für die $\operatorname{Re}(e^{i\varphi} b(x_0, y_0)) = 0$ und $\operatorname{Re}(e^{i\varphi} a(x_0, y_0)) \leqq 0$ gilt, so ergeben sich für $z_0 = e^{i\varphi} x_0 + y_0$ die Beziehungen $z_0 \in \mathfrak{b}^0 \setminus \{0\}$ und $a(z_0, z_0) \leqq 0$ im Widerspruch zur Voraussetzung.

---

[2]) Der Inhalt dieses Lemmas ist in [8], Satz 1.1 enthalten.

[3]) Eine entsprechende Aussage unter etwas allgemeineren Voraussetzungen wurde (in Verallgemeinerung von Lemma 1. 2 aus [5]) in Theorem 1. 1 der folgenden Arbeit bewiesen: M. G. Krein und Ju. L. Šmuljan, Über Plus-Operatoren in einem Raum mit indefiniter Metrik, *Mat. issledovanija Akad. Nauk Moldavskoi SSR*, Kišinev **1** (1) (1966), 131—161 [russ.].

Aus (1. 4) folgt $\mu > -\infty$. Schließlich gilt die Ungleichung (1. 5) für alle $z \in \mathfrak{E}$, da sie nach Definition für $b(z, z) > 0$, auf Grund der Voraussetzung für $b(z, z) = 0$ und infolge (1. 4) für $b(z, z) < 0$ erfüllt ist.

Bemerkung. Ist $\mathfrak{E}$ ein reeller Raum, so bleibt Lemma 1. 2 in vollem Umfang gültig, Lemma 1. 1 hingegen nur unter der zusätzlichen Voraussetzung, daß $\mathfrak{E}$ eine von 2 verschiedene Dimension besitzt (vgl. [8], Satz 1. 1).

2. Voraussetzungen und Definitionen

Nach den vorbereitenden Betrachtungen des Abschnittes 1 in beliebigen komplexen linearen Räumen kehren wir zu den in der Einleitung formulierten Voraussetzungen zurück. Es seien also $\mathfrak{H}$ ein komplexer Hilbertraum, L eine durch (0. 1) auf $\mathfrak{H}$ definierte Schar, die der Bedingung (D) der starken Dämpfung genüge.

Eine einfache Konsequenz von (D) enthält

Lemma 2. 1. *Genügt die Schar L der Bedingung* (D), *so gilt eine der folgenden Relationen:*

$$(Bx, x) > 0 \quad \textit{für alle} \quad x \in \mathfrak{H}, x \neq 0, \quad \textit{mit} \quad (Ax, x) = 0$$

oder

$$(Bx, x) < 0 \quad \textit{für alle} \quad x \in \mathfrak{H}, x \neq 0, \quad \textit{mit} \quad (Ax, x) = 0.$$

Beweis. Aus (D) folgt, daß die quadratischen Formen (Bx, x) $(x \in \mathfrak{H})$ und (Ax, x) $(x \in \mathfrak{H})$ die Bedingung (1. 1) erfüllen. Also liefert Lemma 1. 1 die Behauptung.

Für alles Weitere setzen wir nun anstelle von (D) die Gültigkeit der Bedingung

(D$^+$) Die Schar L genüge der Bedingung (D), und es gelte

$$(Bx, x) > 0 \quad \text{für jedes} \quad x \in \mathfrak{H}, x \neq 0, \quad \text{mit} \quad (Ax, x) = 0.$$

voraus, was insofern keine weitere Einschränkung bedeutet, als nach Lemma 2. 1 unter der Voraussetzung (D) entweder L oder $-L$ sogar der Bedingung (D$^+$) genügt.

An dieser Stelle soll der Fall reeller Hilberträume erwähnt werden. Die soeben vorgenommene Ersetzung der Bedingung (D) durch (D$^+$) beruhte auf Lemma 2. 1 und nach dessen Beweis also auf Lemma 1. 1, das im reellen Fall — wie bereits vermerkt — unter der zusätzlichen Voraussetzung $\dim \mathfrak{H} \neq 2$ richtig bleibt. Somit ist alles bisher Gesagte auch für reelle Räume einer von 2 verschiedenen Dimension gültig. Bei $\dim \mathfrak{H} = 2$ ist die Bedingung (D$^+$) i. a. tatsächlich stärker als (D). Dies ist jedoch der einzige hier auftretende Unterschied zwischen reellen und komplexen Räumen. Alles Weitere gilt vollinhaltlich auch im reellen Fall.

Wir setzen nun abkürzend

$$d(x) = \sqrt{(Bx, x)^2 - 4(Ax, x)(Cx, x)}.$$

Nach der Voraussetzung (D^+) gilt dann

$$d(x) > 0 \qquad (x \in \mathfrak{H},\ x \neq 0). \tag{2. 1}$$

Eine entscheidende Rolle spielen im weiteren die Funktionale

$$p(x) = \begin{cases} -\dfrac{1}{2(Ax,x)}[(Bx,x)-d(x)] & x\in\mathfrak{H},\quad (Ax,x)\neq 0 \\ -\dfrac{(Cx,x)}{(Bx,x)} & x\in\mathfrak{H},\quad x\neq 0,\quad (Ax,x)=0, \end{cases} \tag{2. 2}$$

$$s(x) = \begin{cases} -\dfrac{1}{2(Ax,x)}[(Bx,x)+d(x)] & x\in\mathfrak{H},\quad (Ax,x)\neq 0 \\ \infty\ ^{4)} & x\in\mathfrak{H},\quad x\neq 0,\quad (Ax,x)=0. \end{cases} \tag{2. 3}$$

Dabei heiße p *Funktional erster Art* und s *Funktional zweiter Art* von L (zur Definition und Terminologie s. [1]). Offenbar gilt

$$\begin{matrix} p(\varrho x)=p(x) \\ s(\varrho x)=s(x) \end{matrix} \qquad (x\in\mathfrak{H},\ x\neq 0;\ \varrho\neq 0 \ \text{ beliebig komplex}). \tag{2. 4}$$

Wir definieren noch für jede (komplexe) Zahl λ

$$f_\lambda(x)=(\lambda+p(x))(Ax,x)+(Bx,x) \qquad (x\in\mathfrak{H},\ x\neq 0) \tag{2. 5}$$

und fassen nun die unmittelbar aus den Definitionen folgenden Eigenschaften von p, s und f_λ in den folgenden beiden Lemmata zusammen.

Lemma 2. 2. *Für eine beliebige (komplexe) Zahl λ und beliebiges $x\in\mathfrak{H}$, $x\neq 0$ gilt*

$$f_\lambda(x) = \begin{cases} (Bx,x) & \textit{falls} \quad (Ax,x)=0 \\ (Ax,x)(\lambda-s(x)) & \textit{falls} \quad (Ax,x)\neq 0, \end{cases} \tag{2. 6}$$

$$f_\lambda(x)\neq 0 \quad \textit{genau dann, wenn} \quad \lambda\neq s(x), \tag{2. 7}$$

$$(L(\lambda)x,x)=(\lambda-p(x))f_\lambda(x). \tag{2. 8}$$

Beweis. Ist $(Ax,x)\neq 0$, so folgt aus (2. 2) und (2. 3)

$$p(x)+s(x) = -\frac{(Bx,x)}{(Ax,x)}$$

4) Unter ∞ sei im folgenden stets der die komplexe Zahlenebene (im Sinne von Alexandroff) bikompaktifizierende Punkt zu verstehen. Außerdem wollen wir diesen Punkt als reell definieren.

und somit

$$f_\lambda(x) = \left(\lambda - \frac{(Bx, x)}{(Ax, x)} - s(x)\right)(Ax, x) + (Bx, x) = (Ax, x)(\lambda - s(x)),$$

also (2. 6). Im Falle $(Ax, x)=0$ ergibt sich (2. 6) direkt aus (2. 5).

(2. 7) erhält man aus (2. 6) unter Berücksichtigung der Bedingung (D^+).

(2. 6), (2. 2) und (2. 3) ergeben auf bekannte Weise (2. 8).

Lemma 2. 3. *Für ein (komplexes) λ_0 und ein $x \in \mathfrak{H}$, $x \neq 0$, besteht genau dann die Gleichung $(L(\lambda_0)x, x)=0$, wenn eine der Beziehungen $p(x)=\lambda_0$ oder $s(x)=\lambda_0$ gilt. Ferner ist*

$$\frac{d(L(\lambda)x, x)}{d\lambda}\bigg|_{\lambda=p(x)} = f_{p(x)}(x) = 2p(x)(Ax, x) + (Bx, x) > 0 \quad (x \in \mathfrak{H}, x \neq 0), \tag{2.9}$$

$$\frac{d(L(\lambda)x, x)}{d\lambda}\bigg|_{\lambda=s(x)} = 2s(x)(Ax, x) + (Bx, x) < 0 \quad (x \in \mathfrak{H}, (Ax, x) \neq 0) \tag{2.10}$$

und

$$p(x) \neq s(x) \quad (x \in \mathfrak{H}, x \neq 0). \tag{2.11}$$

Beweis. Die erste Behauptung des Lemmas gilt offensichtlich.

Die Beziehungen (2. 9) und (2. 10) ergeben sich im Falle $(Ax, x)\neq 0$ daraus, daß per definitionem $f_{p(x)}(x)=2p(x)(Ax, x)+(Bx, x) = d(x)>0$ und $2s(x)(Ax, x)+ +(Bx, x) = -d(x)<0$ gilt. Ist $(Ax, x)=0$, so folgt aus (D^+) $f_{p(x)}(x)=(Bx, x)>0$.

(2. 11) erhält man aus (2. 9) und (2. 10).

3. Die Funktionale p und s

Die folgenden Lemmata enthalten Eigenschaften von p und s. Dabei beschränken wir uns im wesentlichen auf die Betrachtung des Funktionals p, da sich Aussagen über p unmittelbar auf s übertragen lassen, wie in Lemma 3. 6 gezeigt wird.

Lemma 3. 1. *Das Funktional p ist auf $\mathfrak{H} \setminus \{0\}$ stetig.*

Beweis. Für jedes $x \in \mathfrak{H}$ mit $(Ax, x)\neq 0$ existiert wegen der Stetigkeit von A eine Umgebung von x, für deren Elemente z ebenfalls $(Az, z)\neq 0$ gilt. Die Stetigkeit von p in x ergibt sich dann nach (2. 2) aus der Stetigkeit von A, B und C.

Ist $x \in \mathfrak{H}$, $x \neq 0$ und $(Ax, x)=0$, so gilt wegen (D^+) die Ungleichung $(Bx, x)>0$. Folglich gibt es eine Umgebung U von x mit $(Bz, z)>0$ für alle $z \in U$.

Nun gilt nach (2. 2) für alle $z \in U$

$$[(Bz, z) + d(z)]p(z) = -2(Cz, z)$$

und daher unter Beachtung von $(Bz, z) > 0$ und (2. 1)

$$p(z) = -\frac{2(Cz, z)}{(Bz, z) + d(z)} \qquad (z \in U).$$

Also ist auch dann p in x stetig, w. z. z. w.

Wir definieren für jede reelle Zahl ϱ die Operatoren

$$T_\varrho = \varrho^2 A - C$$

$$R_\varrho = 2\varrho A + B$$

und setzen

$$\mathsf{P} = \{\varrho \mid \varrho \quad \text{reell, für jedes} \quad x \in \mathfrak{H} \quad \text{mit} \quad (R_\varrho x, x) = 0 \quad \text{gilt} \quad (Ax, x) \geqq 0\}.$$

Lemma 3. 2. *Die quadratische Form (Ax, x) $(x \in \mathfrak{H})$ sei streng indefinit. Dann existiert eine reelle Konstante μ mit*

$$(Bx, x) \geqq \mu(Ax, x) \qquad (x \in \mathfrak{H}). \tag{3. 1}$$

Beweis. Wegen (D^+) genügen die quadratischen Formen (Bx, x) $(x \in \mathfrak{H})$ und (Ax, x) $(x \in \mathfrak{H})$ den Voraussetzungen zu Lemma 1. 2. Daher folgt aus (1. 5) mit $\mu = \inf\limits_{(Ax, x) > 0} \frac{(Bx, x)}{(Ax, x)}$ die Beziehung (3. 1).

Folgerung 1. *Für alle $x \in \mathfrak{H}$ mit $(Ax, x) > 0$ gilt*

$$s(x) < -\frac{\mu}{2}.$$

Beweis. Wegen (3. 1) ist für jedes $x \in \mathfrak{H}$ mit $(Ax, x) > 0$

$$s(x) = -\frac{1}{2(Ax, x)}[(Bx, x) + d(x)] \leqq -\frac{\mu}{2} - \frac{d(x)}{2(Ax, x)} < -\frac{\mu}{2}$$

Folgerung 2. *Für jedes reelle $\varrho < -\frac{\mu}{2}$ gilt $\varrho \in \mathsf{P}$.*

Beweis. Es sei $\varrho < -\frac{\mu}{2}$. Für jedes $x \in \mathfrak{H}$ mit $2\varrho(Ax, x) + (Bx, x) = (R_\varrho x, x) = 0$ folgt aus (3. 1)

$$2\left(-\frac{\mu}{2} - \varrho\right)(Ax, x) = -\mu(Ax, x) + (Bx, x) \geqq 0,$$

also

$$(Ax, x) \geqq 0.$$

Somit gilt $\varrho \in \mathsf{P}$.

Lemma 3. 3. *Sei* $\varrho \in \mathrm{P}$. *Dann gilt für jedes* $x \in \mathfrak{H}$ *mit* $(R_\varrho x, x) > 0$ *und jedes* $y \in \mathfrak{H}$ *mit* $(R_\varrho y, y) < 0$ *die Ungleichung*

$$(3.2) \qquad \frac{(T_\varrho y, y)}{(R_\varrho y, y)} < \frac{(T_\varrho x, x)}{(R_\varrho x, x)}.$$

Beweis. Es sei z ein Element aus $\mathfrak{H}$ mit $(R_\varrho z, z) = 0$ und $(T_\varrho z, z) \leqq 0$. Aus $(R_\varrho z, z) = 0$ folgt $(Bz, z)^2 = 4\varrho^2 (Az, z)^2$ und wegen $\varrho \in \mathrm{P}$ $(Az, z) \geqq 0$. Dies liefert unter Benutzung von $\varrho^2(Az, z) - (Cz, z) = (R_\varrho z, z) \leqq 0$ die Beziehung

$$(Bz, z)^2 \leqq 4(Az, z)(Cz, z).$$

Also gilt nach (D^+) $z = 0$. D. h., für die quadratischen Formen $(T_\varrho z, z)$ und $(R_\varrho z, z)$ sind die Voraussetzungen zu Lemma 1. 2 erfüllt. (3. 2) ergibt sich somit aus (1. 3).

Genauere Auskunft über die Wertebereiche von p und s gibt nun

Lemma 3. 4. [5]) *Es sei* $x \in \mathfrak{H}$, $x \neq 0$, *beliebig. Dann gilt*

$$(3.3) \qquad s(y) < p(x), \quad \text{falls} \quad y \in \mathfrak{H} \quad \text{und} \quad (Ay, y) > 0$$

und

$$(3.4) \qquad s(y) > p(x), \quad \text{falls} \quad y \in \mathfrak{H} \quad \text{und} \quad (Ay, y) < 0.$$

Beweis. (1) Wir beweisen zunächst (3. 3). Angenommen, es gäbe zwei von 0 verschiedene Elemente $x, y \in \mathfrak{H}$, so daß

$$(3.5) \qquad (Ay, y) > 0 \quad \text{und} \quad s(y) \geqq p(x)$$

gilt. Wir setzen $\varrho = p(x)$.

Fall 1: Die Form (Ax, x) $(x \in \mathfrak{H})$ sei streng indefinit. Dann gilt nach Folgerung 1 aus Lemma 3. 2 $p(x) \leqq s(y) < -\dfrac{\mu}{2}$, also wegen Folgerung 2 aus Lemma 3. 2 $\varrho = p(x) \in \mathrm{P}$.

Fall 2: Die Form (Ax, x) $(x \in \mathfrak{H})$ sei (semi-)definit. Auf Grund von $(Ay, y) > 0$ ist dann $(Ax, x) \geqq 0$ $(x \in \mathfrak{H})$. Somit gilt wieder $\varrho = p(x) \in \mathrm{P}$.

Folglich läßt sich Lemma 3. 3 anwenden.

Nun ist nach (2. 9)

$$(R_\varrho x, x) = f_\varrho(x) = f_{p(x)}(x) > 0$$

und wegen $(Ay, y) > 0$ und $s(y) \geqq p(x) = \varrho$ nach (2. 10)

$$(R_\varrho y, y) = 2\varrho(Ay, y) + (By, y) \leqq 2s(y)(Ay, y) + (By, y) < 0.$$

[5]) Dieses Lemma ist für $\dim \mathfrak{H} < \infty$ und $(Bx, x) \geqq 0$ $(x \in \mathfrak{H})$ in [1] enthalten. Die hier allgemeiner formulierte Aussage läßt sich (abweichend vom obigen Beweis) im Prinzip auch durch (schrittweise) Reduktion auf diesen Spezialfall nachweisen.

Daher ergibt die Beziehung (3. 2) des Lemmas 3. 3 die Ungleichung

$$\frac{(T_\varrho y, y)}{(R_\varrho y, y)} < \frac{(T_\varrho x, x)}{(R_\varrho x, x)}. \tag{3. 6}$$

Wir benutzen nun die für alle $z \in \mathfrak{H}$ gültige Gleichung

$$(T_\varrho z, z) = \varrho^2(Az, z) - (Cz, z) = 2\varrho^2(Az, z) + \varrho(Bz, z) - \big(L(\varrho)z, z\big) = \tag{3. 7}$$
$$= \varrho(R_\varrho z, z) - \big(L(\varrho)z, z\big).$$

Da nach Definition für $\varrho = p(x)$ die Gleichung $\big(L(\varrho)x, x\big) = 0$ gilt, liefert (3. 7) mit $z = x$

$$(T_\varrho x, x) = \varrho(R_\varrho x, x).$$

(3. 6) vereinfacht sich so zu

$$\varrho(R_\varrho y, y) - (T_\varrho y, y) < 0. \tag{3. 8}$$

(3. 7), mit $z = y$ in (3. 8) eingesetzt, ergibt dann

$$\big(L(\varrho)y, y\big) < 0. \tag{3. 9}$$

Andererseits folgt aber aus (2. 2) und (2. 3) $p(y) - s(y) = \dfrac{d(y)}{(Ay, y)} > 0$, also unter Beachtung von (3. 5)

$$p(y) > s(y) \geqq \varrho.$$

(2. 6) und (2. 8) liefern so zusammen mit (3. 5) die Ungleichung

$$\big(L(\varrho)y, y\big) \geqq 0$$

im Widerspruch zu (3. 9).

(2) (nach [1]) Zum Beweis von (3. 4) betrachten wir die Schar $L^-(\lambda) = \lambda^2 A^- + \lambda B^- + C^-$ mit $A^- = -A$, $B^- = B$, $C^- = -C$, die offenbar wieder der Bedingung (D^+) genügt. Die Anwendung von Teil (1) des Beweises auf L^- liefert dann (3. 4).

Folgerung 1. *Es sei μ eine reelle Zahl, zu der Elemente $x, y \in \mathfrak{H}$ mit $p(x) \leqq \mu \leqq p(y)$ existieren. Dann gilt für alle $z \in \mathfrak{H}$, $z \neq 0$, die Ungleichung*

$$f_\mu(z) > 0.$$

Beweis. Gilt $(Az, z) = 0$, so folgt aus (D^+) $(Bz, z) > 0$ und somit aus (2. 6) die Behauptung.

Ist $(Az, z) \neq 0$, so ergibt sich aus (3. 3) und (3. 4) $\big(\mu - s(z)\big)(Az, z) > 0$, also nach (2. 6) $f_\mu(z) > 0$.

Folgerung 2. *Ist μ eine reelle Zahl mit* $\inf_{x \in \mathfrak{H} \setminus \{0\}} p(x) \leqq \mu \lneqq \sup_{x \in \mathfrak{H} / \{0\}} p(x)$, *so gilt für alle* $z \in \mathfrak{H}$, $z \neq 0$, *die Ungleichung*

$$f_\mu(z) \geqq 0.$$

Der Beweis ist dem vorstehenden Beweis der Folgerung 1 analog.

Aus dem für das Folgende grundlegenden Lemma 3. 4 ergibt sich

Lemma 3. 5. *Es sei μ eine reelle Zahl mit $f_\mu(z) > 0$ für alle $z \in \mathfrak{H}$, $z \neq 0$, und seien x, y Elemente aus $\mathfrak{H}$ mit $p(x) \leqq p(y)$ und $(L(\mu)x, y) = 0$. Dann gilt*

(1) *Aus $p(y) > p(x) \geqq \mu$ folgt $p(x+y) > \mu$.*

(2) *Aus $p(y) = p(x) = \mu$ und $x \neq -y$ folgt $p(x+y) = \mu$.*

(3) *Aus $p(x) < p(y) \leqq \mu$ folgt $p(x+y) < \mu$.*

Beweis. (0) Wegen $f_\mu(z) > 0$ folgt aus (2. 8)

(3. 10) $$\operatorname{sgn} [(L(\mu)z, z)] = \operatorname{sgn} [\mu - p(z)]^{6)} \qquad (z \in \mathfrak{H},\ z \neq 0).$$

Weiter benutzen wir die nach Voraussetzung gültige Gleichung

(3. 11) $$(L(\mu)(x+y), x+y) = (L(\mu)x, x) + (L(\mu)y, y).$$

(1) Aus $p(y) > p(x) \geqq \mu$ erhält man infolge (3. 10) $(L(\mu)x, x) \leqq 0$ und $(L(\mu)y, y) < 0$, daher wegen (3. 11) $(L(\mu)(x+y), x+y) < 0$, also auf Grund von (3. 10) $p(x+y) > \mu$.

Analog ergeben sich die Beweise von (2) und (3).

Folgerung (s. [1]). *Es seien x, y von 0 verschiedene Elemente aus $\mathfrak{H}$, wobei zu x eine reelle Zahl μ mit $L(\mu)x = 0$ und $p(x) = \mu$ existiere. Dann gilt:*

(1) *Aus $p(y) > p(x)$ folgt $p(y+x) > p(x)$.*

(2) *Aus $p(y) = p(x)$ und $y \neq -x$ folgt $p(y+x) = p(x)$.*

(3) *Aus $p(y) < p(x)$ folgt $p(y+x) < p(x)$.*

Beweis. Nach Voraussetzung gilt $(L(\mu)x, y) = 0$. Außerdem ist wegen der Folgerung 1 zu Lemma 3. 4 $f_\mu(z) > 0$ $(z \in \mathfrak{H}, z \neq 0)$. Also läßt sich das Lemma 3. 5 anwenden.

Wir geben nun noch einen Zusammenhang zwischen Funktionalen erster und zweiter Art an, mit deren Hilfe die bisher bewiesenen Aussagen über p auf s übertragen werden können.

---

[6]) sgn a = 1, 0, oder —1 je nachdem $a > 0$, $a = 0$, oder $a < 0$.

Lemma 3. 6. (1) *Es sei $x_0 \in \mathfrak{H}$, $x_0 \neq 0$, beliebig. Mit Hilfe von $\alpha = p(x_0)$ ordnen wir der Schar L eine Schar $\tilde{L}_\alpha$ durch*

$$\tilde{L}_\alpha(\lambda) = \lambda^2 \tilde{A}_\alpha + \lambda \tilde{B}_\alpha + \tilde{C}_\alpha$$

mit $\tilde{A}_\alpha = L(\alpha)$, $\tilde{B}_\alpha = 2\alpha A + B$, $\tilde{C}_\alpha = A$ zu. Dann erfüllt $\tilde{L}_\alpha$ die Bedingung (D^+) *und für die zugehörigen Funktionale $\tilde{p}_\alpha$, $\tilde{s}_\alpha$ gilt*

$$\tilde{p}_\alpha(x) = \frac{1}{s(x) - \alpha} \qquad (x \in \mathfrak{H},\ x \neq 0)\ ^{7)}, \tag{3.12}$$

$$\tilde{s}_\alpha(x) = \frac{1}{p(x) - \alpha} \qquad (x \in \mathfrak{H},\ x \neq 0). \tag{3.13}$$

(2) *Es sei $B \geqq 0$. Dann gelten die in* (1) *formulierten Aussagen mit $\alpha = 0$.*

Beweis. (1) Die Definitionen ergeben unmittelbar die Beziehung

$$(\tilde{B}_\alpha x, x)^2 - 4(\tilde{A}_\alpha x, x)(\tilde{C}_\alpha x, x) = (Bx, x)^2 - 4(Ax, x)(Cx, x) \qquad (x \in \mathfrak{H}). \tag{3.14}$$

Da außerdem aus Lemma 2. 3

$$(\tilde{A}_\alpha x_0, x_0) = (L(p(x_0))x_0, x_0) = 0$$

und

$$(\tilde{B}_\alpha x_0, x_0) = 2p(x_0)(Ax_0, x_0) + (Bx_0, x_0) > 0$$

folgt, ergibt sich auf Grund von (3. 14) und Lemma 2. 1 aus der Bedingung (D^+) für L sofort die Gültigkeit von (D^+) auch für $\tilde{L}_\alpha$.|

Wir betrachten nun die Schar $L_\alpha(\lambda) = \lambda^2 A_\alpha + \lambda B_\alpha + C_\alpha$ mit $A_\alpha = \tilde{C}_\alpha$, $B_\alpha = \tilde{B}_\alpha$, $C_\alpha = \tilde{A}_\alpha$. Wieder folgt die Gültigkeit von (D^+) für L_α unmittelbar aus der Bedingung (D^+) für L.

Für L_α und die zugehörigen Funktionale d_α, p_α, s_α ergibt sich

$$L_\alpha(\lambda) = \lambda^2 A + \lambda(2\alpha A + B) + \alpha^2 A + \alpha B + C = L(\lambda + \alpha),$$

aus (3. 14)

$$d_\alpha(x) = d(x) \qquad (x \in \mathfrak{H})$$

und daher aus (2. 2) und (2. 3)

$$p_\alpha(x) = p(x) - \alpha,\ s_\alpha(x) = s(x) - \alpha \qquad (x \in \mathfrak{H},\ x \neq 0).$$

(a) Nach Lemma 3. 4 ist $s(x) \neq p(x_0) = \alpha$ $(x \in \mathfrak{H},\ x \neq 0)$ und somit $s_\alpha(x) \neq 0$. Daher gilt nach Definition von L_α und auf Grund von Lemma 2. 3 für alle $x \in \mathfrak{H}$, $x \neq 0$, mit $s_\alpha(x) \neq \infty$

$$(\tilde{L}_\alpha(s_\alpha^{-1}(x))x,\ x) = s_\alpha^{-2}(x)(L_\alpha(s_\alpha(x))x,\ x) = 0 \tag{3.15}$$

---

7) Hierbei setzen wir wie üblich $\infty - \alpha = \infty$, $\frac{1}{\infty} = 0$, $\frac{1}{0} = \infty$.

und nach (2. 10)

$$\left.\frac{d(\tilde{L}_\alpha(\lambda)x, x)}{d\lambda}\right|_{\lambda=s_\alpha^{-1}(x)} = 2s_\alpha^{-1}(x)(\tilde{A}_\alpha x, x) + (\tilde{B}_\alpha x, x) = 2s_\alpha(x)(\tilde{L}_\alpha(s_\alpha^{-1}(x))x, x) -$$

$$(3.16) \qquad -2s_\alpha(x)(\tilde{C}_\alpha x, x) - (\tilde{B}_\alpha x, x) = -[2s_\alpha(x)(A_\alpha x, x) + (B_\alpha x, x)] > 0.$$

Entsprechendes gilt mutatis mutandis für $s_\alpha(x) = \infty$. (3. 15) und (3. 16) ergeben so wegen Lemma 2. 3 die Beziehung (3. 12).

(b) Für $x \in \mathfrak{H}$ gelte $p_\alpha(x) \neq 0$. Analog zu Teil (a) erhält man dann

$$(\tilde{L}_\alpha(p_\alpha^{-1}(x))x, x) = 0$$

und

$$\left.\frac{d(\tilde{L}_\alpha(\lambda)x, x)}{d\lambda}\right|_{\lambda=p_\alpha^{-1}(x)} < 0,$$

also nach Lemma 2. 3 die Gleichung (2. 13).

Ist $p_\alpha(x) = 0$, so folgt aus $(L_\alpha(p_\alpha(x))x, x) = 0$ die Beziehung $(\tilde{A}_\alpha x, x) = (C_\alpha x, x) = 0$ und somit $\tilde{s}_\alpha(x) = \infty$. Daher gilt auch dann (3. 13).

(2) Ist $B \geqq 0$, so folgt aus (3. 14) für $\tilde{L}_0$ unmittelbar die Gültigkeit von (D^+). Alles Weitere liefert der Beweis zu (1) mit $\alpha = 0$ (vgl. [1]).

4. Das Spektrum von L

Wegen der Einheitlichkeit der Darstellung setzen wir für alles Folgende $L(\infty) = A$ [8]) und haben damit die Schar L auf allen Punkten der erweiterten komplexen Zahlenebene, die mit $\overline{\mathfrak{C}}$ bezeichnet werden soll, definiert.

Seien nun für einen beliebigen festen Wert $\lambda \in \overline{\mathfrak{C}}$ das Spektrum $\sigma(L(\lambda))$, dessen Teile $\sigma_p(L(\lambda))$, $\sigma_r(L(\lambda))$, $\sigma_c(L(\lambda))$ und die Resolventenmenge $\varrho(L(\lambda))$ des Operators $L(\lambda)$ wie üblich (s. z.B. [10], S. 292) erklärt.

Wir definieren $\varrho(L) = \{\lambda \in \overline{\mathfrak{C}} \mid 0 \in \varrho(L(\lambda))\}$, $\sigma(L) = \{\lambda \in \overline{\mathfrak{C}} \mid 0 \in \sigma(L(\lambda))\}$, $\sigma_p(L) = \{\lambda \in \overline{\mathfrak{C}} \mid 0 \in \sigma_p(L(\lambda))\}$, $\sigma_c(L) = \{\lambda \in \overline{\mathfrak{C}} \mid 0 \in \sigma_c(L(\lambda))\}$, $\sigma_r(L) = \{\lambda \in \overline{\mathfrak{C}} \mid 0 \in \sigma_r(L(\lambda))\}$.

$\varrho(L)$ heiße *Resolventenmenge*, $\sigma(L)$ *Spektrum* und $\sigma_p(L)$ (bzw. $\sigma_r(L)$ und σ (L)) *Punkt-* (bzw. *Residual- und kontinuierliches*) *Spektrum* der Schar L.

Eine Folge (x_n) von Elementen aus $\mathfrak{H}$ nennen wir eine zu $\lambda \in \overline{\mathfrak{C}}$ gehörige *Folge erster* (bzw. *zweiter*) *Art*, wenn sie den folgenden Bedingungen genügt:

$$(4.1) \qquad \|x_n\| = 1 \qquad (n = 1, 2, \cdots),$$

$$(4.2) \qquad \lim_{n\to\infty} p(x_n) = \lambda \qquad \left(\text{bzw. } \lim_{n\to\infty} s(x_n) = \lambda\right),$$

$$(4.3) \qquad L(\lambda)x_n \to 0 \qquad (n \to \infty).$$

[8]) Vgl. Fußnote [4]).

Damit lassen sich weitere Teilmengen des Spektrums folgendermaßen definieren:

$\sigma^{(1)}(L)=\{\lambda\in\overline{\mathfrak{C}}\mid$ zu λ existiert eine Folge erster Art$\}$,

$\sigma_p^{(1)}(L)=\{\lambda\in\overline{\mathfrak{C}}\mid$ es existiert ein $x\in\mathfrak{H}$, $x\neq 0$, mit $p(x)=\lambda$ und $L(\lambda)x=0\}$,

$\sigma_i^{(1)}(L)=\{\lambda\in\sigma^{(1)}(L)\mid$ für jede zu λ gehörige Folge (x_n) erster Art mit $x_n\rightharpoonup x_0$ $(n\to\infty)$ [9]) gilt $x_0\neq 0$ und $p(x_0)=\lambda\}$ [10])

und in Analogie hierzu (mit $s(x)$ anstelle $p(x)$ und „Folge zweiter Art" anstelle „Folge erster Art") die Teilmengen $\sigma^{(2)}(L)$, $\sigma_p^{(2)}(L)$ und $\sigma_i^{(2)}(L)$.

Zum besseren Verständnis der Definitionen von $\sigma_i^{(1)}(L)$ und $\sigma_i^{(2)}(L)$ soll an ein Kriterium von H. Weyl erinnert werden, das folgendermaßen lautet (s. [9], S. 348):

Eine reelle Zahl ν ist genau dann ein isolierter Punkt endlicher Vielfachheit des Spektrums eines selbstadjungierten Operators T, wenn für jede Folge (x_n), $x_n\in\mathfrak{H}$ $(n=1,2,\cdots)$ mit $(T-\nu I)x_n\to 0$, $x_n\rightharpoonup x_0$ $(n\to\infty)$ und $\|x_n\|=1$ $(n=1,2,\cdots)$ gilt $x_0\neq 0$.

Wir nennen nun $\sigma^{(1)}(L)$ bzw. $\sigma^{(2)}(L)$ *Spektrum erster* bzw. *zweiter Art*, auf entsprechende Benennungen der übrigen Mengen soll der Einfachheit halber verzichtet werden.

Gilt $\lambda\in\sigma_p(L)$ (bzw. $\lambda\in\sigma_p^{(1)}(L)$ oder $\lambda\in\sigma_p^{(2)}(L)$), so heißt λ ein *Eigenwert* (bzw. ein *Eigenwert erster* oder *zweiter Art*) und entsprechend jedes $x\in\mathfrak{H}$, $x\neq 0$, mit $L(\lambda)x=0$ ein zu λ gehöriges *Eigenelement* (bzw. ein *Eigenelement erster* oder *zweiter Art*), Schließlich sagt man, der Eigenwert λ habe die endliche *Vielfachheit n*, wenn es zu λ genau n linear unabhängige Eigenelemente gibt.

Die folgenden Lemmata beinhalten einige einfache Eigenschaften der oben eingeführten Mengen.

Lemma 4.1. *Es gilt*

(1) $\sigma_p^{(1)}(L)\subset\sigma^{(1)}(L)$, $\sigma_p^{(2)}(L)\subset\sigma^{(2)}(L)$. (2) $\sigma^{(1)}(L)\cup\sigma^{(2)}(L)$ *ist reell.*

(3) $\sigma^{(1)}(L)\cup\sigma^{(2)}(L)\subset\sigma(L)$. [11]) (4) $\sigma^{(1)}(L)\cap\sigma^{(2)}(L)\subset\{\inf_{x\neq 0} p(x), \sup_{x\neq 0} p(x)\}$.

(5) $\sigma_p^{(1)}(L)\cap\sigma_p^{(2)}(L)=\emptyset$. (6) $\sigma_p^{(1)}(L)\cup\sigma_p^{(2)}(L)=\sigma_p(L)$.

(7) $\sigma_p(L)$ *ist reell.* (8) $\sigma_r(L)=\emptyset$.

---

[9]) Das Symbol $\rightharpoonup$ soll im folgenden stets die schwache Konvergenz in $\mathfrak{H}$, das Symbol $\to$ ausschließlich die starke Konvergenz in $\mathfrak{H}$ bezeichnen.

[10]) Die Forderung $p(x_0)=\lambda$ ist offenbar automatisch erfüllt, falls $\inf_{x\neq 0} p(x)<\lambda<\sup_{x\neq 0} p(x)$ gilt (s. Lemma 3.4).

[11]) Es gilt i.a. $\sigma^{(1)}(L)\cup\sigma^{(2)}(L)\neq\sigma(L)$.

Beweis. Die Aussagen (1), (2) und (3) gelten per definitionem. (4) und (5) ergeben sich aus Lemma 3. 4. (6) folgt daraus, daß für jedes $\lambda \in \sigma_p(L)$ und ein beliebiges zugehöriges Eigenelement x die Gleichung $(L(\lambda)x, x) = 0$, also wegen Lemma 2. 3, $\lambda = p(x)$ oder $\lambda = s(x)$ gilt. (7) ergibt sich aus (1), (2) und (6). (8) Angenommen, es sei $\lambda \in \sigma_r(L)$, d. h. $0 \in \sigma_r(L(\lambda))$. Dann gilt (s. [10], S. 304) $0 \in \sigma_p([L(\lambda)]^*) = \sigma_p(L(\bar{\lambda}))$, d.h. $\bar{\lambda} \in \sigma_p(L)$. Wegen (7) gilt aber $\lambda = \bar{\lambda}$ und deshalb $\lambda \in \sigma_p(L)$ im Widerspruch zu $\lambda \in \sigma_r(L)$.

Lemma 4. 2. *Der Schar L werde auf die in Lemma 3. 6, (1) oder (2), angegebene Weise die Schar $\tilde{L}_\alpha$ zugeordnet. Dann sind die Beziehungen*

$$\lambda \in \sigma^{(2)}(L) \quad \textit{und} \quad \frac{1}{\lambda - \alpha} \in \sigma^{(1)}(\tilde{L}_\alpha),\ ^{12)} \qquad \lambda \in \sigma_p^{(2)}(L) \quad \textit{und} \quad \frac{1}{\lambda - \alpha} \in \sigma_p^{(1)}(\tilde{L}_\alpha),$$

ebenso wie

$$\lambda \in \sigma^{(1)}(L) \quad \textit{und} \quad \frac{1}{\lambda - \alpha} \in \sigma^{(2)}(\tilde{L}_\alpha), \qquad \lambda \in \sigma_p^{(1)}(L) \quad \textit{und} \quad \frac{1}{\lambda - \alpha} \in \sigma_p^{(2)}(\tilde{L}_\alpha),$$

zueinander äquivalent.

Der Beweis des Lemmas folgt unmittelbar aus den entsprechenden Definitionen und sei deshalb dem Leser überlassen.

Im folgenden werden wir nun alle Aussagen ausschließlich für das Spektrum erster Art formulieren. Die entsprechenden Aussagen für das Spektrum zweiter Art ergeben sich dann sofort mittels Lemma 4. 2.

Lemma 4. 3. *Jeder Wert $\lambda \in \sigma_i^{(1)}(L)$ ist ein Eigenwert endlicher Vielfachheit von L, und es gilt $\lambda \in \sigma_p^{(1)}(L)$.*

Beweis. (1) sei $\lambda \in \sigma_i^{(1)}(L)$. Nach Definition existiert eine zu λ gehörige Folge (x_n) erster Art. Wegen $\|x_n\| = 1$ $(n = 1, 2, \cdots)$ besitzt diese Folge in $\mathfrak{H}$ eine schwach konvergente Teilfolge (x_{n_k}); es sei $x_{n_k} \rightharpoonup x_0$ $(k \to \infty)$. Auf Grund von $\lambda \in \sigma_i^{(1)}(L)$ gilt dann $x_0 \neq 0$ und $p(x_0) = \lambda$. Außerdem folgt aus $x_{n_k} \rightharpoonup x_0$ $(k \to \infty)$ die Beziehung $L(\lambda)x_{n_k} \rightharpoonup L(\lambda)x_0$ $(k \to \infty)$. Nach Voraussetzung gilt aber $L(\lambda)x_{n_k} \to 0$ $(k \to \infty)$, somit erhalten wir $L(\lambda)x_0 = 0$, also $\lambda \in \sigma_p^{(1)}(L)$.

(2) Angenommen, λ wäre ein Eigenwert unendlicher Vielfachheit. Dann existierte ein unendliches Orthonormalsystem (e_n) von Eigenelementen der Schar L zum Eigenwert λ, das offenbar eine zu λ gehörige Folge erster Art ist. Bekanntlich gilt aber $e_n \rightharpoonup 0$ $(n \to \infty)$ im Widerspruch zu $\lambda \in \sigma_i^{(1)}(L)$.

Lemma 4. 4. *Ist λ ein Häufungspunkt der Menge $\sigma^{(1)}(L)$, so gilt*

$$\lambda \in \sigma^{(1)}(L) \setminus \sigma_i^{(1)}(L).$$

---

[12]) Vgl. Fußnote [7]).

Beweis. (1) Sei $\lambda \neq \infty$. Nach Voraussetzung existiert eine Folge (λ_n), $\lambda_n \in \sigma^{(1)}(L)$ $(n=1, 2, \cdots)$, mit $\lambda_n \neq \lambda$ und

(4. 4) $$\lim_{n\to\infty} \lambda_n = \lambda.$$

Dann gibt es wegen $\lambda_n \in \sigma^{(1)}(L)$ $(n=1, 2, \cdots)$ zu jeder natürlichen Zahl n eine Folge $(x_m^{(n)})$ aus $\mathfrak{H}$, so daß für die Beziehungen

(4. 5) $$\|L(\lambda_n)x_m^{(n)}\| < \frac{1}{m}|\lambda_m - \lambda|, \qquad |p(x_m^{(n)}) - \lambda_n| < \frac{1}{m} \quad \text{und} \quad \|x_m^{(n)}\| = 1$$

für alle $n, m = 1, 2, \cdots$ gelten.

Setzen wir $y_n = x_n^{(n)}$ $(n=1, 2, \cdots)$, so folgt

$$|p(y_n) - \lambda| \leqq |p(y_n) - \lambda_n| + |\lambda_n - \lambda| < \frac{1}{n} + |\lambda_n - \lambda|,$$

also wegen (4. 4)

$$\lim_{n\to\infty} p(y_n) = \lambda$$

und unter Beachtung von $\|y_n\| = 1$ $(n=1, 2, \cdots)$

$$\|L(\lambda)y_n\| \leqq \|L(\lambda_n)y_n\| + \|(L(\lambda_n) - L(\lambda))y_n\| < \frac{1}{n}|\lambda_n - \lambda| + |\lambda_n^2 - \lambda^2|\,\|A\| + |\lambda_n - \lambda|\,\|B\|,$$

also wegen (4. 4)

(4. 6) $$L(\lambda)y_n \to 0 \qquad (n \to \infty).$$

Daher gilt $\lambda \in \sigma^{(1)}(L)$, wobei (y_n) eine zu λ gehörige Folge erster Art ist. Diese enthält eine schwach konvergente Teilfolge (y_{n_k}), $y_{n_k} \rightharpoonup y$ $(k \to \infty)$, die offenbar wieder eine zu λ gehörige Folge erster Art ist.

Angenommen, es wäre $\lambda \in \sigma_i^{(1)}(L)$. Dann gilt $y \neq 0$ und $p(y) = \lambda$. Aus $y_{n_k} \rightharpoonup y$ $(k \to \infty)$ ergibt sich $L(\lambda)y_{n_k} \rightharpoonup L(\lambda)y$ $(k \to \infty)$ und somit aus (4. 6)

(4. 7) $$L(\lambda)y = 0.$$

Weiter gilt

$$\frac{1}{\lambda_n - \lambda}(L(\lambda_n) - L(\lambda)) = \frac{1}{\lambda_n - \lambda}[(\lambda_n^2 - \lambda^2)A + (\lambda_n - \lambda)B] =$$
$$= (\lambda_n + \lambda)A + B \qquad (n = 1, 2, \cdots).$$

Hieraus folgt

$$([(\lambda_n + \lambda)A + B]y_n, y) = \frac{1}{\lambda_n - \lambda}[(L(\lambda_n)y_n, y) - (L(\lambda)y_n, y)] \qquad (n = 1, 2, \ldots).$$

Unter Beachtung der nach (4. 7) gültigen Beziehung $(L(\lambda)y_n, y) = (y_n, L(\lambda)y) = 0$ erhält man so auf Grund von (4. 5)

(4. 8) $$|([\lambda_{n_k} + \lambda)A + B]y_{n_k}, y)| \leqq \frac{1}{|\lambda_{n_k} - \lambda|}\|L(\lambda_{n_k})y_{n_k}\|\,\|y\| < \frac{1}{n_k}\|y\|.$$

Nun gilt aber wegen $y_{n_k} \rightharpoonup y$ $(k \to \infty)$ und $\lim\limits_{k\to\infty} \lambda_{n_k} = \lambda$

$$([(\lambda_{n_k} + \lambda)A + B]y_{n_k}, y) \to ([2\lambda A + B]y, y).$$

Zusammen mit (4. 8) ergibt sich dann

$$f_\lambda(y) = ([2\lambda A + B]y, y) = 0.$$

Dies steht im Widerspruch dazu, daß für y als Eigenelement erster Art zu λ nach (2.9) $f_\lambda(y) > 0$ gilt.

(2) Sei $\lambda = \infty$. Wird $\tilde{L}_\alpha$ wie in Lemma 3. 6 definiert, so ist nach Voraussetzung wegen Lemma 4. 2 0 ein Häufungspunkt der Menge $\sigma^{(2)}(\tilde{L}_\alpha)$. Wie in (1) zeigt man dann $0 \in \sigma^{(2)}(\tilde{L}_\alpha)$. Aus Lemma 4. 2 ergibt sich so $\infty \in \sigma^{(1)}(L)$. Offenbar gilt aber $\infty \notin \sigma_i^{(1)}(L)$, da das Funktional $p(x)$ per definitionem auf $\mathfrak{H} \setminus \{0\}$ nur endliche Werte annimmt.

Folgerung. *Jedes* $\lambda \in \sigma_i^{(1)}(L)$ *ist ein isolierter Punkt von* $\sigma^{(1)}(L)$.

5. Minimaxprinzipe

Lemma 5. 1 (s. [1]). *Sind* $\lambda_1 < \lambda_2 < \cdots < \lambda_n$ *Eigenwerte erster Art von* L *und* $x_1, x_2, \cdots, x_n$ *zugehörige Eigenelemente* [13]), *so gilt*

$$\lambda_1 < p(x_1 + x_2 + \cdots + x_n) < \lambda_n.$$

Folgerung 1. *Es sei* $\mathfrak{S}_i = \{x \in \mathfrak{H} \mid L(\lambda_i)x = 0\}$ $(i = 1, 2, \cdots, n)$. *Dann gilt für jedes* $x \in \mathfrak{S}_1 + \mathfrak{S}_2 + \cdots + \mathfrak{S}_n$, $x \neq 0$,

$$\lambda_1 \leqq p(x) \leqq \lambda_n.$$

Folgerung 2. *Die Elemente* $x_1, x_2, \cdots, x_n$ *sind linear unabhängig.*

Den Beweis des Lemmas und seiner Folgerungen erhält man wie in [1].

Für das Folgende werde mit $\mathfrak{L}_i$ $(i = 0, 1, \cdots)$ die Gesamtheit aller der Teilräume von $\mathfrak{H}$ bezeichnet, deren orthogonales Komplement in $\mathfrak{H}$ ein i-dimensionaler Teilraum ist. Wir nennen die reelle Zahl

$$k_i = \sup_{\mathfrak{L} \in \mathfrak{L}_i} \inf_{x \in \mathfrak{L} \setminus \{0\}} p(x) \qquad (i = 0, 1, \cdots)$$

den *i-ten Minimum-Maximum-Wert* (erster Art) von L.

Das wesentliche Ergebnis dieses Kapitels soll nun darin bestehen, unter geeigneten Voraussetzungen die Eigenwerte erster Art von L als Minimum-Maximum-Werte erster Art zu charakterisieren. Zuvor beweisen wir einige vorbereitende Lemmata.

[13]) Man beachte, daß nach Definition und wegen Lemma 4. 1, (5), für jedes Eigenelement x zu einem Eigenwert λ erster Art gilt $p(x) = \lambda$.

Lemma 5. 2. *Zu jedem $(i+1)$-dimensionalen Teilraum $\mathfrak{B}_{i+1} \subset \mathfrak{H}$ $(i=0, 1, \cdots)$ existiert ein Element $y_{i+1} \in \mathfrak{B}_{i+1}$ mit*

$$p(y_{i+1}) \geqq k_i.$$

Beweis. Angenommen, für alle $y \in \mathfrak{B}_{i+1}$, $y \neq 0$, gelte

$$p(y) < k_i. \tag{5. 1}$$

Wir betrachten das Supremum $s_{i+1} = \sup\limits_{x \in \mathfrak{B}_{i+1} \setminus \{0\}} p(x)$. Da $p(x)$ nach Lemma 3. 1 auf der Einheitskugel $\mathfrak{K}_{i+1}$ von $\mathfrak{B}_{i+1}$ stetig ist, existiert ein Element $x_{i+1} \in \mathfrak{K}_{i+1}$ mit $p(x_{i+1}) = \sup\limits_{x \in \mathfrak{K}_{i+1}} p(x) = s_{i+1}$. Auf Grund von (5. 1) folgt so

$$s_{i+1} = p(x_{i+1}) < k_i. \tag{5. 2}$$

Bekanntlich gibt es aber zu jedem $\mathfrak{L} \in \mathfrak{L}_i$ ein $y_{\mathfrak{L}} \in \mathfrak{L} \cap \mathfrak{B}_{i+1}$ mit $y_{\mathfrak{L}} \neq 0$ (s. z.B. [9], S. 223). Also gilt für jedes $\mathfrak{L} \in \mathfrak{L}_i$ $\inf\limits_{x \in \mathfrak{L} \setminus \{0\}} p(x) \leqq s_{i+1}$ und daher

$$k_i = \sup_{\mathfrak{L} \in \mathfrak{L}_i} \inf_{x \in \mathfrak{L} \setminus \{0\}} p(x) \leqq s_{i+1},$$

im Widerspruch zu (5. 2).

Lemma 5. 3. *Es seien $x_0, x_1, \cdots, x_n$ linear unabhängige Eigenelemente erster Art von L und $\lambda_0 \leqq \lambda_1 \leqq \cdots \leqq \lambda_n$ die zugehörigen Eigenwerte. Ferner sei μ eine reelle Zahl mit den Eigenschaften $\lambda_n \leqq \mu$ und $f_\mu(x) > 0$ für alle $x \in \mathfrak{H}$, $x \neq 0$. Dann gilt für*

$$\mathfrak{H}_{n+1} = \{x \in \mathfrak{H} \mid ([(\mu+\lambda_i)A + B]x_i, x) = 0 \quad (i=0, 1, \cdots, n)\}.$$

(1) *die Elemente $y_i = [(\mu+\lambda_i)A + B]x_i$ $(i=0, 1, \cdots, n)$ sind linear unabhängig, d.h., es ist $\mathfrak{H}_{n+1} \in \mathfrak{L}_{n+1}$;*

(2) *für jedes $z \in \mathfrak{H}_{n+1}$, $z \neq 0$, sind $x_0, x_1, \cdots, x_n, z$ linear unabhängig;*

(3) *es ist*

$$\mathfrak{H} = \mathfrak{X}_{n+1} \dotplus \mathfrak{H}_{n+1} \text{ [14]}, \tag{5. 3}$$

wobei $\mathfrak{X}_{n+1}$ die lineare Hülle der Elemente $x_0, x_1, \cdots, x_n$ bezeichnet.

Beweis. Wir vermerken für das Folgende die Beziehung

$$L(\mu)x_i = [L(\mu) - L(\lambda_i)]x_i = (\mu - \lambda_i)[(\mu+\lambda_i)A + B]x_i \tag{5. 4}$$

und definieren $J_n' = \{i \mid i=0, 1, \cdots, n \text{ und } \lambda_i \neq \mu\}$ bzw. $J_n = \{i \mid i=0, 1, \cdots, n \text{ und } \lambda_i = \mu\}$.

(1) Für die komplexen Zahlen α_i $(i=0, 1, \cdots, n)$ gelte

$$\sum_{i=0}^{n} \alpha_i y_i = 0. \tag{5. 5}$$

---

[14]) Durch das Symbol $\dotplus$ seien im folgenden direkte Summen gekennzeichnet.

Wir setzen $z_1 = \sum_{i \in J'_n} \frac{\alpha_i}{\mu - \lambda_i} x_i$ und $z_2 = \sum_{i \in J_n} \alpha_i x_i$. Offenbar gilt nach (5. 4) $L(\mu) z_1 = \sum_{i \in J'_n} \alpha_i y_i$ und per definitionem $(2\mu A + B) z_2 = \sum_{i \in J_n} \alpha_i y_i$. (5. 5) erhält so die Gestalt

$$L(\mu) z_1 = -(2\mu A + B) z_2. \tag{5. 6}$$

Angenommen, es sei $z_2 \neq 0$. Dann ergäbe sich aus Punkt (2) der Folgerung von Lemma 3. 5 $p(z_2) = \mu$. Da weiter nach Definition $L(\mu) z_2 = 0$ gilt, folgte somit aus (5. 6)

$$f_\mu(z_2) = ([(\mu + p(z_2))A + B] z_2, z_2) = ([2\mu A + B] z_2, z_2) =$$
$$= -(L(\mu) z_1, z_2) = -(z_1, L(\mu) z_2) = 0.$$

Dies steht im Widerspruch zur Voraussetzung. Also gilt $\sum_{i \in J_n} \alpha_i x_i = z_2 = 0$ und daher wegen der linearen Unabhängigkeit der x_i

$$\alpha_i = 0 \quad \text{für alle} \quad i \in J_n. \tag{5. 7}$$

Außerdem erhält man aus 5. 6 $L(\mu) z_1 = 0$.

Wäre nun $z_1 \neq 0$, so ergäbe sich nach Voraussetzung $f_\mu(z_1) > 0$ und daher aus (2. 8) $p(z_1) = \mu$; z_1 wäre also ein Eigenelement erster Art zu μ. Dann sind aber auf Grund der Folgerung 2 zu Lemma 5. 1 die Elemente z_1 und x_i $(i \in J'_n)$ linear unabhängig im Widerspruch zur Definition von z_1.

Es gilt also $z_1 = 0$. Dies ergibt wegen der linearen Unabhängigkeit der x_i $\alpha_i = 0$ für alle $i \in J'_n$ und zusammen mit (5. 7) schließlich $\alpha_i = 0$ für alle $i = 0, 1, \cdots, n$, w. z. z. w.

(2) Angenommen, es sei $z = \sum_{i=0}^{n} \alpha_i x_i$. Dann folgt aus (5. 4)

$$(L(\mu) z, z) = \sum_{i=0}^{n} \alpha_i (\mu - \lambda_i)([(\mu + \lambda_i) A + B] x_i, z) = 0.$$

(2. 8) liefert so zusammen mit $f_\mu(z) > 0$ die Beziehung $p(z) = \mu$. Hieraus ergibt sich $\alpha_i = 0$ für alle $i \in J'_n$; andernfalls wäre nämlich nach Lemma 5. 1 $\mu = p(z) = p\left(\sum_{i=0}^{n} \alpha_i x_i\right) < \lambda_n$ im Widerspruch zur Voraussetzung $\lambda_n \leq \mu$.

Daher gilt $z = \sum_{i \in J_n} \alpha_i x_i$. Aus $z \in \mathfrak{H}_{n+1}$ erhält man aber wegen $p(z) = \mu$

$$f_\mu(z) = ([2\mu A + B] z, z) = \left(\sum_{i \in J_n} \alpha_i ([(\mu + \lambda_i) A + B] x_i, z)\right) = 0$$

im Widerspruch zu $f_\mu(z) > 0$.

(3) Wir betrachten die — nach (2) direkte — Summe $\mathfrak{H}' = \mathfrak{X}_{n+1} \dotplus \mathfrak{H}_{n+1}$. $\mathfrak{H}'$ ist dann infolge der endlichen Dimension von $\mathfrak{X}_{n+1}$ ein abgeschlossener Teilraum von $\mathfrak{H}$ (s. [10]).

Für jedes $i=0, 1, \cdots, n$ bezeichne $[x_i]$ die lineare Hülle des Elementes x_i. Wegen (2) existiert ein $z_0 \in [x_0] + \mathfrak{H}_{n+1}$, $\|z_0\| = 1$, mit $(z_0, \mathfrak{H}_{n+1}) = \{0\}$. Entsprechend findet man ein

$$z_1 \in [x_1] \dot{+} [x_0] \dot{+} \mathfrak{H}_{n+1}, \; \|z_1\| = 1, \quad \text{mit} \quad (z_1, [x_0] \dot{+} \mathfrak{H}_{n+1}) = \{0\}.$$

$(n+1)$-malige Anwendung dieses Verfahrens liefert ein in $\mathfrak{H}'$ zu $\mathfrak{H}_{n+1}$ orthogonales Orthonormalsystem $z_0, z_1, \cdots, z_n$; $\mathfrak{H}_{n+1}$ besitzt also in $\mathfrak{H}'$ einen Defekt $\geqq n+1$. Da aber $\mathfrak{H}_{n+1}$ in $\mathfrak{H}$ den Defekt $n+1$ hat, folgt

$$\mathfrak{H} = \mathfrak{H}' = \mathfrak{X}_{n+1} \dot{+} \mathfrak{H}_{n+1}, \qquad \text{w.z.z.w.}$$

Wir kommen nun zur Formulierung eines Minimaxprinzips.

Satz 5. 1. (Minimum-Maximum-Prinzip.) *Es sei* $\alpha = \inf_{x \in \mathfrak{H} \setminus \{0\}} p(x) \neq \infty$, *und es existiere eine reelle Konstante* $\beta > \alpha$ *mit der Eigenschaft* $\sigma^{(1)}(L) \cap [\alpha, \beta) \subset \sigma_i^{(1)}(L)$. *Dann gilt*

(1) *Jeder Minimum-Maximum-Wert* $k_n \in [\alpha, \beta)$ *ist ein Eigenwert erster Art und endlicher Vielfachheit von* L.

(2) *Jeder Punkt* $\lambda \in \sigma^{(1)}(L) \cap [\alpha, \beta)$ *ist ein Eigenwert erster Art von endlicher Vielfachheit und tritt in der Menge der* k_n $(n=0, 1, \cdots)$ *seiner Vielfachheit entsprechend oft auf.*

(3) *Es sei* $k_n \in [\alpha, \beta)$. *Dann existiert ein System von* $n+1$ *linear unabhängigen Eigenelementen* $x_0, x_1, \cdots, x_n$ *mit den zugehörigen Eigenwerten* $\lambda_0 = k_0, \lambda_1 = k_1, \cdots, \cdots, \lambda_n = k_n$.[15]) *Definieren wir* $\mathfrak{H}_n = \{x \in \mathfrak{H} \mid ([k_n + \lambda_j)A + B]x_j, x) = 0 \; (j=0, \cdots, n-1)\}$, *so gilt überdies*

(3_1) $k_n = \min_{x \in \mathfrak{H}_n \setminus \{0\}} p(x)$.

(3_2) *Jedes Element* $y \in \mathfrak{H}_n$, $y \neq 0$, *mit* $p(y) = k_n$ *ist ein zu* k_n *gehöriges Eigenelement und ist außerdem linear unabhängig von den Elementen* $x_0, x_1, \cdots, x_{n-1}$.

Beweis. Zu (3) (vollständige Induktion). Es sei $k_l \in [\alpha, \beta)$. Ferner sei die Aussage (3) des Satzes für $n = l-1$ richtig. Dann können wir uns beim Beweis von (3) auf den Beweis von (3_1) und (3_2) beschränken, da aus diesen Beziehungen unmittelbar die Existenz eines zu k_l gehörigen Eigenelementes x_l folgt, das von den nach Induktionsvoraussetzung existierenden Eigenelementen $x_0, x_1, \cdots, x_{l-1}$ linear unabhängig ist.

(a) Wir zeigen als erstes

$$k_l = \inf_{x \in \mathfrak{H} \setminus \{0\}} p(x). \tag{5.8}$$

(a_1) Es sei zunächst $k_{l-1} < k_l$. Offenbar existiert dann ein $y_0 \in \mathfrak{H}$ mit $p(y_0) \leqq k_l$ und außerdem nach Lemma 5. 2 ein $y_1 \in \mathfrak{H}$ mit $k_l \leqq p(y_1)$. Daher sind wegen der

[15]) Es sei daran erinnert, daß nach Definition $\alpha = k_0 \leqq k_1 \leqq \ldots$ gilt.

Folgerung 1 aus Lemma 3. 4 für $\mu=k_l$ die Voraussetzungen zu Lemma 3. 5 und mit $x_0, x_1, \cdots, x_{l-1}$ und $\mu=k_l$ zu Lemma 5. 3 erfüllt. Diese beiden Lemmata sollen im folgenden angewendet werden.

Wir nehmen an, es gäbe ein $z_l\in\mathfrak{H}_l$, $z_l\neq 0$, mit $p(z_l)<k_l$. Ist dann x ein (beliebiges) Element aus $\mathfrak{H}$ der Gestalt $x=\sum_{j=0}^{l-1}\alpha_j x_j$, so gilt wegen $z_l\in\mathfrak{H}_l$ unter Benutzung der auch hier gültigen Beziehung (5. 4)

$$(L(k_l)x, z_l) = \sum_{j=0}^{l-1}\alpha_j(k_l-\lambda_j)([(k_l+\lambda_j)A+B]x_j, z_l) = 0.$$

Weiter ist wegen der Folgerung 1 aus Lemma 5. 1

$$p(x) = p\left(\sum_{j=0}^{l-1}\alpha_j x_j\right) \leqq \lambda_{l-1} = k_{l-1} < k_l$$

und nach Annahme $p(z_l)<k_l$. Somit liefert Lemma 3. 5, (3),

$$p(x+z_l)<k_l.$$

Bezeichnet $\mathfrak{Z}_l$ die lineare Hülle der Elemente $x_0, x_1, \cdots, x_{l-1}, z_l$, so gilt also

$$(5.\,9) \qquad p(y)<k_l \quad \text{für alle} \quad y\in\mathfrak{Z}_l,\, y\neq 0.$$

Auf Grund von Lemma 5. 3, (2), ist aber $\dim\mathfrak{Z}_l=l+1$, daher ergibt sich aus Lemma 5. 2 die Existenz eines Elementes $y_{l+1}\in\mathfrak{Z}_l$ mit $p(y_{l+1})\geqq k_l$. Widerspruch zu (5. 9).

Damit ist die Ungleichung

$$(5.\,10) \qquad k_l \leqq \inf_{x\in\mathfrak{H}_l\setminus\{0\}} p(x)$$

bewiesen.

Andererseits ist infolge Lemma 5. 3, (1), $\mathfrak{H}_1\in\mathfrak{L}_1$. Daher gilt

$$(5.\,11) \qquad \inf_{x\in\mathfrak{H}_l\setminus\{0\}} p(x) \leqq \sup_{\mathfrak{L}\in\mathfrak{L}_l}\inf_{x\in\mathfrak{L}\setminus\{0\}} p(x) = k_l.$$

(5. 10) und (5. 11) ergeben (5. 8).

(a$_2$) Es sei $k_{l-1}=k_l$. Auch dann gilt offenbar die Ungleichung (5. 11). Außerdem ist nach Definition $\mathfrak{H}_l\subset\mathfrak{H}_{l-1}$, somit erhält man

$$(5.\,12) \qquad \inf_{x\in\mathfrak{H}_l\setminus\{0\}} p(x) \geqq \inf_{x\in\mathfrak{H}_{l-1}\setminus\{0\}} p(x) = k_{l-1} = k_l.$$

Wieder liefern (5. 11) und (5. 12) die Gleichung (5. 8).

(b) Wir beweisen nun die Existenz eines Elementes $x_l\in\mathfrak{H}_l$, $x_l\neq 0$, mit $p(x_l)=k_l$ und $L(x_l)=0$.

Da nach (5. 8) für alle $x \in \mathfrak{H}_l$, $x \neq 0$, die Ungleichung $k_l - p(x) \leqq 0$ gilt, ergibt sich aus (2. 8) zusammen mit Folgerung 2 von Lemma 3. 4

(5. 13) $$(L(k_l)x, x) \leqq 0 \qquad (x \in \mathfrak{H}_l).$$

Wegen (5. 8) existiert eine Folge (x_n^l), $x_n^l \in \mathfrak{H}_l$ $(n = 1, 2, \cdots)$, mit $\|x_n^l\| = 1$ und $\lim_{n\to\infty} p(x_n^l) = k_l$. Da dann unter Beachtung von $(L(p(x_n^l))x_n^l, x_n^l) = 0$

$$(L(k_l)x_n^l, x_n^l) = ([L(k_l) - L(p(x_n^l))]x_n^l, x_n^l) \leqq$$
$$\leqq |k_l^2 - p^2(x_n^l)| \, \|A\| + |k_l - p(x_n^l)| \, \|B\| \qquad (n = 1, 2, \cdots)$$

gilt, folgt

(5. 14) $$\lim_{n\to\infty} (L(k_l)x_n^l, x_n^l) = 0.$$

Wir definieren mit P_l die orthogonale Projektion von $\mathfrak{H}$ auf $\mathfrak{H}_l$. Die infolge (5. 13) für die Bilinearform $(P_l L(k_l)x, y)$ $(x, y \in \mathfrak{H}_l)$ gültige Schwarzsche Ungleichung liefert

(5. 15) $$\|P_l L(k_l)x\|^4 = (P_l L(k_l)x, P_l L(k\,)x)^2 \leqq |(P_l L(k_l)x, x)| \, |([P_l L(k_l)]^2 x, P_l L(k_l)x)| \leqq$$
$$\leqq |(L(k_l)x, x)| \, \|P_l L(k_l)\|^3 \|x\|^2 \qquad (x \in \mathfrak{H}_1).$$

Hieraus folgt für (x_n^l) nach (5. 14)

(5. 16) $$P_l L(k_l)x_n^l \to 0 \qquad (n \to \infty).$$

Per definitionem ist nun $(I - P_l)(\mathfrak{H})$ der von den Elementen $y_j = [(k_l + \lambda_j)A + B]x_j$ $(j = 0, \cdots, l-1)$ aufgespannte Teilraum; es gilt also $\dim (I - P_l)(\mathfrak{H}) = l < \infty$. Somit existiert wegen der Kompaktheit beschränkter Mengen in $(I - P_l)\mathfrak{H}$ eine Teilfolge $(x_{n_k}^l)$, so daß die Folge $(L(k_l)x_{n_k}^l)$ auf Grund von

$$L(k_l)x_{n_k}^l = P_l L(k_l)x_{n_k}^l + (I - P_l)L(k_l)x_{n_k}^l$$

und nach (5. 16) konvergiert.

Wir zeigen jetzt $L(k_l)x_{n_k}^l \to 0$ $(k \to \infty)$. Für $P_l = I$ folgt dies sofort aus (5. 16). Im Falle $P_l \neq I$ gilt $\lambda_0 = p(x_0) \leqq k_l$, außerdem existiert nach (5. 8) ein $z_0 \in \mathfrak{H}_l$ mit $k_l \leqq p(z_0)$. Auf Grund der Folgerung 1 aus Lemma 3. 4 ist also Lemma 5. 3 mit $\mu = k_l$ anwendbar und liefert entsprechend (3) eine Zerlegung jedes Elementes $y \in \mathfrak{H}$ in die Summe

$$y = y' + y''$$

mit $y' = \sum_{i=0}^{l-1} \alpha_i x_i$ und $y'' \in \mathfrak{H}_l$.

Da wegen (5. 4) und $x_{n_k}^l \in \mathfrak{H}_l$ $(k = 1, 2, \cdots)$

$$(L(k_l)x_{n_k}^l, y') = \left(x_{n_k}^l, L(k_l)\left(\sum_{i=0}^{l-1} \alpha_i x_i\right)\right) =$$
$$= \left(x_{n_k}^l, \sum_{i=0}^{l-1} \alpha_i (k_l - \lambda_i)[(k_l + \lambda_i)A + B]x_i\right) = 0 \qquad (k = 1, 2, \cdots)$$

gilt, ergibt sich so für jedes $y \in \mathfrak{H}$ die Beziehung

(5.17) $\quad \big(L(k_l)x^l_{n_k}, y\big) = \big(L(k_l)x^l_{n_k}, y''\big) = \big(P_l L(k_l)x^l_{n_k}, y''\big) \qquad (k = 1, 2, \cdots).$

Also folgt aus (5. 16) und (5. 17) $L(k_l)x^l_{n_k} \rightharpoonup 0$ $(k \to \infty)$ und deshalb

$$L(k_l)x^l_{n_k} \to 0 \qquad (k \to \infty),$$

da die Folge $\big(L(k_l)x^l_{n_k}\big)$ nach Konstruktion stark konvergiert. Somit ist $(x^l_{n_k})$ eine zu k_l gehörige Folge erster Art; demnach gilt $k_l \in \sigma^{(1)}(L)$ und wegen $k_l \in [\alpha, \beta)$ auf Grund der Voraussetzungen des Satzes schließlich

(5. 18) $$k_l \in \sigma_i^{(1)}(L).$$

Wie in Teil (1) des Beweises zu Lemma 4. 3 erhalten wir nun durch Übergang zu einer Teilfolge (y_j), $y_j = x^l_{n_{k_j}}$ $(j = 1, 2, \cdots)$, eine zu k_l gehörige Folge erster Art mit $y_j \rightharpoonup x_l \neq 0$ $(j \to \infty)$. Dann gilt

$$L(k_l)x_l = 0 \quad \text{und} \quad p(x_l) = k_l,$$

und wegen $y_j \in \mathfrak{H}_l$, $j = 1, 2, \cdots$, ist $x_l \in \mathfrak{H}_l$.

(c) Die Beziehung (3_1) folgt nun unmittelbar aus den in (a) und (b) bewiesenen Aussagen. Ferner folgt für jedes $y \in \mathfrak{H}_l$ mit $p(y) = k_l$ aus (5. 15) $P_l L(k_l)y = 0$ und entsprechend aus (5. 17) $L(k_l)y = 0$. Die lineare Unabhängigkeit der Elemente $x_0, x_1, \cdots, x_{l-1}, y$ ergibt sich aus Lemma 5. 3, (2). Also gilt auch (3_2).

(d) Zur Vervollständigung unseres Induktionsbeweises bleibt noch die Richtigkeit von (3) für $n = 0$ zu zeigen. Wegen $\mathfrak{H} = \mathfrak{H}_0$ gilt aber $k_0 = \inf_{x = \mathfrak{H}_0 \setminus \{0\}} p(x)$. Wie in (b) läßt sich dann die Existenz eines $x_0 \in \mathfrak{H}_0$, $x_0 \neq 0$, mit $p(x_0) = k_0$ beweisen. Also gilt (3_1) für $n = 0$. Der Beweis für (3_2) im Falle $n = 0$ erfolgt wie in (c).

Zu (1). Nach (5. 18) und Lemma 4. 3 gilt (1).

Zu (2). (a) Es sei $\dim \mathfrak{H} = m < \infty$. Dann gilt $\sigma^{(1)}(L) = \sigma_i^{(1)}(L) = \sigma_p^{(1)}(L)$ und daher (3) mit $\beta = \infty$. Also existieren m linear unabhängige Eigenelemente erster Art zu den Eigenwerten $k_0, k_1, \cdots, k_{m-1}$; hieraus folgt (2).

(b) Es sei $\mathfrak{H}$ unendlichdimensional und sei $\lambda \in \sigma^{(1)}(L) \cap [\alpha, \beta)$. Nach Voraussetzung und wegen Lemma 5. 3 ist dann λ ein Eigenwert erster Art und endlicher Vielfachheit von L.

Wir beweisen zunächst, daß es zwei benachbarte Minimum-Maximum-Werte k_n und k_{n+1} gibt, so daß $k_n \leqq \lambda < k_{n+1}$ gilt. Andernfalls wäre nämlich $k_l \in [\alpha, \lambda] \subset \subset [\alpha, \beta)$ für alle $l = 0, 1, \cdots$; es gäbe also nach (1) und (3) in $[\alpha, \lambda)$ unendlich viele voneinander verschiedene Eigenwerte erster Art und somit einen Häufungspunkt $\mu \in [\alpha, \beta)$ der Menge $\sigma^{(1)}(L)$, für den wegen Lemma 4. 4 $\mu \in \sigma^{(1)}(L) \setminus \sigma_i^{(1)}(L)$ wäre. Dies steht jedoch im Widerspruch zur Voraussetzung $\sigma^{(1)}(L) \cap [\alpha, \beta) \subset \sigma_i^{(1)}(L)$.

Also gibt es unter allen natürlichen Zahlen n mit $k_n > \lambda$ eine kleinste, die mit n_0 bezeichnet werden soll. Weiter sei $S = \{x_0, x_1, \cdots, x_{n_0 - 1}\}$ ein nach (3) existierendes System linear unabhängiger Eigenelemente zu den Eigenwerten $k_0, k_1, \cdots, \cdots, k_{n_0-1}$. Gäbe es nun ein zu λ gehöriges Eigenelement y, das von dem System S linear unabhängig ist, so wäre die lineare Hülle $\mathfrak{L}$ von S und y ein (n_0+1)-dimensionaler Teilraum von $\mathfrak{H}$, für den einerseits nach Lemma 5. 2 $s = \sup\limits_{x\in\mathfrak{L}\setminus\{0\}} p(x) \geqq k_{n_0}$, also $s > \lambda$ und andererseits nach Folgerung 1 aus Lemma 5. 1 $s \leqq \lambda$ wäre. Widerspruch.

Folglich tritt λ unter der Menge der k_n $(n=0, 1, \cdots)$ seiner Vielfachheit entsprechend oft auf, w.z.z.w.

Das folgende Lemma gibt mit Hilfe von Satz 5. 1 eine andere Charakterisierung der Minimum-Maximum-Werte an. Hierbei bezeichne $\mathfrak{M}_n$ die Gesamtheit aller n-dimensionalen Teilräume von $\mathfrak{H}$.

Lemma 5. 4. *Unter den Voraussetzungen von Satz* 5. 1 *gilt für jedes* $k_i \in [\alpha, \beta)$

(5. 19) $$k_i = \min_{\mathfrak{L}\in\mathfrak{M}_{i+1}} \max_{x\in\mathfrak{L}\setminus\{0\}} p(x).$$

Beweis. Nach Satz 5. 1 existieren $i+1$ linear unabhängige Eigenelemente $x_0, x_1, \cdots, x_i$ erster Art von L zu den Eigenverten $\lambda_0 = k_0, \lambda_1 = k_1, \cdots, \lambda_i = k_i$. Dann gehört die lineare Hülle $\mathfrak{X}_{i+1}$ der Elemente $x_0, x_1, \cdots, x_i$ zu $\mathfrak{M}_{i+1}$ und es gilt nach Folgerung 1 aus Lemma 5. 1

(5. 20) $$k_i = \lambda_i = \max_{x\in\mathfrak{X}_{i+1}\setminus\{0\}} p(x).$$

Aus Lemma 5. 2 folgt aber $\max\limits_{x\in\mathfrak{L}\setminus\{0\}} p(x) \geqq k_i$ [16]) für jedes $\mathfrak{L}\in\mathfrak{M}_{i+1}$. Also gilt

(5. 21) $$\inf_{\mathfrak{L}\in\mathfrak{M}_{i+1}} \max_{x\in\mathfrak{L}\setminus\{0\}} p(x) \geqq k_i.$$

(5. 20) und (5. 21) ergeben schließlich (5. 19).

Wir geben nun noch ein zweites Minimaxprinzip zur Bestimmung von Eigenwerten der Schar L, begonnen beim größten Eigenwert erster Art, an. Zu diesem Zwecke definieren wir unter Beibehaltung der vorn eingeführten Bezeichnungen

$$m_i = \inf_{\mathfrak{L}\in\mathfrak{L}_i} \sup_{x\in\mathfrak{L}\setminus\{0\}} p(x) \qquad (i = 0, 1, \cdots).$$

Dabei heiße m_i der *i-te Maximum-Minimum-Wert.*

Satz 5. 2 (Maximum-Minimum-Prinzip). *Es sei* $\alpha = \sup\limits_{x\in\mathfrak{H}\setminus\{0\}} p(x) \neq \infty$, *und es existiere eine reelle Konstante* $\beta < \alpha$ *mit der Eigenschaft* $\sigma^{(1)}(L) \cap (\beta, \alpha] \subset \sigma_i^{(1)}(L)$. *Dann gilt:*

---

[16]) Im Beweis zu Lemma 5. 2 wurde gezeigt, daß auf jedem endlichdimensionalen Teilraum $\mathfrak{M} \subset \mathfrak{H}$ $\sup\limits_{x\in\mathfrak{M}\setminus\{0\}} p(x) = \max\limits_{x\in\mathfrak{M}\setminus\{0\}} p(x)$ gilt.

(1) *Jeder Maximum-Minimum-Wert* $m_n \in (\beta, \alpha]$ *ist ein Eigenwert erster Art und endlicher Vielfachheit von* L.

(2) *Jeder Punkt* $\lambda \in \sigma^{(1)}(L) \cap (\beta, \alpha]$ *ist ein Eigenwert erster Art von endlicher Vielfachheit und tritt in der Menge der* m_n $(n=0, 1, \cdots)$ *seiner Vielfachheit entsprechend oft auf.*

(3) *Es sei* $m_n \in (\beta, \alpha]$. *Dann existiert ein System von* $n+1$ *linear unabhängigen Eigenelementen* $x_0, x_1, \cdots, x_n$ *mit den zugehörigen Eigenwerten* $\lambda_0 = m_0, \lambda_1 = m_1, \cdots$ $\cdots, \lambda_n = m_n$.[17]) *Definieren wir* $\mathfrak{H}_n = \{x \in \mathfrak{H} \mid ([(m_n+\lambda_j)A+B]x_j, x) = 0 \quad (j=0, \cdots$ $\cdots, n-1)\}$, *so gilt überdies*

(3_1) $m_n = \max_{x \in \mathfrak{H}_n \setminus \{0\}} p(x)$.

(3_2) *Jedes Element* $y \in \mathfrak{H}_n$, $y \neq 0$, *mit* $p(y) = m_n$ *ist ein zu* m_n *gehöriges Eigenelement und ist außerdem linear unabhängig von den Elementen* $x_0, x_1, \cdots, x_{n-1}$.

Beweis. Anwendung von Satz 5. 1 auf die Schar (vgl. [1])

$$L^-(\lambda) = \lambda^2 A^- + \lambda B^- + C^-$$

mit $A^- = -A$, $B^- = B$, $C^- = -C$ liefert unmittelbar die Behauptung.

6. Anwendbarkeit der Minimaxprinzipe

Der vorliegende Abschnitt gibt einige einfache hinreichende Bedingungen für die Anwendbarkeit der Minimaxprinzipe an.

Zur Abkürzung werden wir im folgenden sagen, daß die Schar L für eine reelle Zahl $\lambda \neq \infty$ [18]) der Bedingung (V_λ) genügt, wenn gilt:

(V_λ) Für jede zu λ gehörige Folge (x_n) erster Art mit $x_n \rightharpoonup x_0$ $(n \to \infty)$ gilt $Cx_n \to Cx_0$ $(n \to \infty)$.

Außerdem werden im weiteren stellenweise einige der nachstehenden Bedingungen (I_A), (I_B), (I_C) und (B) benutzt.

(I_A) (bzw. (I_B) oder (I_C)). Die Schar L sei so beschaffen, daß, falls $0 \in \sigma(A)$ (bzw. $0 \in \sigma(B)$ oder $0 \in \sigma(C)$) gilt, 0 ein isolierter Punkt des Spektrums von A (bzw. B oder C) von endlicher Vielfachheit ist.

(B) Der Operator B der Schar $L = \lambda^2 A + \lambda B + C$ sei positiv [19]) oder vollstetig.

Lemma 6. 1. *Jede der nachstehenden Bedingungen ist hinreichend für die Beschränktheit des Funktionals* p:

(1) *Die Form* (Ax, x) $(x \in \mathfrak{H})$ *ist streng indefinit.*

(2) L *genügt der Bedingung* (I_A).

---

[17]) Es sei daran erinnert, daß nach Definition $\alpha = m_0 \geqq m_1 \geqq \ldots$ gilt.

[18]) Im weiteren werden stets nur endliche reelle Zahlen betrachtet.

[19]) D. h. $(Bx, x) \geqq 0$ für alle $x \in \mathfrak{H}$ (Bezeichnung: $B \geqq 0$).

Beweis. Im Falle (1) folgt die Aussage direkt aus Lemma 3. 4.

(2) L genüge der Bedingung (I_A). Angenommen, p wäre unbeschränkt, d.h., es gäbe eine Folge (x_n), $x_n \in \mathfrak{H}$, $\|x_n\|=1$ $(n=1, 2, \cdots)$ mit

$$|p(x_n)| \geqq n \qquad (n=1, 2, \cdots). \tag{6. 1}$$

Wegen (1) ist dann die Form (Ax, x) (semi-) definit. Weiter existiert nach Definition (2. 2) auf Grund von (6. 1) eine Teilfolge (x_{n_k}), für die $\lim\limits_{k\to\infty}(Ax_{n_k}, x_{n_k})=0$ gilt.

Aus der (beschränkten) Folge (x_{n_k}) wählen wir eine schwach konvergente Teilfolge (y_l), $y_l = x_{n_{k_l}}$ $(l=1, 2, \cdots)$: $y_l \rightharpoonup y_0$, für die wegen $\lim\limits_{l\to\infty}(Ay_l, y_l)=0$

$$Ay_l \to 0 \quad (l\to\infty)$$

gilt, wie man durch Anwendung der wegen der Definitheit von (Ax, x) $(x \in \mathfrak{H})$ gültigen Schwarzschen Ungleichung aus

$$(Ay_l, Ay_l)^2 \leqq (Ay_l, y_l)\|A\|^3 (y_l, y_l) \qquad (l=1, 2, \cdots)$$

erkennt.

Es bezeichne nun N den Nullraum von A und R dessen orthogonales Komplement. R und N reduzieren bekanntlich A und es gilt $\mathfrak{H}=R\oplus N$; $y_l = y_l^N + y_l^R$ $(l=1, 2, \cdots)$ sei die hierzu gehörige Zerlegung von y_l. Dann folgt aus $Ay_l \to 0$ $(l\to\infty)$ wegen $Ay_l^N=0$ $(l=1, 2, \cdots)$ die Beziehung $Ay_l^R \to 0$ $(l\to\infty)$ und auf Grund der Voraussetzung (I_A) hieraus $y_l^R \to 0$ $(l\to\infty)$. Wegen $y_l^N + y_l^R = y_l \rightharpoonup y_0$ $(l\to\infty)$ ergibt sich somit $y_l^N \rightharpoonup y_0$ und wegen der nach (I_A) endlichen Dimension von N also $y_l \to y_0$ $(l\to\infty)$, $\|y_0\|=1$.

Infolge der Stetigkeit des Funktionals p (s. Lemma 3. 1) erhält man schließlich

$$\lim_{l\to\infty} p(x_{n_{k_l}}) = \lim_{l\to\infty} p(y_l) = p(y_0)$$

in Widerspruch zu (6. 1).

Lemma 6. 2. *Jede der folgenden beiden Bedingungen ist hinreichend dafür, daß die Schar L für jedes reelle λ der Bedingung (V_λ) genügt:*

(1) *Der Operator C ist vollstetig.*

(2) *Die Operatoren A und B sind beide vollstetig.*

Beweis. Im Falle (1) ist die Aussage evident. Bei (2) folgt wegen der Vollstetigkeit von A und B für jede Folge (x_n) erster Art mit $x_n \rightharpoonup x_0$ $(n\to\infty)$ aus $L(\lambda)x_n = (\lambda^2 A + \lambda B)x_n + Cx_n \to 0 = L(\lambda)x_0 = (\lambda^2 A + \lambda B)x_0 + Cx_0$ $(n\to\infty)$ die Beziehung $Cx_n \to Cx_0$ $(n\to\infty)$, w.z.z.w.

Lemma 6. 3. *Genügt die Schar L in einem Punkt $\lambda \neq 0$ der Bedingung (V_λ) und außerdem der Bedingung* (B), *so gilt für jede zu λ gehörige Folge (x_n) erster Art mit $x_n \rightharpoonup 0$ $(n\to\infty)$ gleichzeitig $Ax_n \to 0$ $(n\to\infty)$ und $Bx_n \to 0$ $(n\to\infty)$.*

Beweis. Es sei (x_n) eine zu λ gehörige Folge erster Art mit $x_n \rightharpoonup 0$ $(n \to \infty)$. Dann folgt zunächst aus (V_λ)

(6. 2) $$Cx_n \to 0 \qquad (n \to \infty).$$

Der Voraussetzung (B) entsprechend unterscheiden wir zwei Fälle.

(1) B sei vollstetig. Dann gilt $Bx_n \to 0$ $(n \to \infty)$ und wegen $L(\lambda)x_n = \lambda^2 Ax_n + \lambda Bx_n + Cx_n \to 0$ $(n \to \infty)$ infolge (6. 2) und $\lambda \neq 0$ schließlich $Ax_n \to 0$ $(n \to \infty)$, w.z.z.w.

(2) B sei positiv. Wir betrachten die nach (2. 2) gültige Beziehung

(6. 3) $$[(Bx_n, x_n) + d(x_n)]p(x_n) = -2(Cx_n, x_n).$$

Da aus (6. 2) wegen $\|x_n\| = 1$ $(n = 1, 2, \cdots)$ die Gleichung $\lim_{n \to \infty} (Cx_n, x_n) = 0$ folgt, liefert (6. 3) auf Grund von $\lim_{n \to \infty} p(x_n) = \lambda \neq 0$

$$\lim_{n \to \infty} [(Bx_n, x_n) + d(x_n)] = 0$$

und somit wegen $B \geqq 0$ und $d(x_n) > 0$ $(n = 1, 2, \cdots)$

$$\lim_{n \to \infty} (Bx_n, x_n) = 0.$$

Infolge $B \geqq 0$ ergibt sich hieraus

$$Bx_n \to 0 \qquad (n \to \infty).$$

Wie in (1) erhält man nun $Ax_n \to 0$ $(n \to \infty)$, w.z.z.w.

Lemma 6. 4. *Die Schar L genüge für eine reelle Zahl λ der Bedingung* (V_λ). *Ferner gelte* (B). *Ist dann (x_n) eine zu λ gehörige Folge erster Art mit $x_n \rightharpoonup x_0 \neq 0$ $(n \to \infty)$, so gilt $p(x_0) = \lambda$.*

Beweis. Wir vermerken zunächst, daß aus $x_n \rightharpoonup x_0$ $(n \to \infty)$ die Beziehung $L(\lambda)x_n \rightharpoonup L(\lambda)x_n$ $(n \to \infty)$ folgt und somit wegen $L(\lambda)x_n \to 0$ die Gleichung

$$L(\lambda)x_0 = 0$$

gilt. Daher genügt es auf Grund von Lemma 2. 3, zum Beweis von $p(x_0) = \lambda$ die Gültigkeit der Ungleichung

(6. 4) $$2\lambda(Ax_0, x_0) + (Bx_0, x_0) \geqq 0$$

zu zeigen. (B) entsprechend unterscheiden wir hierzu zwei Fälle.

(1) Es sei $B \geqq 0$.

(a) Im Falle $\lambda = 0$ gilt offenbar (6. 4).

(b) Es sei $\lambda \neq 0$. Wir setzen ohne Einschränkung der Allgemeinheit voraus, daß die offenbar beschränkte Zahlenfolge $((Bx_n, x_n))$ konvergiert.

Aus $L(\lambda)x_n \to L(\lambda)x_0 = 0$ folgt, da wegen (V_λ) $Cx_n \to Cx_0$ $(n \to \infty)$ gilt, die Beziehung $\lambda^2 Ax_n + \lambda Bx_n \to \lambda^2 Ax_0 + \lambda Bx_0$ $(n \to \infty)$ und unter Beachtung von $\lambda \neq 0$ somit

$$2\lambda Ax_n + 2Bx_n \to 2\lambda Ax_0 + 2Bx_0 \qquad (n \to \infty).$$

Deshalb ergibt sich auf Grund der Relationen $x_n \rightharpoonup x_0$ $(n\to\infty)$ und $\lim_{n\to\infty} 2(\lambda - p(x_n))(Ax_n, x_n) = 0$

$$\lim_{n\to\infty} [2p(x_n)(Ax_n, x_n) + 2(Bx_n, x_n)] = 2\lambda(Ax_0, x_0) + 2(Bx_0, x_0).$$

Mit $f_{p(x_n)}(x_n) = 2p(x_n)(Ax_n, x_n) + (Bx_n, x_n)$ erhält man so

$$\text{(6. 5)} \quad 2\lambda(Ax_0, x_0) + (Bx_0, x_0) = \lim_{n\to\infty} f_{p(x_n)}(x_n) + \lim_{n\to\infty} (Bx_n, x_n) - (Bx_0, x_0).$$

Weiter gilt, weil aus $x_n \rightharpoonup x_0$ die Beziehung $\lim_{n\to\infty}(Bx_n, x_0) = (Bx_0, x_0)$ folgt, infolge $B \geqq 0$

$$\text{(6. 6)} \quad \begin{aligned} \lim_{n\to\infty} (Bx_n, x_n) - (Bx_0, x_0) &= \lim_{n\to\infty} [(Bx_n, x_n) - (Bx_n, x_0) - (Bx_0, x_n) + (Bx_0, x_0)] = \\ &= \lim_{n\to\infty} (B(x_n - x_0), (x_n - x_0)) \geqq 0. \end{aligned}$$

Nun ist nach (2. 9) $f_{p(x_n)}(x_n) > 0$. Somit liefern (6. 5) und (6. 6) die Ungleichung (6. 4), w.z.z.w.

(2) B sei vollstetig. Dan ist

$$\text{(6. 7)} \quad Bx_n \to Bx_0 \qquad (n\to\infty).$$

(a) Gilt $\lambda = 0$, so folgt aus der nach (2. 9) gültigen Ungleichung $f_{p(x_n)}(x_n) = 2p(x_n)(Ax_n, x_n) + (Bx_n, x_n) > 0$ wegen der Beziehungen $\lim_{n\to\infty} p(x_n) = \lambda = 0$, $|(Ax_n, x_n)| \leqq \|A\|$ $(n = 1, 2, \cdots)$ und $\lim_{n\to\infty}(Bx_n, x_n) = (Bx_0, x_0)$ die Ungleichung

$$(Bx_0, x_0) = \lim_{n\to\infty} f_{p(x_n)}(x_n) \geqq 0,$$

also (6. 4).

(b) Es sei $\lambda \neq 0$. Wie in (1) gelangt man zu (6. 5). Daraus folgt auf Grund von (6. 7) $\lim_{n\to\infty} f_{p(x_n)}(x_n) = 2\lambda(Ax_0, x_0) + (Bx_0, x_0)$. Wieder gilt (6. 4) wegen $f_{p(x_n)}(x_n) > 0$.

Lemma 6. 5. *Es sei $\lambda \in \sigma^{(1)}(L)$. Ferner gelte für die Schar L die Bedingung* (B).

(1) *Genügt L im Punkte λ der Bedingung* (V_λ) *und erfüllt L außerdem die Bedingung* (I_C), *so gilt $\lambda \in \sigma_i^{(1)}(L)$.*

(2) *Ist $\lambda \neq 0$ und genügt L neben der Bedingung* (V_λ) *wenigstens einer der Bedingungen* (I_A) *oder* (I_B), *so gilt $\lambda \in \sigma_i^{(1)}(L)$.*

Beweis. Wegen $\lambda \in \sigma^{(1)}(L)$ existiert eine zu λ gehörige Folge erster Art (y_m), die infolge $\|y_m\| = 1$ $(m = 1, 2, \cdots)$ eine schwach konvergente Teilfolge (x_n) mit $x_n \rightharpoonup x$ $(n\to\infty)$ enthält.

Angenommen, es wäre $\lambda \notin \sigma_i^{(1)}(L)$. Dann ergäbe sich wegen Lemma 6. 4 $x = 0$, d. h. $x_n \rightharpoonup 0$ $(n\to\infty)$.

Fall (1). Aus (V_λ) folgte nunmehr $Cx_n \to 0$ $(n \to \infty)$. Dies steht aber nach dem auf Seite 51 zitierten Kriterium von H. WEYL im Widerspruch zur Bedingung (I_C).

Fall (2). Aus Lemma 6. 3 erhielte man dann sowohl $Ax_n \to 0$ $(n \to \infty)$ als auch $Bx_n \to 0$ $(n \to \infty)$. Dies ist wieder nach dem Kriterium von H. WEYL zu (I_A) bzw. (I_B) im Widerspruch.

Wir können nun die Minimaxprizipe auf spezielle Scharen anwenden.

Satz 6. 1. *Die Schar $L = \lambda^2 A + \lambda B + C$ genüge den (auf Seite 42 bzw. 62 genannten) Voraussetzungen* (D^+), (B) *und* (I_A). *Ferner sei der Operator C vollstetig. Dann sind die Sätze* 5. 1 *und* 5. 2 *mit* $\beta = 0$ *anwendbar und es ergibt sich:*

Die Menge $\sigma^{(1)}(L) \setminus \{0\}$ besteht aus höchstens abzählbar vielen isolierten Punkten, die sich nur im Nullpunkt häufen können, und jedes $\lambda \in \sigma^{(1)}(L)$ mit $\lambda \neq 0$ ist ein Eigenwert erster Art von endlicher Vielfachheit.

Gilt $C \neq 0$, so ist $\sigma^{(1)}(L) \setminus \{0\} \neq \emptyset$; das Minimum-Maximum-Prinzip (*Satz* 5. 1) *liefert dann alle negativen und das Maximum-Minimum-Prinzip* (*Satz* 5. 2) *alle positiven Eigenwerte erster Art von L.*

Beweis. Auf Grund der Voraussetzung des Satzes folgt aus Lemma 6. 1 $\inf_{x \in \mathfrak{H} \setminus \{0\}} p(x) \neq \infty$, $\sup_{x \in \mathfrak{H} \setminus \{0\}} p(x) \neq \infty$ und aus Lemma 6. 5, (2), mittels Lemma 6. 2, (1), $\sigma^{(1)}(L) \setminus \{0\} \subset \sigma_i^{(1)}(L)$. Daher sind die Voraussetzungen zu den Sätzen 5. 1 und 5. 2 mit $\beta = 0$ erfüllt. Hieraus folgen alle übrigen Aussagen des Satzes, wobei man beim Beweis von $\sigma^{(1)}(L) \setminus \{0\} \neq \emptyset$ im Falle $C \neq O$ beachte, daß dann entweder $k_0 = \inf_{x \in \mathfrak{H} \setminus \{0\}} p(x) < 0$ oder $m_0 = \sup_{x \in \mathfrak{H} \setminus \{0\}} p(x) > 0$ gilt.

Satz 6. 2. *Die Schar $L = \lambda^2 A + \lambda B + C$ genüge den (auf Seite 42 bzw. 62 genannten) Bedingungen* (D^+) *und* (I_C). *Ferner seien die Operatoren A und B vollstetig, und es gelte $A \neq O$. Existiert dann ein $x \in \mathfrak{H}$ mit $(Ax, x) > 0$, so ist Satz* 5. 1 *und andernfalls Satz* 5. 2 *mit unendlichem β anwendbar. Dabei ergibt sich:*

Gilt $\dim \mathfrak{H} = n < \infty$, so besteht $\sigma^{(1)}(L)$ aus endlich vielen Eigenwerten, zu denen ein System von genau n linear unabhängigen Eigenelementen erster Art gehört.

Ist $\mathfrak{H}$ unendlichdimensional, so besteht $\sigma^{(1)}(L) \setminus \{\infty\}$ aus genau abzählbar vielen isolierten Punkten, die sich nur gegen ∞ häufen. Jedes $\lambda \in \sigma^{(1)}(L) \setminus \{\infty\}$ ist ein Eigenwert erster Art von endlicher Vielfachheit. Der Operator A ist dann positiv oder negativ.

Beweis. Wir vermerken zunächst, daß nach Voraussetzung für L die Bedingung (B) gilt und somit das Lemma 6. 5 angewendet werden kann.

(1) Es existiere ein $x \in \mathfrak{H}$ mit $(Ax, x) > 0$. Dann folgt aus (3. 3) $\inf_{x \in \mathfrak{H} \setminus \{0\}} p(x) \neq \infty$ und wegen Lemma 6. 2, (2) aus Lemma 6. 5, (1), $\sigma^{(1)}(L) \setminus \{\infty\} \subset \sigma_i^{(1)}(L)$. Daher sind die Voraussetzungen zu Satz 5. 1 mit unendlichem β erfüllt, woraus die Aussagen des Satzes folgen, wenn man beachtet, daß $p(x)$ und somit alle Minimum-Maximum-Werte k_n $(n = 0, 1, \cdots)$ nur endliche Werte annehmen.

Im Falle unendlicher Dimension von $\mathfrak{H}$ ergibt sich aus der Häufung von $\sigma^{(1)}(L)\setminus\{\infty\}$ gegen ∞ die Unbeschränktheit von p und daher nach Lemma 6. 1, (1), die Definitheit der Form (Ax, x) $(x\in\mathfrak{H})$; also ist A entweder positiv oder negativ.

(2) Existiert kein $x\in\mathfrak{H}$ mit $(Ax, x)>0$, so gibt es wegen $A\neq O$ ein $y\in\mathfrak{H}$ mit $(Ay, y)<0$. (3. 4) liefert dann $\sup\limits_{x\in\mathfrak{H}\setminus\{0\}} p(x)\neq\infty$. (1) entsprechend zeigt man, daß jetzt Satz 5. 2 anwendbar ist, w.z.z.w.

Folgerung. *Die Schar L genüge den Bedingungen* (D^+) *und* (I_A). *Ferner seien die Operatoren B und C vollstetig, und es sei $B\geqq O$, $C\neq O$. Dann gilt:*

Ist $\dim\mathfrak{H}=n<\infty$, *so besteht $\sigma^{(2)}(L)$ aus endlich vielen Eigenwerten, zu denen ein System von genau n linear unabhängigen Eigenelementen zweiter Art gehört.*

Besitzt $\mathfrak{H}$ unendliche Dimension, so besteht $\sigma^{(2)}(L)\setminus\{0\}$ aus genau abzählbar vielen isolierten Punkten, die sich nur gegen 0 *häufen. Jedes $\lambda\in\sigma^{(2)}(L)\setminus\{0\}$ ist ein Eigenwert zweiter Art von endlicher Vielfachheit. Der Operator C ist dann positiv oder negativ.*

Beweis. Wir ordnen der Schar L entsprechend Lemma 3. 6, (2), (bzw. Lemma 4. 2) die Schar $\tilde{L}_0(\lambda)=\lambda^2\tilde{A}_0+\lambda\tilde{B}_0+\tilde{C}_0$ mit $\tilde{A}_0=C$, $\tilde{B}_0=B$, $\tilde{C}_0=A$ zu. Offenbar erfüllt $\tilde{L}_0$ die Bedingungen (D^+), ($I_{\tilde{C}_0}$) und genügt daher wegen $\tilde{A}_0=C\neq O$ und auf Grund der Vollstetigkeit von $B=\tilde{B}_0$, $C=\tilde{A}_0$ den Voraussetzungen zu Satz 5. 2, der dann unter Benutzung von Lemma 4. 2 die Behauptung liefert.

An dieser Stelle sei nochmals auf die bereits zitierten Arbeiten [1] und [2] hingewiesen. In [1] wird in einem endlichdimensionalen Raum die Schar L unter der Bedingung (D) und der zusätzlichen Voraussetzung $(Bx, x)\geqq 0$ $(x\in\mathfrak{H})$ betrachtet; in [2] werden u.a. stark gedämpfte Scharen untersucht, die sich durch eine Parametertransformation auf diesen Fall zurückführen lassen. Stets genügt hierbei die Schar L der Bedingung (D^+). Wegen der endlichen Dimension von $\mathfrak{H}$ sind auch die übrigen Voraussetzungen zu den Sätzen 6. 1 und 6. 2 erfüllt. Wir erhalten so als Spezialfall die in [1] bewiesenen Aussagen.

Literatur

[1] R. J. Duffin, A minimax theory for overdamped networks, *J. Rat. Mech. and Anal.*, **4** (1955), 221—233.

[2] R. J. Duffin, The Rayleigh—Ritz method for dissipative or gyroscopic systems, *Quarterly J. Appl. Math.*, **18** (1960), 215—221.

[3] M. G. Krein u. H. Langer, Zur Theorie quadratischer Scharen selbstadjungierter Operatoren, *Dokl. Akad. Nauk SSSR*, **154** (1964), 1258—1261 [russisch].

[4] M. G. Krein u. H. Langer, Über einige mathematische Prinzipien der linearen Theorie kleiner Schwingungen der Kontinua, *Arbeiten des intern. Symp. über Anwendungen der Funktionentheorie i. d. Kontinuumsmechanik* (Moskau, 1965) [russisch].

[5] R. Kühne, Über eine Klasse *J*-selbstadjungierter Operatoren, *Math. Ann.*, **154** (1964), 56—69.

[6] P. H. Müller, Eine neue Methode zur Behandlung nichtlinearer Eigenwertaufgaben, *Math. Zeitschr.*, **70** (1959), 381—406.

[7] P. H. Müller, Eigenwertabschätzungen für Gleichungen vom Typ $(\lambda^2 I - \lambda A - B)x = 0$, *Archiv d. Math.*, **12** (1961), 307—310.

[8] E. Pesonen, Über die Spektraldarstellung quadratischer Formen in linearen Räumen mit indefiniter Metrik, *Ann. Acad. Sci. Fenn.* A I, **227** (1956).

[9] F. Riesz u. B. Sz.-Nagy, *Vorlesungen über Funktionalanalysis* (Berlin, 1956).

[10] A. C. Zaanen, *Linear analysis* (Amsterdam, 1953).

(Eingegangen am 10. April 1967)

The shell of a Hilbert-space operator

By CHANDLER DAVIS in Toronto (Canada)

For an arbitrary closed linear operator in Hilbert space, I will define a subset of real 3-space which summarizes much information about it: its point spectrum, its numerical range, many of its spectral sets, and more besides.

1. Notations and principal ideas

Let $\mathfrak{H}$ be a complex Hilbert space. For any $x \in \mathfrak{H}$, let the corresponding linear functional be x^*; thus y^*x is the inner product of x by y.

Let $\mathfrak{A}$ be a closed linear relation in $\mathfrak{H}$; that is, a closed linear subspace of $\mathfrak{H}_1 \oplus \mathfrak{H}_2$, where each $\mathfrak{H}_i$ is a replica of $\mathfrak{H}$. (No distinction is to be made between $\mathfrak{H}_1$ and $\mathfrak{H}_1 \oplus \{0\}$, or between $\mathfrak{H}_2$ and $\{0\} \oplus \mathfrak{H}_2$.) The most important case is that of an operator A, i.e. when $(y, x) \in \mathfrak{A}$ means that $y = Ax$. The 'domain' of $\mathfrak{A}$ is $\{x\colon (\exists y)(y, x) \in \mathfrak{A}\}$, its 'range' is $\{y\colon (\exists x)(y, x) \in \mathfrak{A}\}$. (This reversal of the customary order in the notation for relations will save me, in § 5, from having to reverse order in a more troublesome way.)

Before giving the novel ideas I must also fix the notations for stereographic projection. Let $\boldsymbol{C}$ denote the complex plane and $\overline{\boldsymbol{C}} = \boldsymbol{C} \cup \{\infty\}$. Define $\tau\colon \overline{\boldsymbol{C}} \to \boldsymbol{R}^3$ by

$$\tau(z) = \left(\frac{2z}{1+|z|^2}, \frac{-1+|z|^2}{1+|z|^2}\right) \quad (z \in \boldsymbol{C}), \quad \text{and} \quad \tau(\infty) = (0, 1). \tag{1.1}$$

The first two co-ordinates in $\boldsymbol{R}^3$ are here collapsed into a single complex number; this will be done frequently throughout. In this notation, the Riemann sphere $\tau(\overline{\boldsymbol{C}})$ is the unit sphere $S = \{(\zeta, h) \in \boldsymbol{R}^3\colon |\zeta|^2 + h^2 = 1\}$. Letting $B = \{(\zeta, h) \in \boldsymbol{R}^3\colon |\zeta|^2 + h^2 \leqq 1\}$, the unit ball, define $\pi\colon B \to \overline{\boldsymbol{C}}$ by taking $\pi(\zeta, h)$ to be that $z \in \boldsymbol{C}$ for which (ζ, h) is on the line joining $(z, 0)$ to $(0, 1)$; explicitly,

$$\pi(\zeta, h) = \frac{\zeta}{1-h} \quad (h \neq 1), \quad \text{and} \quad \pi(0, 1) = \infty. \tag{1.2}$$

For any subset $E \subseteqq B$, $\pi(E)$ will be called the 'shadow' of E. On S, π coincides with τ^{-1}.

Definition 1. 1. The 'shell' of the relation $\mathfrak{A}$, denoted $s(\mathfrak{A})$, is defined as the set of points

$$\varphi(y, x) = \left(\frac{2x^*y}{\|x\|^2 + \|y\|^2}, \frac{-\|x\|^2 + \|y\|^2}{\|x\|^2 + \|y\|^2} \right) \tag{1.3}$$

in $\boldsymbol{R}^3$, where (y, x) runs over all non-zero elements of $\mathfrak{A}$.

In case of an operator, I will write alternatively $s(A)$. This is the case where $(0, 1)$ does not belong to the set. $s(\mathfrak{A})$ is void if and only if $\mathfrak{A} = \{(0, 0)\}$.

One finds by direct computation that $s(\mathfrak{A}) \subseteqq B$. Further geometric properties of the set, in relation to spectral properties of $\mathfrak{A}$, will be developed below, especially in §§ 2—3. Examples are treated in § 4. In § 5, I will state the essential facts on the transformation of $s(\mathfrak{A})$ when $\mathfrak{A}$ is subjected to a Möbius transformation.

Next, consider spectral sets [16], [17, § 154].

Let us adopt the following terminology:

(i) a set of the form $\{z \in \boldsymbol{C}: |z - z_0| \leqq r\}$ $(z_0 \in \boldsymbol{C}, r > 0)$ will be called a 'finite disk';

(ii) a set of the form $\{z \in \boldsymbol{C}: |z - z_0| \geqq r\} \cup \{\infty\}$ $(z_0 \in \boldsymbol{C}, r > 0)$ will be called a 'complementary disk';

(iii) a set of the form $\{z \in \boldsymbol{C}: \operatorname{Re}(\zeta z) \geqq a\} \cup \{\infty\}$ $(|\zeta| = 1, a \in \boldsymbol{R})$ will be called a 'half-plane';

(iv) a set of any of the types (i)—(iii) will be called a 'disk'. Thus the disks are exactly those subsets X of $\overline{\boldsymbol{C}}$ for which $\tau(X)$ is a proper spherical cap.

The main result of the paper, Theorem 7. 2, may be stated roughly as follows: A disk X is a spectral set for A if and only if $s(A)$ is contained in the convex hull of $\tau(X)$. That is, the support planes of (the convex closure of) $s(A)$ correspond naturally one-one to the minimal disk spectral sets of A. This is exploited in § 8 to give a description in terms of the shell of the operator classes occurring in [22].

In order to formulate my results for arbitrary closed linear relations, I had to supplement the basic results of the paper [1] [1]) of Arens with a study of the spectrum (§ 2, below) and of the rational functional calculus (§ 6).

My grateful acknowledgement is due to H. S. M. Coxeter, C. Foiaş, G. Kalisch, and C. A. McCarthy, for stimulating conversations.

[1]) The reader is warned of the uncommonly pesky misprints in the article [1]. Professor Arens has pointed out also that Theorem 3. 7 of [1] is not true in quite the generality claimed.

2. The shell and the spectrum

It has already been remarked that $s(\mathfrak{A})\subseteq B$. The shell may have points on the boundary, the sphere S; this depends on spectral properties of $\mathfrak{A}$, see Theorem 2. 2 below. I define the spectrum in a form convenient for the purpose; cf. [13, § 2. 16], [1, § 2]. (In the following definitions, recall that it is $\mathfrak{H}_2$ which contains the domain.)

Definition 2. 1. $0\in\sigma_p(\mathfrak{A})$, the 'point spectrum' of $\mathfrak{A}$, in case $\mathfrak{A}\cap\mathfrak{H}_2\neq\{(0,0)\}$. The 'null-space' $\mathfrak{N}(\mathfrak{A})$ is the set of $x\in\mathfrak{H}$ such that $(0,x)\in\mathfrak{A}$, thus $0\in\sigma_p(\mathfrak{A})$ if and only if $\mathfrak{N}(\mathfrak{A})\neq\{0\}$.

Definition 2. 2. $0\in\sigma_c(\mathfrak{A})$, the 'continuous spectrum' of $\mathfrak{A}$, in case there exist $x_n\in\mathfrak{H}\ominus\mathfrak{N}(\mathfrak{A})$, $y_n\in\mathfrak{H}$, such that $(y_n, x_n)\in\mathfrak{A}$, $\|x_n\|=1$, and $y_n\to 0$.

Definition 2. 3. The 'range' $\mathfrak{R}(\mathfrak{A})$ is the set of $y\in\mathfrak{H}$ such that for some $x\in\mathfrak{H}$, $(y,x)\in\mathfrak{A}$. $0\in\sigma_r(\mathfrak{A})$, the 'residual spectrum' of $\mathfrak{A}$, in case $\overline{\mathfrak{R}(\mathfrak{A})}\neq\mathfrak{H}$.

According to these definitions, 0 may belong to any one of σ_p. σ_c, σ_r, independently of whether it belongs to the others.

Let $\mathfrak{J}$ denote the identity relation: $(y,x)\in\mathfrak{J}$ if and only if $x=y$.

Definition 2. 4. For $z\in\boldsymbol{C}$, $z\in\sigma_p(\mathfrak{A})$ if and only if $0\in\sigma_p(\mathfrak{A}-z\mathfrak{J})$, and similarly for σ_c, σ_p. $\infty\in\sigma_p(\mathfrak{A})$ if and only if $0\in\sigma_p(\mathfrak{A}^{-1})$, and similarly for σ_c, σ_r.

Thus $\infty\in\sigma_r(\mathfrak{A})$ if and only if the domain $\mathfrak{D}(\mathfrak{A})$ is not dense.

Definition 2. 5. The 'approximate point spectrum' $\sigma_\pi(\mathfrak{A})$ is $\sigma_p(\mathfrak{A})\cup\sigma_c(\mathfrak{A})$. The 'approximate residual spectrum' $\sigma_\varrho(\mathfrak{A})$ is $\sigma_c(\mathfrak{A})\cup\sigma_r(\mathfrak{A})$. The 'spectrum' $\sigma(\mathfrak{A})$ is $\sigma_p(\mathfrak{A})\cup\sigma_c(\mathfrak{A})\cup\sigma_r(\mathfrak{A})$.

Proposition 2. 1. *$0\in\sigma_\pi(\mathfrak{A})$ if and only if there exist $x_n, y_n\in\mathfrak{H}$ such that $(y_n, x_n)\in\mathfrak{A}$, $\|x_n\|=1$, and $y_n\to 0$. $0\in\sigma_\varrho(\mathfrak{A})$ if and only if $\mathfrak{R}(\mathfrak{A})\neq\mathfrak{H}$.*

The first property is the familiar justification for the term 'approximate point spectrum', cf. [10]. The second has no counterpart under the usual definitions.

Proposition 2. 2. *$\sigma_\pi(\mathfrak{A})$ is closed.*

I give only the key step in the (familiar) proof: suppose we have chosen, as we may if $0\in\overline{\sigma_\pi(\mathfrak{A})}$, numbers $z_n\in\sigma_\pi(\mathfrak{A})$ with $z_n\to 0$, and elements $(y_n, x_n)\in\mathfrak{A}-z_n\mathfrak{J}$ with $\|x_n\|=1$ and $\|y_n\|\to 0$. Then $0\in\sigma_\pi(\mathfrak{A})$ is established by using, as the sequence in Prop. 2. 1, $(y_n+z_nx_n, x_n)\in\mathfrak{A}$.

Definition 2. 6. The 'adjoint' $\mathfrak{A}^*$ of a relation $\mathfrak{A}$ is $(-\mathfrak{U}^{-1})^\perp$; that is, $(w,z)\in\mathfrak{A}^*$ if and only if, for all $(y,x)\in\mathfrak{A}$, $w^*x=z^*y$.

It follows easily that $(a\mathfrak{A}+b\mathfrak{J})^*=\bar{a}\mathfrak{A}^*+\bar{b}\mathfrak{J}$.

The following property of σ_p, σ_c, σ_r may justify the peculiar way I have defined the types of spectral point.

Theorem 2. 1. *$z\in\sigma_p(\mathfrak{A})$ if and only if $\bar z\in\sigma_r(\mathfrak{A}^*)$. $z\in\sigma_c(\mathfrak{A})$ if and only if $\bar z\in\sigma_c(\mathfrak{A}^*)$. Similar assertions hold for ∞ in place of z.*

Proof. The first statement follows at once from the definitions. As for ∞, it presents no special problem: merely consider $\mathfrak{A}^{-1}$ in place of $\mathfrak{A}$. It remains to prove the assertion concerning $z\in\sigma_c$, and we may evidently take $z=0$. The proof is essentially familiar.

Let P denote the projector of $\mathfrak{H}_1\oplus\mathfrak{H}_2$ onto $\mathfrak{H}_2$, and let Q denote the projector of $\mathfrak{H}_1\oplus\mathfrak{H}_2$ onto $\mathfrak{A}$. The pairs (y_n, x_n) occurring in Definition 2. 2 will be sums of $(y_n, 0)\in\mathfrak{H}_1\perp\mathfrak{H}_2$ and $(0, x_n)\in\mathfrak{H}_2$; because $x_n\perp\mathfrak{N}(\mathfrak{A})$, we have also $(0, x_n)\perp\mathfrak{H}_2\cap\mathfrak{A}$; hence $(y_n, x_n)\perp\mathfrak{H}_2\cap\mathfrak{A}$. Also it is clear that both (y_n, x_n) and $(0, x_n)$ are orthogonal to $\mathfrak{H}_1\cap\mathfrak{A}^\perp$. We now confine attention for the moment to a certain subspace $\mathfrak{K}$ of $\mathfrak{H}_1\oplus\mathfrak{H}_2$, namely,

$$\mathfrak{K}=(\mathfrak{H}_2\cap\mathfrak{A}\oplus\mathfrak{H}_1\cap\mathfrak{A}^\perp)^\perp=\mathfrak{N}(P-Q)^\perp. \tag{2. 1}$$

It reduces both P and Q.

We have seen that $(y_n, x_n)\in Q\mathfrak{K}$ and $(0, x_n)\in P\mathfrak{K}$, though both of norm $\geqq 1$, satisfy $\|(y_n, x_n)-(0, x_n)\|=\|y_n\|\to 0$. It is required to prove the existence of w_n, z_n such that $(z_n, w_n)\in(1-Q)\mathfrak{K}$ and $(z_n, 0)\in(1-P)\mathfrak{K}$, both (z_n, w_n) and $(z_n, 0)$ have norm $\geqq 1$, and yet they satisfy $\|(z_n, w_n)-(z_n, 0)\|=\|w_n\|\to 0$; for then the sequence of pairs $(-w_n, z_n)\in\mathfrak{A}^*$ will show $0\in\sigma_c(\mathfrak{A}^*)$ from Def. 2. 2.

To prove this, define, as in [7] and [14, I. 4. 6 and I. 6. 8],

$$C=(P+Q-1)^2,\quad S=(P-Q)^2$$

(commuting operators $\geqq 0$). We are already restricted (see (2. 1)) to $\mathfrak{K}=\mathfrak{N}(S)^\perp$, and it is easy to see that there is no loss in restricting further to $\mathfrak{L}=\mathfrak{K}\ominus\mathfrak{N}(C)$. ($\mathfrak{N}(C)=\mathfrak{H}_1\cap\mathfrak{A}\oplus\mathfrak{H}_2\cap\mathfrak{A}^\perp$.) Then

$$(CS)^{-\frac{1}{2}}(QP-PQ) \tag{2. 2}$$

turns out to be a unitary operator $\mathfrak{L}\to\mathfrak{L}$ taking $P\mathfrak{L}$ to $(1-P)\mathfrak{L}$ and QL to $(1-Q)\mathfrak{L}$. Therefore we are able to specify (z_n, w_n) and $(z_n, 0)$ having the properties desired: the images under (2. 2) of (y_n, x_n) and $(0, x_n)$, respectively.

The proof is complete.

Proposition 2. 3. *$\sigma(\mathfrak{A})$ is closed.*

This is immediate from Prop. 2. 2 and Thm. 2. 1.

The main relevance of mentioning the duality between point and residual spectra in the present connection is that, as will now be explained, only σ_π is involved n matters concerning the shell.

Theorem 2. 2. *The set $S\cap s(\mathfrak{A})$ consists exactly of the image of $\sigma_p(\mathfrak{A})$ under the stereographic projection τ.*

Indeed, for C this follows easily from (1. 1) and (1. 3). As for ∞, $\infty\in\sigma_p(\mathfrak{A})$ is equivalent to $(y, 0)\in\mathfrak{A}$ for some non-zero y, which has already been remarked to be equivalent to $(0, 1)\in s(\mathfrak{A})$.

Theorem 2. 3. *The set* $S\cap\overline{s(\mathfrak{A})}$ *consists exactly of the image of* $\sigma_\pi(\mathfrak{A})$ *under the stereographic projection* τ.

Proof. Part I. Let $z\in C\cap\sigma_c(\mathfrak{A})$; it is to be proved that $\tau(z)\in\overline{s(\mathfrak{A})}$. We can choose $(y_n, x_n)\in\mathfrak{A}$ such that $\|x_n\|=1$ and $y_n-zx_n\to 0$. For such a sequence, it is easy to compute that $\varphi(y_n, x_n)\to\tau(z)$.

Let $\infty\in\sigma_c(\mathfrak{A})$; it is to be proved that $\tau(\infty)\in\overline{s(\mathfrak{A})}$. The above paragraph (for $z=0$) gives a proof by exchanging x with y in (1. 3).

Part II. Assume $z\in C$, $\tau(z)\in\overline{s(\mathfrak{A})}$. This means that there exist $(y_n, x_n)\in\mathfrak{A}$ such that $\varphi(y_n, x_n)\to\tau(z)$.

If $\|x_n\|\neq o(\|y_n\|)$, there is no loss of generality in assuming that no x_n is 0; replacing (y_n, x_n) by $(y_n/\|x_n\|, x_n/\|x_n\|)$ gives still a point of $\mathfrak{A}$, having the same image under φ, so in this case there is no loss of generality in assuming that $\|x_n\|=1$. Now

$$\frac{-1+\|y_n\|^2}{1+\|y_n\|^2}\to\frac{-1+|z|^2}{1+|z|^2}$$

implies $\|y_n\|\to|z|$; then also $x_n^*y_n\to z$. This implies that $\|y_n-zx_n\|^2\to 0$. Hence $z\in\sigma_c(\mathfrak{A})$ (provided, to be sure, that $x_n\perp\mathfrak{N}(\mathfrak{A}-z\mathfrak{I})$; but if $\mathfrak{N}(\mathfrak{A}-z\mathfrak{I})$ is non-trivial then $z\in\sigma_p(\mathfrak{A})$, which is also all right).

If, on the other hand, $\|x_n\|=o(\|y_n\|)$, then

$$\frac{-\|x_n\|^2+\|y_n\|^2}{\|x_n\|^2+\|y_n\|^2}\to 1,$$

$\varphi(y_n, x_n)\to(0, 1)$, and we must be in the remaining case, $\tau(\infty)\in\overline{s(\mathfrak{A})}$. This is easily disposed of, and the proof is complete.

3. The shell and the numerical range

First, the following key observation, which follows immediately from the definitions.

Theorem 3. 1. *The shadow of the shell is the numerical range.*

The definitions to be consulted are those in § 1, and the following

Definition 3. 1. The 'numerical range' of $\mathfrak{A}$, denoted $w(\mathfrak{A})$, is

$$\{x^*y\colon \|x\|=1,\ (y, x)\in\mathfrak{A}\},$$

with ∞ adjoined in case $\infty\in\sigma_p(\mathfrak{A})$.

In light of the last theorem, the following may be regarded as a sharpening of the Hausdorff—Toeplitz theorem, which asserts that the numerical range of an operator is convex:

Theorem 3. 2. *For every pair of points of $s(\mathfrak{A})$, there is an ellipsoid (perhaps degenerate) containing them and lying in $s(\mathfrak{A})$.*

Proof. Let us normalize every vector (y, x) entering in (1. 3), $\|(y, x)\|^2 = \|x\|^2 + \|y\|^2 = 1$, so that (1. 3) reads simply $\varphi(y, x) = (2x^* y, -\|x\|^2 + \|y\|^2)$.

Now argue as in the proof of the Hausdorff—Toeplitz theorem [9]. The two given points of $s(\mathfrak{A})$ come from two unit vectors (y_1, x_1), (y_2, x_2), spanning a space $\mathfrak{S} \subseteqq \mathfrak{A}$. The set of points

$$(3.1) \qquad \big(\mathrm{Re}(2x^* y), \quad \mathrm{Im}(2x^* y), \quad -x^* x + y^* y\big)$$

in $\boldsymbol{R}^3$, as (y, x) ranges over all unit vectors in the 2-dimensional space $\mathfrak{S}$, will be shown to be an ellipsoid, and it is $\subseteqq s(\mathfrak{A})$ by definition.

Each component in (3. 1) is a hermitian form in the variable (y, x). By suitably choosing co-ordinates in $\boldsymbol{R}^3$ we may assume all have trace zero; and then by suitable choice of co-ordinates in $\mathfrak{S}$ we may write (3. 1) as

$$(3.2) \qquad \big(a_1(|\xi|^2 - |\eta|^2) + 2\,\mathrm{Re}\,(b_1 \xi\eta), \quad a_2(|\xi|^2 - |\eta|^2) + \\ + 2\,\mathrm{Re}\,(b_2 \bar{\xi}\eta), \quad a_3(|\xi|^2 - |\eta|^2)\big) \quad (|\xi|^2 + |\eta|^2 = 1),$$

with a_1, a_2, a_3 real. It is elementary to express the set of points (3. 2) explicitly as an ellipsoid, and the proof is complete.

Thus the shell is in general not convex, as the examples immediately following will illustrate, but it becomes convex if its "holes are filled up". That is, the unbounded component of the complement of the shell, is the complement of a convex set.

Proposition 3. 1. *Assume $\mathfrak{A}$ is not densely defined. It may be extended to $\mathfrak{B} = \mathfrak{A} \oplus (\mathfrak{H}_2 \ominus \mathfrak{D}(\mathfrak{A}))$ (that is, $\mathfrak{B}$ is the zero operator on $\mathfrak{D}(\mathfrak{A})^\perp$). Then $s(\mathfrak{B})$ is a union of (perhaps degenerate) ellipsoids joining points of $s(\mathfrak{A})$ with $(0, -1)$.*

Proof. Of course $s(\mathfrak{B}) \supseteqq s(\mathfrak{A})$ just because $\mathfrak{B} \supseteqq \mathfrak{A}$. The general point of $\mathfrak{B}$ is $(y, x + x')$, where $(y, x) \in \mathfrak{A}$ and $x' \in \mathfrak{D}(\mathfrak{A})^\perp$. Now use (y, x) and $(0, x')$ as the vectors spanning $\mathfrak{S}$, in the construction of Thm. 3. 2.

As a corollary of Thm. 3. 2, we have

Theorem 3. 3. *Unless $\sigma_\pi(\mathfrak{A})$ is void, every point of $\overline{s(\mathfrak{A})}$ lies on a (perhaps degenerate) ellipsoid lying in $\overline{s(\mathfrak{A})}$ and intersecting S.*

Proof. Imbed $\mathfrak{H}$ in a space $\mathfrak{H}^0$ of approximate proper vectors of $\mathfrak{A}$. This is done in the same manner as given by BERBERIAN [3] for the case of an operator. Let $\mathfrak{A}^0$ denote the corresponding relation in $\mathfrak{H}^0 \oplus \mathfrak{H}^0$. It is easy to see that $s(\mathfrak{A}^0) = \overline{s(\mathfrak{A})}$. Now $s(\mathfrak{A}^0) \cap S$ is just $\tau(\sigma_p(\mathfrak{A}^0)) = \tau(\sigma_\pi(\mathfrak{A}))$, which is not empty. Given any point of $s(\mathfrak{A}^0)$, choose an arbitrary point of $s(\mathfrak{A}^0) \cap S$, and invoke Thm. 3. 2 to show there is an ellipsoid in $s(\mathfrak{A}^0)$ joining them, as desired.

The proof could have been accomplished by approximation without mentioning $\mathfrak{H}^0$, as was done in Thm. 2. 3.

Many theorems are known drawing conclusions on the structure of an operator from failure of its numerical range to have a smooth boundary [9], [18], [12]. The key seems to be the degeneracy of the ellipsoid in Thm. 3. 2, and these theorems should have sharper forms in which the shell would figure in place of the numerical range. There should also be theorems relating the shell with dilations. A small beginning is made in § 8 of this paper.

4. Examples

In this section various results are listed, illustrating the sort of shells which occur. The justifications of the results are mostly easy, and are not given.

Example 1. Let A be a normal matrix with eigenvalues $\lambda_1, \cdots, \lambda_n$. Then $s(A)$ is the convex hull of the points $\tau(\lambda_j) \in S$.

Example 2. Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Then $s(A)$ is the set of $(\zeta, h) \in \boldsymbol{R}^3$ satisfying $\frac{1}{2}|\zeta|^2 + (h + \frac{1}{2})^2 = \frac{1}{4}$. This is a non-degenerate ellipsoid tangent to the unit sphere at the point corresponding to the eigenvalue of A, viz., $(0, -1) = \tau(0)$.

Example 3. Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$. Then $s(A)$ is the set of $(\xi, \eta, h) \in \boldsymbol{R}^3$ satisfying $\xi^2 + \eta^2 - \xi h + h^2 - \xi + h = 0$. This is a non-degenerate ellipsoid tangent to the unit spere at the two points corresponding to the eigenvalues of A, viz., $(1,0) = \tau(1)$ and $(0, -1) = \tau(0)$.

From these results, $s(A)$ can be found for all 2×2 matrices by using the results of § 5. In particular, the shell of a 2×2 matrix is a degenerate ellipsoid if and only if the matrix is normal; otherwise, a non-degenerate ellipsoid.

Example 4. Let A be a normal operator. Then $s(A)$ is the convex hull of $\tau(\sigma(A))$, except that the points $\tau(\sigma(A) \setminus \sigma_p(A))$ are not in $s(A)$. In particular, the shell of the bilateral shift is $\{(\zeta, 0): |\zeta| < 1\}$.

Example 5. Let V denote the unilateral shift. Then $s(V) = \{(\zeta, 0): |\zeta| < 1\}$, but $s(V^*) = \{(\zeta, h) \in B: |\zeta| < 1, h \leqq 0\}$. (To prove all the points with $h < 0$ are in

$s(V^*)$, it suffices to consider sequences of the form $(\delta, \zeta, \zeta^2, \zeta^3, \cdots).)$] Note that V^{-1} is a not-everywhere-defined, bounded operator. It happens also that V and V^{-1} have the same shell. Nonetheless, the two operators are quite different: $\sigma_r(V) = \{z: |z|<1\}$, while $\sigma_r(V^{-1}) = \{z: |z|>1\} \cap \{\infty\}$. This is a pleasingly simple instance of Thm. 5. 1 and Prop. 6. 4 below. Note also that V^* happens to be the same as the extension of V^{-1} obtained as in Prop. 3. 1; the relationship between their shells is as described there.

The preceding examples show some empty σ_p, hence some shells disjoint from S. It is easy to see from familiar results that if $\mathfrak{D}(\mathfrak{A})$ is dense then $\sigma_\pi(\mathfrak{A})$ is not empty, hence $\overline{s(\mathfrak{A})}$ is not disjoint from S. However, we have

Example 6. In 2-space, let $Ax_2 = x_1 \perp x_2$, and let A be otherwise undefined. Then $s(A) = \{(0, 0)\}$, $\sigma_\pi(A)$ is void, $\sigma(A) = \sigma_r(A) = \bar{C}$.

5. Transformation properties of the shell

The facts to be presented will emerge as immediate consequences of the definitions, once these have been expressed in the notations appropriate to the purpose.

Möbius transformations $\bar{C} \to \bar{C}$ are most simply expressed in terms of complex-homogeneous co-ordinates. Parametrize $\bar{C}$ in terms of pairs (z_1, z_2), with z represented by $(z, 1)$ and ∞ by $(1, 0)$. Then the linear transformation

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad (ad-bc \neq 0) \tag{5. 1}$$

of the space of pairs represents the Möbius transformation usually written

$$z \to \mu(z) = \frac{az+b}{cz+d}.$$

The Riemann sphere $S = \tau(\bar{C})$ may also be written in a way better suited to present purposes:

$$\tau(z_1, z_2) = (|z_1|^2, z_1\bar{z}_2, z_2\bar{z}_1, |z_1|^2). \tag{5. 2}$$

(To give these second and third components, not necessarily real but conjugate to each other, is more convenient than the equivalent procedure of giving the corresponding real and imaginary parts. Indeed, it makes possible the especially simple form of (5. 4) below.) These are real-homogeneous co-ordinates $(\delta_1, \delta_2, \delta_3, \delta_4)$ for the points of S, which will be the locus of non-zero quadruples with $\bar{\delta}_2 = \delta_3$, $\delta_2\delta_3 = \delta_1\delta_4$. In terms of the previous co-ordinates (ζ, h), S was the locus of $|\zeta|^2 + h^2 = 1$. Composing (1. 2) with (5. 2) it turns out that

$$(\zeta, h) \quad \text{may be written} \quad (1+h, \zeta, \bar{\zeta}, 1-h). \tag{5. 3}$$

The representation (5. 3) can be applied to all of B, that is, to all (ζ, h) with $|\zeta|^2+h^2\leqq 1$. It gives all non-zero $(\delta_1, \delta_2, \delta_3, \delta_4)$ such that $\bar{\delta}_2=\delta_3$, $\delta_2\delta_3\leqq\delta_1\delta_4$. Planes, in the δ-co-ordinates, have homogeneous equations $\sum\bar{\Delta}_j\delta_j=0$, subject to Δ_1, Δ_4 real, and $\bar{\Delta}_2=\Delta_3$. Halfspaces have representations $\sum\bar{\Delta}_j\delta_j\geqq 0$, subject to the same conditions, with the additional one: $\delta_1+\delta_4$ non-negative. To handle linear inequalities we need positively homogeneous co-ordinates; Δ and $-\Delta$ will not give the same half-space.

Now the Möbius transformation μ above gives a transformation $S\to S$ taking $\tau(z_1, z_2)$ to $\tau(az_1+bz_2, cz_1+dz_2)$. Clearly the δ-co-ordinates are transformed by the matrix

(5. 4)
$$\begin{pmatrix} a\bar{a} & a\bar{b} & b\bar{a} & b\bar{b} \\ a\bar{c} & a\bar{d} & b\bar{c} & b\bar{d} \\ c\bar{a} & c\bar{b} & d\bar{a} & d\bar{b} \\ c\bar{c} & c\bar{d} & d\bar{c} & d\bar{d} \end{pmatrix}.$$

The correspondence between (5. 1) and (5. 4) is a representation of $GL(2, \mathbf{C})$, viz., the tensor product of the natural representation and its complex conjugate. Specializing (5. 1) to have determinant 1 (as for present purposes we may), we give (5. 4) determinant 1 also. These matrices comprise the proper Lorentz group in its natural representation by linear transformations of $\mathbf{R}^4$ (though not in the customary co-ordinate system). Indeed this group is well known to be isomorphic to the group of Möbius transformations (e.g., [5, § 17]).

We are interpreting the δ_j as homogeneous co-ordinates, so for us the matrices (5. 4) give a representation ϱ of the group by non-linear transformations of $\mathbf{R}^3$. The invariant cone (or half-cone!) under the Lorentz group is for us replaced by S in $\mathbf{R}^3$, which was known to be invariant from the start. Now note that (corresponding to the fact that the "future" is invariant under Lorentz transformations) all the $\varrho(\mu)$ also take the ball B onto itself.

On $B\setminus S$ this gives a group of plane-preserving transformations — the group of congruences of the Beltrami model [4, § 16. 2] of hyperbolic 3-space. Cf. [19, § 15, ex. 5], [2, Abschnitt IV]. The rigid rotations of S are given in the particular case $d=\bar{a}$, $c=-\bar{b}$. Since all the $\varrho(\mu)$ are plane-preserving, those which are rigid rotations of S will also be ordinary rotations of all of B.

The one ingredient still lacking is the definition of a Möbius transformation of a relation. For μ as above, define

(5. 5)
$$\mu(\mathfrak{A})=\{(ay+bx, cy+dx):(y, x)\in\mathfrak{A}\}.$$

This agrees with the usual definition $\mu(A)=(aA+b)(cA+d)^{-1}$ for the most important case, that in which A is an everywhere defined operator and $-d/c$ is not in its spectrum. On the other hand, it agrees for all $\mathfrak{A}$ with the usual definition of $\mathfrak{A}^{-1}=\{(x, y):(y, x)\in\mathfrak{A}\}$.

Theorem 5. 1. *The shell of $\mu(\mathfrak{A})$ is obtained from the shell of $\mathfrak{A}$ by the transformation $\varrho(\mu)$.*

Proof. Write everything in the δ-co-ordinates and this falls out. The shell of $\mathfrak{A}$ is

$$\{(\|y\|^2, x^*y, y^*x, \|x\|^2):(y, x)\in A\}.$$

Similarly, $s(\mu(\mathfrak{A}))$ is

$$\{(\|ay+bx\|^2, (\bar{c}y^*+\bar{d}x^*)(ay+bx), (\bar{a}y^*+\bar{b}x^*)(cy+dx), \|cy+dx\|^2):(y, x)\in\mathfrak{A}\}.$$

Clearly the quadruples in $s(\mu(\mathfrak{A}))$ are got from those in $s(\mathfrak{A})$ by applying the matrix (5. 4).

This proves the theorem as stated, but leaves open the question whether, for two Möbius transformations μ_1, μ_2, we need have $\mu_1(\mu_2(\mathfrak{A}))=\mu_1\circ\mu_2(\mathfrak{A})$. This is one of a whole class of questions which will be treated in the next section.

6. Rational functional calculus of relations

The main result of the paper was stated in the introduction for operators only. Before stating it for general closed linear relations $\mathfrak{A}$, spectral sets must be defined for them, and this means that we need discussion of rational functions of relations.

The definition of powers is standard:

$$\mathfrak{A}^2=\mathfrak{A}\circ\mathfrak{A}=\{(y, x):(\exists w)(y, w)\in\mathfrak{A}\,\&\,(w, x)\in\mathfrak{A}\},$$

etc. Similarly for linear combinations:

$$a_1\mathfrak{A}_1+a_2\mathfrak{A}_2=\{(a_1y_1+a_2y_2, x):(y_1, x)\in\mathfrak{A}_1\,\&\,(y_2, x)\in\mathfrak{A}_2\}.$$

This defines polynomials. Arens [1] explains their properties. In particular, it is important to use only polynomials with leading coefficient non-zero, because $0\mathfrak{A}$ (say) may not be the zero relation: it may be properly contained in it. With this understanding, Arens proves the following:

If p is a polynomial

$$p(z)=a_nz^n+\cdots+a_1z+a_0 \qquad (a_n\neq 0),$$

then $(y, x)\in p(\mathfrak{A})$ if and only if there exist $w_0=x, w_1, \cdots, w_n$ with $(w_j, w_{j-1})\in\mathfrak{A}$ and $y=\sum_{j=0}^{n} a_jw_j$. If p and q are polynomials then $p(\mathfrak{A})q(\mathfrak{A})=(pq)(\mathfrak{A})$ — this despite the fact that the distributive law fails in general. If p and q are polynomials then $p(q(\mathfrak{A}))=p\circ q(\mathfrak{A})$. If p and q are polynomials such that, in forming $p+q$, the leading terms do not cancel, then $(p+q)(\mathfrak{A})=p(\mathfrak{A})+q(\mathfrak{A})$.

We want to extend the ideas to rational functions. Again, care is required because (for instance) $\mathfrak{A}\mathfrak{A}^{-1}$ need not be $\mathfrak{I}$, but may be a proper subset. There are therefore two inequivalent possible definitions, of which I choose this one (already used in (5. 5) in the case of a Möbius transformation):

Definition 6. 1. Let $f=p/r$ be a rational function, where p and r are polynomials without common factor: $p(z)=a_n z^n+\cdots+a_1 z+a_0$ $(a_n\neq 0)$, $r(z)= =b_m z^m+\cdots+b_1 z+b_0$ $(b_m\neq 0)$. Then $(y, x)\in f(\mathfrak{A})$ if and only if there exist w_k $(k=0, 1, \cdots, \max\{n, m\})$ with $(w_k, w_{k-1})\in\mathfrak{A}$, such that $y=\sum_{j=0}^{n} a_j w_j$ and $x=\sum_{j=0}^{m} b_j w_j$.

By ARENS's result just quoted, this gives the usual result in case f is a polynomial $(r=1)$. For general polynomial r, the $f(\mathfrak{A})$ determined by Def. 6. 1 is $\subseteqq p(\mathfrak{A})(r(\mathfrak{A}))^{-1}$ and the inclusion may be proper, though equality holds for $\mathfrak{A}$ an operator. We do have some of the expected relations.

Proposition 6. 1. *Let* $p=q_1 r+r_1$, *with* $m_1=\text{degree}\,(r_1)<n=\text{degree}\,(p)$. *Then* $(p/r)(\mathfrak{A})=q_1(\mathfrak{A})+(r_1/r)(\mathfrak{A})$.

(The hypothesis implies that $m=\text{degree}\,(r)\leqq n$, but not that $m_1<m$.)

Proof. Let $r_1(z)=\sum_{j=0}^{m_1} c_j z^j$, so that $(q_1 r)(z)=\sum_{j=0}^{n}(a_j-c_j)w_j$. The general pair in $q_1(\mathfrak{A})+(r_1/r)(\mathfrak{A})$ is expressible as $(u+v, x)$ where $(u, x)\in q_1(\mathfrak{A})$, and $v=\sum c_j w_j$, $x=\sum b_j w_j$, for some $w_0, \cdots, w_m$, such that $(w_j, w_{j-1})\in\mathfrak{A}$. From $(u, x)\in q_1(\mathfrak{A})$ and $(x, w_0)\in r(\mathfrak{A})$ follows that $(u, w_0)\in(q_1 r)(\mathfrak{A})$, by one of ARENS's results just cited. Indeed, by reference to the proof of that result [1, 2. 3] we see that we can even use

$$u=\sum_{j=0}^{n}(a_j-c_j)w_j \qquad (w_j, w_{j-1})\in\mathfrak{A},$$

with the same w_j as before as far as $j=m'$. But then $(u+v, x)\in(p/r)(\mathfrak{A})$. This proves „$\supseteqq$" in the conclusion.

In the other direction no subtleties are involved: we are given $y=\sum a_j w_j$, $x=\sum b_j w_j$ as in Def. 6. 1, and we define $u=\sum(a_j-c_j)w_j$, $v=\sum c_j w_j$. By definition $(v, x)\in(r_1/r)(\mathfrak{A})$, it remains to prove that $(u, x)\in q_1(\mathfrak{A})$. Let $q_1(z)=\sum d_j z^j$, then for each i, $a_i-c_i=\sum_j d_{i-j}b_j$. Define $x_k=\sum_j b_j w_{j+k}$ $(k=0, 1, \cdots, \text{degree}\,(q_1))$, so that $x_0=x$, $(x_k, x_{k-1})\in\mathfrak{A}$, and $\sum d_k x_k=u$. Then by definition $(u, x)\in q_1(\mathfrak{A})$.

In the following propositions, Möbius transformations $\mu(z)=\dfrac{az+b}{cz+d}$ again play a special role.

Proposition 6. 2. $\mu(\mathfrak{A})$ *is closed.*

(We are assuming $\mathfrak{A}$ closed throughout.)

Proof. To say that $(y_\nu, x_\nu)\in\mu(\mathfrak{A})$ is to say that $y_\nu=aw_{1\nu}+bw_{0\nu}$ and $x_\nu=cw_{1\nu}+dw_{0\nu}$, with $(w_{1\nu}, w_{0\nu})\in\mathfrak{A}$. We assume in addition that $y_\nu\to y$, $x_\nu\to x$. Then solving for $w_{1\nu}$ and $w_{0\nu}$ in terms of y_ν and x_ν, we deduce that $w_{1\nu}\to w_1$ and $w_{0\nu}\to w_0$ for w_i such that $y=aw_1+bw_0$, $x=cw_1+dw_0$. Since $\mathfrak{A}$ is closed, $(w_1, w_0)\in\mathfrak{A}$. This completes the proof.

I don't know how far this remains true when μ is replaced by the more general f considered earlier; cf. [13, § 2. 16], [1, 3. 8].

Proposition 6. 3. $(f\circ\mu)(\mathfrak{A})=f(\mu(\mathfrak{A}))$.

Proof. This is clear when μ is just multiplication by a constant. When μ is a translation, $\mu(z)=z+b$, a calculation is needed which I will only summarize: in $y=\sum a_j w_j$ change to an expression $y=\sum a'_k w'_k$ by the substitution $w'_k=\sum_j\binom{k}{j}b^{k-j}w_j$ (and similarly for x, of course), and make the verification that $(w'_k, w'_{k-1})\in\mathfrak{A}+b\mathfrak{I}$.

When μ is reciprocation, $\mu(z)=z^{-1}$, there is again a little calculation. It is convenient to depart from previous practice and write $p(z)=\sum_{j=0}^{n} a_j z^j$, $r(z)=\sum_{j=0}^{n} b_j z^j$, where not both a_n and b_n are 0 and not both a_0 and b_0 are 0. With this convention there is notational symmetry between f and $f\circ\mu$, and reference to definitions will verify the conclusion.

The observation that any Möbius transformation is obtained from these types by composition, completes the proof.

Proposition 6. 4. (Detailed spectral mapping theorem for Möbius transformations.) $\sigma_p(\mu(\mathfrak{A}))=\mu(\sigma_p(\mathfrak{A}))$, $\sigma_c(\mu(\mathfrak{A}))=\mu(\sigma_c(\mathfrak{A}))$, $\sigma_r(\mu(\mathfrak{A}))=\mu(\sigma_r(\mathfrak{A}))$.

Again, we are at liberty to consider simple special Möbius transformations and then compose them to give the general μ.

Under linear transformation, σ_p, σ_c, and σ_r all behave as desired by Def. 2. 4, with the exception of ∞. Assume $\infty\in\sigma_p(\mathfrak{A})$ and let us prove $\infty\in\sigma_p(\mathfrak{A}+b\mathfrak{I})$. The assumption is equivalent to the existence of non-zero x such that $(x, 0)\in\mathfrak{A}$; but then $(x, 0)\in\mathfrak{A}+b\mathfrak{I}$ as well, and this gives the conclusion. Similarly for $\infty\in\sigma_c(\mathfrak{A})$. That $\infty\in\sigma_r(\mathfrak{A})$ entails $\infty\in\sigma_r(\mathfrak{A}+b\mathfrak{I})$ is a consequence of the remark following Def. 2. 4.

If we consider reciprocation, it is the λ other than 0 and ∞ which require checking. Assume $\lambda\in\sigma_c(\mathfrak{A})$, so that there are $(y_\nu, x_\nu)\in\mathfrak{A}$ such that $\|x_\nu\|=1$ but $\|y_\nu-\lambda x_\nu\|\to 0$ (so that $\|y_\nu\|\not\to 0$; without loss of generality, $y_\nu\neq 0$). Let $x'_\nu=x_\nu/\|y_\nu\|$, $y'_\nu=y_\nu/\|y_\nu\|$. Clearly $(x'_\nu, y'_\nu)\in\mathfrak{A}^{-1}$, $\|y'_\nu\|=1$, and $\|x'_\nu-\lambda^{-1}y'_\nu\|\to 0$, so $\lambda^{-1}\in\sigma'_c(\mathfrak{A}^{-1})$. For σ_p, it is even easier. Finally, take $\lambda\in\sigma_r(\mathfrak{A})$. This means $\mathfrak{R}(\mathfrak{A}-\lambda\mathfrak{I})$ is not dense. Choosen non-zero z such that, for all $(y, x)\in\mathfrak{A}$, $z\perp y-\lambda z$. Then also for all $(x, y)\in\mathfrak{A}^{-1}$, $z\perp x-\lambda^{-1}y$, so $\mathfrak{R}(\mathfrak{A}^{-1}-\lambda^{-1}\mathfrak{I})$ is not dense either. This completes the proof.

Again, I don't know how much of this result holds for more general f; I have partial results. Of course it can't hold in toto. For example, it is easy to construct an operator A whose square is O and hence has void continuous spectrum, while A itself has non-void continuous spectrum. (For the case of operators see [13, Theorem 5.12. 2].)

7. The shell and spectral sets

Definition 7. 1. The 'norm' of a relation $\mathfrak{A}$ is

$$\|\mathfrak{A}\| = \sup \{\|y\|/\|x\| : (0, 0) \neq (y, x) \in \mathfrak{A}\}.$$

Definition 7. 2. Let X be a closed subset of $\overline{\boldsymbol{C}}$. X 'is a spectral set for' $\mathfrak{A}$ or 'is s.s. for' $\mathfrak{A}$) in case every rational function f having modulus $\leqq 1$ on X has also the property $\|f(\mathfrak{A})\| \leqq 1$.

It is clear by Prop. 2. 1 that $\|\mathfrak{A}\|$ is finite if and only if $\infty \notin \sigma_\pi(\mathfrak{A})$, that is, $\mathfrak{A}$ corresponds to an operator A which is bounded, and in this case $\|\mathfrak{A}\| = \|A\|$. Furthermore, as will appear in the course of developing the basic properties, statements about spectral sets for relations can be reduced to statements involving operators without much trouble. However there are two closely related virtues in the present definition: it applies to not-everywhere-defined operators; and, secondly, it brings in σ_π instead of σ (see in particular Prop. 7. 2 below).

Proposition 7. 1. *X is s.s. for $\mathfrak{A}$ if and only if $\mu(X)$ is s.s. for $\mu(\mathfrak{A})$.*

This follows easily from Prop. 6. 3.

Proposition 7. 2. *If X is s.s. for $\mathfrak{A}$ then $\sigma_\pi(\mathfrak{A}) \subseteqq X$.*

Proof. By Prop. 7. 1, it is enough to consider a special point of $\overline{\boldsymbol{C}}$. Assume, then, $\infty \in \sigma_\pi(\mathfrak{A}) \setminus X$, so that X is a compact subset of $\boldsymbol{C}$ and $\mathfrak{A}$ is not a bounded operator. To show Def. 7. 2 is not satisfied, choose $f(z) = az$, for a sufficiently small constant $a > 0$.

Theorem 7. 1. *The unit disk $D = \{z \in \boldsymbol{C} : |z| \leqq 1\}$ is s.s. for $\mathfrak{A}$ if and only if $\|\mathfrak{A}\| \leqq 1$.*

Proof. Either condition implies we are dealing with a bounded operator A. If $\mathfrak{D}(A) = \mathfrak{H}$, the theorem is just von Neumann's basic result, as in [17, § 154]. More generally, define $\tilde{A}$ as the extension of A which is zero on $\mathfrak{D}(A)^\perp$. By applying von Neumann's theorem to $\tilde{A}$, it is easy to deduce (the non-trivial half of) the present theorem for A.

This, theorem, together with Prop. 7. 1, tells which disks are s.s. for $\mathfrak{A}$. The result is familiar in general outlines: it resembles that in [17] except that no special exemptions need to be made for functions with poles in $\sigma(\mathfrak{A})$ — even in $\sigma_\pi(\mathfrak{A})$. Namely,

(i) the finite disk $\{z\in \boldsymbol{C}\colon |z-z_0|\leqq r\}$ is s.s. for $\mathfrak{A}$ if and only if $\|\mathfrak{A}-z_0\mathfrak{J}\|\leqq r$;

(ii) the complementary disk $\{z\in \boldsymbol{C}\colon |z-z_0|\geqq r\}\cup\{\infty\}$ is s.s. for $\mathfrak{A}$ if and only if $\|(\mathfrak{A}-z_0\mathfrak{J})^{-1}\|\leqq r^{-1}$;

(iii) the half-plane $\{z\in \boldsymbol{C}\colon \operatorname{Re}(\bar{\zeta}z)\geqq a\}\cup\{\infty\}$ is s.s. for $\mathfrak{A}$ if and only if it contains $w(\mathfrak{A})$.

The following theorem contains all three of these, in a geometric form, invariant under Möbius transformations of $\overline{\boldsymbol{C}}$.

Theorem 7. 2. *Let X be a disk. Then X is s.s. for $\mathfrak{A}$ if and only if the shell $s(\mathfrak{A})$ is contained in the convex hull of the stereographic projection $\tau(X)$.*

Proof. Let μ be the Möbius transformation such that $\mu(X)=D$, the unit disk. We have just seen that X is s.s. for $\mathfrak{A}$ if and only if $\|\mu(\mathfrak{A})\|\leqq 1$. Comparing Defs. 7. 1 and 1. 1, this is seen to be equivalent to saying that $s(\mu(\mathfrak{A}))$ lies in the half-space $\{(\zeta, h)\in \boldsymbol{R}^3\colon h\leqq 0\}$; and this half-space is the convex hull of $\tau(D)=\tau(\mu(X))$. To return from $\mu(\mathfrak{A})$ and $\mu(X)$ to $\mathfrak{A}$ and X, we want to consider the transformation μ' inverse to μ, and the projective transformation $\varrho(\mu')$ of $\boldsymbol{R}^3$ to which it gives rise. This was discussed in § 5. $\varrho(\mu')$ preserves B, and preserves planes; consequently it takes convex hulls to convex hulls. It takes $\tau(\mu(X))$ to $\tau(X)$. Finally, by Thm. 5. 1, it takes $s(\mu(\mathfrak{A}))$ to $s(\mathfrak{A})$. The theorem is proved.

Alternatively, I could have confined consideration to rigid rotations of B (cf. § 5) with only slight modification in the proof.

The set of disks which are s.s. for fixed $\mathfrak{A}$ is hereby represented as a closed convex set $s(\mathfrak{A})^\dagger$, the dual of $s(\mathfrak{A})$. The convexity of this set is inherent in the following easily proved fact, a generalization of the Lemma of [8, § 3. 1]: if two disks are s.s. for $\mathfrak{A}$, then so is any disk containing their intersection. The closedness of $s(A)^\dagger$ is also easy to prove directly.

However, if it was a question only of representing all disks s.s. for $\mathfrak{A}$ in a simple geometric way, the closed convex hull of $s(\mathfrak{A})$ would do exactly as well as $s(\mathfrak{A})$, for it has the same dual. It is like the situation for the numerical range $w(\mathfrak{A})$. Only $\overline{w(\mathfrak{A})}$ is needed in criterion (iii) above, telling which half-planes are spectral sets. Still for some purposes the richer structure of $w(\mathfrak{A})$ itself is interesting. The shell is like the numerical range with one dimension added (cf. Thm. 3. 1), and its structure is very much richer than that of its closed convex hull.

In spite of these remarks, Thm. 7. 2 throws emphasis on $s(\mathfrak{A})^\dagger$, and this will persist in the final section of the paper. Let me therefore say a few more words about this set.

It has already been pointed out in § 5 that half-spaces in $\boldsymbol{R}^3$ may be for present purposes conveniently represented by quadruples $\varDelta=(\varDelta_1, \varDelta_2, \varDelta_3, \varDelta_4)$; here

$\Delta_1, \Delta_4 \in \boldsymbol{R}$, $\Delta_2 = \bar{\Delta}_3 \in \boldsymbol{C}$, and if $\lambda > 0$ then $\lambda\Delta$ represents the same half-space as Δ. Namely, it is the half-space of all $(\zeta, h) \in \boldsymbol{R}^3$ such that $\Delta_1(1+h) + \bar{\Delta}_2\zeta + \bar{\Delta}_3\bar{\zeta} + \Delta_4(1-h) \geqq 0$.

Let us regard the dual $C^\dagger$ of an arbitrary set $C \subseteqq \boldsymbol{R}^3$ as the convex cone of all $\Delta \in \boldsymbol{R}^4$ representing half-spaces $\supseteqq C$. This sort of dual has been used often in the theory of convex sets [23].

Proposition 7. 3. *$\Delta \in B^\dagger$ if and only if Δ gives homogeneous co-ordinates for a point of B in the δ-representation of § 5, i.e., if and only if $\Delta_1 + \Delta_4 > 0$ and $\Delta_2\Delta_3 \leqq \Delta_1\Delta_4$.*

This is just self-duality of the unit ball; I will not bother translating the familiar proof into this notation.

8. Generalization of Berger's theorem

The most striking results relating disk spectral sets to dilation theory are the well-known theorem of Sz.-Nagy [20] on dilations of contractions, and C. Berger's theorem [11], [21] on dilations of operators with $w(A)$ in the unit disk. Sz.-Nagy and Foiaş [22] have recently given a common generalization of the two, which I will relate to the ideas of this paper. Their theorem may be stated as follows. [2])

Theorem 8. 1 (Sz.-Nagy—Foiaş). *For a bounded, everywhere-defined operator A, and for any $\varrho > 0$, consider these conditions:*

(i) *$A \in C_\varrho$, that is, there exists a unitary U on a Hilbert space $\mathfrak{K} \supseteqq \mathfrak{H}$ such that $A^n = \varrho \cdot \mathrm{pr}\, U^n$ $(n = 1, 2, \cdots)$, where* pr *denotes compression to $\mathfrak{H}$:*

(ii) *for every z with $|z| < 1$, $\|zA((\varrho-1)zA-\varrho)^{-1}\| \leqq 1$;*

(iii) *for every $x \in \mathfrak{H}$ and every $t \in [0, 1]$,*

$$2|\varrho-1|t \cdot |x^*Ax| \leqq \varrho\|x\|^2 - (2-\varrho)t^2\|Ax\|^2.$$

Conclusions: Conditions (ii) *and* (iii) *are equivalent. Condition* (i) *is equivalent to* (iii) *together with*

(iv) *$\sigma(A) \subseteqq D$, the closed unit disk.*

If $\varrho \leqq 2$, (iv) *follows from* (iii).

This formulation is not quite that of Sz.-Nagy and Foiaş; let me bridge the short gap. My (ii) is equivalent to their (5). My (iii) is obtained from their (I_ϱ), an inequality which must be asserted for every $z \in D$, by rewriting it in such a way that the phase of z need no longer be kept in view.

---

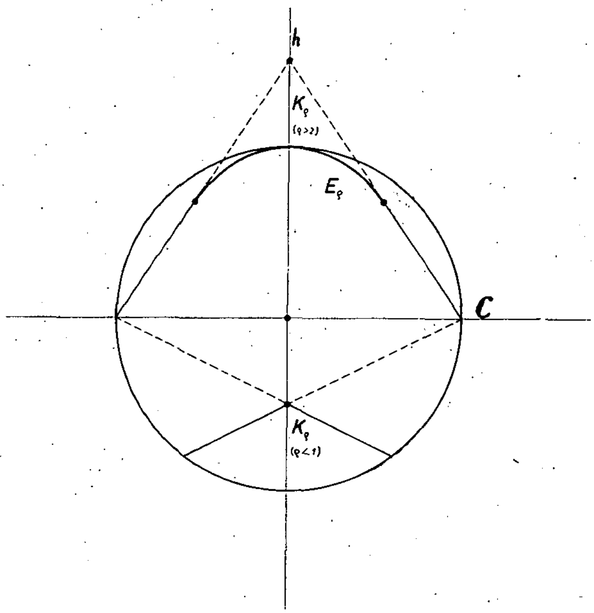
[2]) The real parameter ϱ is not to be confused with the symbol $\varrho(\mu)$ already introduced in § 5.

Now for each t, the inequality (iii) is a homogeneous linear inequality relating the moduli of the homogeneous co-ordinates of points of $s(A)$. As such, it is readily interpreted geometrically: it restricts $s(A)$ to lie in a certain solid half-cone (from now on I will say simply 'cone'). The cone may degenerate into a half-space, in particular it does so for $t=0$. To complete the account, we must examine the consequences for $s(A)$ of imposing the restrictions (iii) for all t simultaneously.

For this purpose, here are a few special notations. Let

$$f_x(t) = (2-\varrho)\|Ax\|^2 t^2 + 2|\varrho - 1| \cdot |x^* Ax| t - \varrho \|x\|^2.$$

For each non-zero x, f_x is a real quadratic polynomial, considered as a function on [0, 1]; (iii) asserts that every f_x is $\leqq 0$ on the whole interval.



Secondly, let K_ϱ denote the cone $\{(\zeta, h) \in \boldsymbol{R}^3 \colon |\varrho - 1| \cdot |\zeta| \leqq \varrho - 1 - h\}$. The assertion $f_x(1) \leqq 0$ is readily transformed into the assertion $\varphi(Ax, x) \in K_\varrho$. This can be done directly from (1. 3) (normalizing by assuming $\|x\|^2 + \|Ax\|^2 = 1$).

The first part of the picture can now be completed.

Proposition 8. 1. *For* $\varrho \leqq 2$, *A satisfies condition* (iii) *above if and only if* $s(A) \subseteqq K_\varrho$.

Proof. Refer to the definition of f_x. We have just noted that $s(A) \subseteqq K_\varrho$ if and only if $f_x(1) \leqq 0$ for all x; and $f_x(0) \leqq 0$ in any case. But f_x is a quadratic polynomial with leading coefficient $\geqq 0$; therefore it is non-positive at the endpoints of the interval [0, 1] if and only if it is non-positive throughout. This gives the equivalence.

K_ϱ is of course symmetric about the h-axis, and has apex at $h=\varrho-1$. Its generators pass through the equatorial points $(e^{i\theta}, 0)$; but for $\varrho<1$, the equator lies in the other nappe of the cone, and K_ϱ itself lies entirely below the equatorial plane. For $\varrho=1$ (SZ.-NAGY's case), K_ϱ is the half-space $h\leqq 0$. For $\varrho=2$ (BERGER's case), the apex is at the north pole, and $s(A)\subseteqq K_\varrho$ is equivalent to $w(A)\subseteqq D$ by Thm. 3. 1.

For $\varrho>2$, the apex of K_ϱ is above the north pole. But then further restrictions on $s(A)$ result from (iii).

Let E_ϱ be obtained by removing from K_ϱ all points lying between the apex and the upper half of the ellipsoid $(\varrho-1)^2|\zeta|^2=\varrho(\varrho-2)(1-h^2)$. This ellipsoid, in addition to being evidently symmetric with respect to the h-axis and with respect to the ζ-plane, is tangent to ∂K_ϱ. The definition of E_ϱ requires the following interpretation. The "upper half" of the ellipsoid is the portion above the circle of tangency.

The ellipsoid lies entirely in B, and is tangent to S at the poles. Hence in $h>0$, E_ϱ has no points in common with S except (0, 1).

Proposition 8. 2. *For $\varrho>2$, A satisfies condition* (iii) *above if and only if* $s(A)\subseteqq E_\varrho$.

Proof. The cones to which $s(A)$ is restricted by (iii) are as follows: for $t=0$, the half-space $h\leqq 1$; for $t=1$, K_ϱ; and for intermediate t, the intermediate cones tangent to the ellipsoid. To see this, one reduces by symmetry to consideration of $\zeta>0$, and then makes an elementary computation which will not be reproduced here. Now the equivalence of $s(A)\subseteqq E_\varrho$ becomes clear.

Proposition 8. 3. *In Thm.* 8. 1, *condition* (iv) *follows from* (iii) *in case* $\varrho>2$ *also. Hence* (i) *and* (iii) *are equivalent for all values of* ϱ.

Proof. For a bounded, everywhere-defined operator, $\sigma(A)\subseteqq D$ is known to follow from $\sigma_\pi(A)\subseteqq D$. By Thm. 2. 3 and the above description of E_ϱ, any $\mathfrak{A}$ with $s(\mathfrak{A})\subseteqq E_\varrho$ must have $\sigma_\pi(\mathfrak{A})\subseteqq D\cup\{\infty\}$. Here ∞ is ruled out because we are in the bounded, single-valued case. The conclusion therefore follows from Prop. 8. 2.

These ideas lead to the following question: "If $s(\mathfrak{A})\subseteqq K_\varrho$ if $\varrho\leqq 2$ (or $\subseteqq E_\varrho$ if $\varrho>2$), what can we conclude about dilations of $\mathfrak{A}$?" They do not, however, lead to an answer. The difficulties arise even for $\varrho<2$; although we then know we have to deal with operators, we do not know that they are everywhere defined. Some new idea seems to be needed to cope with this dilation problem.

References

[1] R. Arens, Operational calculus of linear relations, *Pacific J. Math.*, **11** (1961), 9—23.
[2] R. Baldus, *Nichteuklidische Geometrie. Hyperbolische Geometrie der Ebene*, Sammlung Göschen 970 (Berlin—Leipzig, 1927).
[3] S. K. Berberian, Approximate proper vectors, *Proc. Amer. Math. Soc.*, **13** (1962), 111—114.
[4] H. S. M. Coxeter, *Introduction to geometry* (New York, 1961).
[5] ——— The inversive plane and hyperbolic space, *Abh. Math. Sem. Univ. Hamburg*, **29** (1966), 217—242.
[6] C. W. Curtis and I. Reiner, *Representation theory of finite groups and associative algebras* (New York, 1962).
[7] Ch. Davis, Separation of two linear subspaces, *Acta Sci. Math.*, **19** (1958), 172—187.
[8] Ch. Davis and D. G. Rider, Spectral sets and numerical range, *Revue Roumaine de Math. Pures et Appl.*, **10** (1965), 125—131.
[9] W. F. Donoghue, On the numerical range of a bounded operator, *Michigan Math. J.*, **4** (1957), 261—263.
[10] P. R. Halmos, *Introduction to Hilbert space and the theory of spectral multiplicity* (New York, 1951).
[11] ——— Positive definite sequences and the miracle of *w*. Mimeographed lecture notes, University of Michigan, 1965.
[12] S. Hildebrandt, Über den numerischen Wertebereich eines Operators, *Math. Ann.*, **163** (1966), 230—247.
[13] E. Hille and R. S. Phillips, *Functional analysis and semigroups*, American Mathematical Society Colloquium Publications, Vol. 31 (Providence, R. I., 1957).
[14] T. Kato, *Perturbation theory for linear operators* (Berlin, 1967).
[15] J. von Neumann, Über adjungierte Funktionaloperatoren, *Annals of Math.*, **33** (1932), 294—310.
[16] ——— Eine Spektraltheorie für allgemeine Operatoren eines unitären Raumes, *Math. Nachr.*, **4** (1950/51), 258—281.
[17] F. Riesz and B. Sz.-Nagy, *Leçons d'analyse fonctionnelle*, 2nd ed. (Budapest, 1953).
[18] M. Schreiber, Numerical range and spectral sets, *Michigan Math. J.*, **10** (1963), 283—288.
[19] H. Schwerdtfeger, *Geometry of complex numbers* (Toronto, 1962).
[20] B. Sz.-Nagy, Sur les contractions de l'espace de Hilbert, *Acta Sci. Math.*, **15** (1953), 87—92.
[21] ——— Positiv-definite, durch Operatoren erzeugte Funktionen, *Wiss. Zeitschrift Techn. Univ. Dresden*, **15** (1966), 219—222.
[22] B. Sz.-Nagy and C. Foiaş, On certain classes of power-bounded operators in Hilbert space, *Acta Sci. Math.*, **27** (1966), 17—25.
[23] F. A. Valentine, *Convex sets* (New York, 1964).

(Received May 9, 1967)

The inner function in Rota's model

By J. W. MOELLER in Chicago (Illinois, U. S. A.)

Let K be a Hilbert space of dimension $\aleph_0$ with inner product $(\cdot,\cdot)$, and let $H^2(K)$ denote the Hardy class of vector-valued functions

$$k(z) = \sum_{n=0}^{\infty} k_n z^n \qquad \left(k_n \in K;\ \sum_{n=0}^{\infty} (k_n, k_n) < \infty\right).$$

An inner product for these Hardy functions can be defined by setting

$$\langle k(z), h(z) \rangle = \sum_{n=0}^{\infty} (k_n, h_n),$$

and $H^2(K)$ becomes a Hilbert space in its own right under this new inner product. It is well known from the work of BEURLING, LAX, and HALMOS (see [2], pp. 115—116) that every closed subspace of $H^2(K)$ which is invariant under multiplication by z has a representation of the form $G\,H^2(K)$, where

$$G(z) = \sum_{n=0}^{\infty} G_n z^n$$

and the Taylor coefficients of G are linear operators from K into K. In addition, the operator norm of G is bounded from above by one, and the radial limits of G on the boundary of the unit disc are equal almost everywhere to partial isometries. (Such functions are commonly called "inner functions" in the literature.) If S^* designates the operation of multiplication by z in $H^2(K)$, a straightforward calculation reveals that the adjoint of this operation is given by

$$(Sh)(z) = z^{-1}(h(z) - h(0)).$$

Henceforth we will call a closed subspace of $H^2(K)$ which is invariant under S a *left translation invariant subspace*, and a closed subspace which is invariant under S^* a *right translation invariant subspace*. It is not difficult to show that the orthogonal complement of a right translation invariant subspace is a left translation invariant subspace, and conversely.

In [3], G. C. ROTA established the following interesting result.

Theorem. *Let $A: K \to K$ be a bounded linear operator whose spectrum is contained in the interior of the unit disc. Then the set*

$$L_A = \{(I - zA)^{-1}k \mid k \in K\}$$

is a left translation invariant subspace of $H^2(K)$ and S acting on L_A is similar to A acting on K.

According to the Beurling—Lax theorem, we may write

$$L_A = H^2(K) \ominus G\, H^2(K)$$

for some G whose Taylor coefficients depend only on A. Whenever $\|A\| < 1$, Helson proved ([1], pp. 104—106) that G is always equal to a unitary operator on the rim of the unit disc, and he further derived the explicit formula

$$G(z) = G_0 + z(I - zA)^{-1}G_1.$$

He did not, however, relate the operators G_1 and G_0 to A in any way. Our aim here is to determine this relationship.

We begin by computing the orthogonal projection of the constant functions in $H^2(K)$ onto L_A in two different ways. If $k \in K$, we have

$$\langle (I - G(z)G^*(0))k,\, z^n G(z)G^*(0)k\rangle = \langle k,\, z^n G(z)G^*(0)k\rangle - \langle G(z)G^*(0)k,\, z^n G(z)G^*(0)k\rangle$$
$$= \langle k,\, z^n G(z)G^*(0)k\rangle - \langle G^*(0)k,\, z^n G^*(0)k\rangle = 0 \quad \text{for} \quad n = 0, 1, \ldots .$$

From this relation and the simple identity

$$k = (I - G(t)G^*(0))k + G(z)G^*(0)k$$

we quickly deduce that

$$Pk = (I - G(z)G^*(0))k,$$

where P denotes the orthogonal projection onto L_A.

To complete the last part of our task, we will express P in terms of A alone. If we set

$$\tilde{A} = \sum_{n=0}^{\infty} A^{*n} A^n,$$

it follows immediately that

$$\langle k,\, (I - zA)^{-1}f\rangle \langle (I - zA)^{-1}f,\, (I - zA)^{-1}f\rangle^{-\frac{1}{2}} = (k, f)(\tilde{A}f, f)^{-\frac{1}{2}}.$$

For fixed k, the right hand side of the preceding expression assumes its maximum when $f = \tilde{A}^{-1}k$, so we conclude that

$$Pk = (I - zA)^{-1}\tilde{A}^{-1}k.$$

After identifying the Taylor coefficients in HELSON's formula, we find

(1) $$I-G_0G_0^*=\tilde{A}^{-1} \quad \text{and} \quad G_1G_0^*=-A\tilde{A}^{-1}.$$

Since $G(z)$ is a unitary operator on the boundary of the unit disc,

(2) $$G^*(e^{i\theta})G(e^{i\theta})=I \quad \text{and} \quad G(e^{i\theta})G^*(e^{i\theta})=I.$$

Setting $\theta=0$ in the last identity gives

$$(G_0+(I-A)^{-1}G_1)(G_0^*+G_1^*(I-A^*)^{-1})=I,$$

which, together with (1), implies

$$G_1G_1^*=(I-A)\tilde{A}^{-1}(I-A^*)+(I-A)\tilde{A}^{-1}A^*+A\tilde{A}^{-1}(I-A^*)^{-1}.$$

Thus we finally have

(3) $$G_1G_1^*=\tilde{A}^{-1}-A\tilde{A}^{-1}A^*.$$

The first equation in (2) may be rewritten in the form

$$I=G_0^*G_0+G_1^*\tilde{A}G_1,$$

and premultiplication by G_0 gives

$$G_0=(G_0G_0^*)G_0+(G_0G_1^*)\tilde{A}G_1.$$

In other words, $(I-G_0G_0^*)G_0=(G_0G_1^*)\tilde{A}G_1$, so

(4) $$G_0=-A^*\tilde{A}G_1.$$

Hence, $I=G_1^*(\tilde{A}AA^*\tilde{A}+\tilde{A})G_1$ and we infer that the operator

(5) $$U=(\tilde{A}AA^*\tilde{A}+\tilde{A})^{\frac{1}{2}}G_1$$

is an isometry.

An easy application of the identity $\tilde{A}=I+A^*\tilde{A}A$ reveals that

$$(\tilde{A}^{-1}-A\tilde{A}^{-1}A^*)(\tilde{A}AA^*\tilde{A}+\tilde{A})=I \quad \text{and} \quad (\tilde{A}AA^*\tilde{A}+\tilde{A})(\tilde{A}^{-1}-A\tilde{A}^{-1}A^*)=I.$$

We now see from (3) and (5) that U is actually a unitary operator and

(6) $$G_1=(\tilde{A}^{-1}-A\tilde{A}^{-1}A^*)^{\frac{1}{2}}U.$$

Equations (4) and (6) thus determine the inner function associated with L_A up to multiplication on the right by a constant unitary factor.

Bibliography

[1] H. Helson, *Lectures on Invariant Subspaces* (New York, 1964).
[2] K. Hoffman, *Banach Spaces of Analytic Functions* (New Jersey, 1962).
[3] G. C. Rota, On Models for Linear Operators, *Comm. Pure Appl. Math.*, **13** (1960), 469—472.

(Received June 9, 1967)

On interpolation functions. II

By J. PEETRE in Lund (Sweden)

The purpose of this note is to indicate a certain simplification of the results of [2]. (We use throughout the terminology of that note.)

If $h=h(u)$ is a positive Borel measurable function on $(0, \infty)$ we agree to say that h is *pseudo-concave* if it is equivalent to a concave function, i.e. if there exist a concave function h_0 and a constant C such that $C^{-1}h_0(u) \leqq h(u) \leqq Ch_0(u)$ for all u.

Lemma. *h is pseudo-concave if and only if for some C holds*

(1) $$h(v) \leqq C \max(1, v/u)\, h(u) \quad \textit{for all} \quad u, v.$$

Proof (necessity). It suffices to show that (1) holds for h_0. But from the concavity it follows that $h_0(u)$ is increasing but $\frac{h_0(u)}{u}$ is decreasing. Therefore (1) follows with $C=1$.

Proof (sufficiency). Let $\alpha_i \geqq 0$, $\sum \alpha_i = 1$, $u = \sum \alpha_i u_i$ (finite sums). Using (1) we obtain

$$\Sigma\alpha_i h(u_i) \leqq C\Sigma\alpha_i \max\left(1, \frac{u_i}{u}\right) h(u) \leqq C\left(\Sigma\alpha_i + \Sigma \frac{\alpha_i u_i}{u}\right) h(u) \leqq 2Ch(u).$$

This establishes the pseudo-concavity because we may now take

$$h_0(u) = \sup \sum \alpha_i h(u_i),$$

which is by the way the least concave majorant of h.

Corollary. *If $h(u)$ is pseudo-concave so is $k(u) = (h(u^r))^{\frac{1}{r}}$ where r is any real number $\neq 0$.*

Proof. Obvious. — This corollary was obtained in [2] as a by-product of the main result there which we now reformulate as follows.

Theorem. *A function $H = H(z_0, z_1)$ homogeneous of degree 1 and satisfying condition (2) in [2] is an interpolation function (of any power p) if and only if it is of the form $H(z_0, z_1) = z_0 h(z_1/z_0)$ with h pseudo-concave.*

Proof (necessity). Consider the space $X=\{x, y\}$ provided with the measure μ such that each of the two points x and y carries the mass 1. Take

$$\zeta_0(x)=1, \quad \zeta_0(y)=1; \quad \zeta_1(x)=u, \quad \zeta_1(y)=v$$

and define a linear mapping π by $\pi a(x)=0$, $\pi a(y)=a(x)$. The corresponding operator norms of π (considered as a mapping from $L^p_{\zeta_0}$ into $L^p_{\zeta_0}$ etc.) are

$$M_0=1, \qquad M_1=\frac{v}{u}, \qquad M=\frac{h(v)}{h(u)}.$$

By the inequality (*cf.* inequality (1) in [2]) $M \leqq C \max(M_0, M_1)$ we now get

$$\frac{h(v)}{h(u)} \leqq C \max\left(1, \frac{v}{u}\right)$$

which establishes (1).

Proof (sufficiency). By the lemma H is equivalent to a function of the form (4) in [2]. Therefore we may use the corresponding part of the proof of theorem 1 in [2] (*cf.* in particular pp. 168—169).

We conclude by illustrating our new result in a concrete case (*cf.* [1]).

Let $X=(0, \infty)$, $\mu=$ Haar measure, $\zeta_0(x)=1$, $\zeta_1(x)=e^{x^\alpha}$ $(\alpha>0)$. With

$$h(u) = e^{(\log u)^\lambda}, \quad \lambda = \frac{\beta}{\alpha} \qquad (\alpha > \beta > 0)$$

we then have

$$e^{x^\beta} = h(e^{x^\alpha}) = \zeta_0 h\left(\frac{\zeta_1}{\zeta_0}\right).$$

We claim that h is pseudo-concave. It is clear that $h(u)$ is increasing. It suffices therefore to show that $\frac{h(u)}{u}$ is decreasing if u is large, since the values $u<1$ do not interfere. To this end we consider the corresponding logarithmic derivative; we have

$$\frac{d}{du} \log(h(u)/u) = (\lambda(\log u)^{\lambda-1} - 1)/u$$

which, since $\lambda<1$, eventually becomes < 0. Applying now our theorem we thus get the following interpolation theorem: *If π is a continuous linear mapping from L^p_0 into L^p_0 and from L^p_α into L^p_α it is also a continuous linear mapping from L^p_β into L^p_β* $(\alpha>\beta>0)$. Here we have set $L^p_\gamma = L^p_{(e^{x^\gamma})}$ with $\gamma=0, \alpha, \beta$.

References

[1] C. Goulaouic, Interpolation entre les espaces L^p avec poids, *C. R. Acad. Sci. Paris*, **262** (1966), 333—336.

[2] J. Peetre, On interpolation functions, *Acta Sci. Math.*, **27** (1966), 167—171.

(Received April 5, 1967)

A Cauchy problem for a generalized wave equation

By H. BERENS and U. WESTPHAL in Aachen (B. R. Deutschland)[1])

Introduction

The equation

$$\frac{\partial^2 w(x,t)}{\partial x^2} = \frac{\partial^2 w(x,t)}{\partial t^2} \qquad (x, t > 0)$$

with the initial conditions

$$w(x,0) = \left.\frac{\partial w(x,t)}{\partial t}\right|_{t=0} = 0 \qquad (x>0)$$

and the boundary conditions

$$w(0,t) = f(t), \quad \lim_{x\to\infty} w(x,t) = 0 \qquad (t>0)$$

is no doubt the simplest type of a boundary value problem which may be formulated for the wave equation. Its solution is given by $w(x, t) = f(t-x)$ for $0<x<t$ and $=0$ for $x>t$, in case $f(t)$ is twice continuously differentiable. The object here is to study a problem of this type for a generalized wave equation in the Lebesgue space $L^p(0, \infty)$ $(1 \leqq p < \infty)$. For the sake of precision we first restate the original problem as follows:

Let the operator J^{-2}, with domain $D(J^{-2}; p) = \{f \in L^p(0, \infty);\ f$ and f' locally absolutely continuous on $[0, \infty)$ with $f(0) = f'(0) = 0$ and $f'' \in L^p(0, \infty)\}$ and range in $L^p(0, \infty)$ $(1 \leqq p < \infty)$, be defined by $J^{-2}f = f''$.

Cauchy problem I. Given an element $f_0 \in L^p(0, \infty)$, find a vector-valued function $w(x) = w(x; f_0)$ on $[0, \infty)$ to $L^p(0, \infty)$ such that

(i) $w(x)$ is twice continuously differentiable in the L^p-norm on $(0, \infty)$;

(ii) $w(x) \in D(J^{-2}; p)$ for each $x>0$ and $\frac{d^2}{dx^2} w(x) = J^{-2} w(x)$;

(iii) there is a constant $M = M_{f_0} > 0$ such that $\|w(x)\|_p \leqq M$ $(x>0)$.

(iv) $\lim_{x\to 0+} \|w(x) - f_0\|_p = 0$.

---

[1]) The preparation of this paper was partially supported by a DFG grant.

One easily proves that for any given $f_0 \in D(J^{-2}; p)$ there is a unique solution $w(x; f_0) = W(x) f_0$, where $W(x)$ is the semi-group of right translations on $L^p(0, \infty)$:

$$[W(x)f](t) = \begin{cases} 0, & 0 < t \leqq x, \\ f(t-x), & x < t < \infty. \end{cases} \tag{1}$$

The Cauchy problem I will be generalized in such a way that the operator $J^{-2} = (\partial/\partial t)^2$ is replaced by a differential operator of order 2γ $(0 < \gamma \leqq 1)$, namely by $J^{-2\gamma} = (\partial/\partial t)^{2\gamma}$. (For the exact definition of $J^{-\gamma}$ and its domain in $L^p(0, \infty)$ see the following section.) In case $\gamma = 1/2$, this leads to a boundary value problem of the heat-conduction equation for a semi-infinite rod.

Abstract Cauchy problems of higher orders (especially of order two) are studied in E. Hille and R. S. Phillips [9, Sec. 23. 9]. Also we refer to Yu. I. Lyubich [11]. In particular for the wave equation we mention two papers [7, 8] of E. Hille. However, the cited papers always deal with initial value problems, while the above Cauchy problem I is a proper boundary value problem (cf. conditions (iii) and (iv)).

The Cauchy problem for the generalized wave equation will be treated in Sec. 2. To this end we first solve a corresponding Cauchy problem of order one via the Hille—Yosida theorem of semi-group theory. Thus we collect in Sec. 1 some results on semi-groups of operators on Banach spaces and then define integral and differential operators of fractional order $\gamma > 0$ acting on functions defined on the positive real axis. One-to-one characterizations of these notions are given via Laplace transforms, which also play an important role in the proofs. In Sec. 3 we finally prove several equivalent characterizations of the domain of the differential operator $J^{-\gamma}$ $(\gamma > 0)$ in the function space $L^p(0, \infty)$ $(1 \leqq p < \infty)$. The results obtained are generalizations of those presented by the authors in [4]. The most interesting characterization here will be in terms of the vector-valued integral

$$\int_0^\infty \frac{\Delta_u^n f}{u^{1+\gamma}}\, du \qquad (0 < \gamma < n), \tag{2}$$

where

$$[\Delta_u^n f](\cdot) = \sum_{j=0}^{n} (-1)^j \binom{n}{j} f(\cdot - ju) \tag{3}$$

is the n-th Riemann difference for the semi-group of right translations (1) on $L^p(0, \infty)$.

The results established in Sec. 3 possess far-reaching generalizations to fractional powers of infinitesimal generators of semi-groups of operators on Banach spaces. For further details, we refer to Sec. 5 in [4], to [3] and the literature cited there.

The authors are indebted to Professor P. L. Butzer for his kind interest in this paper.

1. Auxiliary results

Let X be a real or complex Banach space with elements $f, g, \ldots$ and norm $\|\cdot\|$, and let $\mathfrak{E}(X)$ be the Banach algebra of endomorphisms of X. If $T \in \mathfrak{E}(X)$, $\|T\|$ denotes the norm of T. A family of operators $\{T(x); x \geqq 0\}$ in $\mathfrak{E}(X)$ is said to be a contraction semi-group of class $(\mathfrak{C}_0)$, if it is subject to the following conditions: (i) $T(0)=I$ (identity operator); (ii) $T(x_1+x_2)=T(x_1)T(x_2)$ for $x_1, x_2 \in [0, \infty)$; (iii) $\|T(x)\| \leqq 1$ uniformly with respect to $x \geqq 0$; (iv) $\lim_{x\to 0+} \|T(x)f-f\|=0$ for all $f \in X$. The infinitesimal generator A of $\{T(x); x \geqq 0\}$, defined by $Af = \text{s-lim}_{x\to 0+} x^{-1} \cdot [T(x)-I]f$ whenever the strong limit (s-lim) exists, is a closed linear operator with domain $D(A)$ dense in X. The powers A^r of A $(r=2, 3, \ldots)$, are defined inductively. If f belongs to $D(A^r)$, so does $T(x)f$ for each $x \geqq 0$ and

$$\frac{d^r}{dx^r} T(x)f = A^r T(x)f = T(x)A^r f.$$

If $\{(Tx); x \geqq 0\}$ has a holomorphic extension $T(z)$ $(z=x+iy)$ in a sector $\{z; 0<x<\infty, |\arg z| \leqq \alpha_0 < \pi/2\}$ of the complex plane, we speak of a holomorphic semi-group. A necessary and sufficient condition for this is that $T(x)[X] \subset D(A)$ for $x>0$ and that there is a constant N such that $x\|AT(x)\| \leqq N$ for $x>0$.

Under the above hypotheses upon $\{T(x); x \geqq 0\}$ the set $\{\lambda; \lambda>0\}$ belongs to the resolvent set $\varrho(A)$ of the generator A, and the resolvent operator $R(\lambda; A)$ is given by

$$R(\lambda; A)f = \int_0^\infty e^{-\lambda x} T(x)f\,dx \qquad (f \in X; \lambda>0). \tag{4}$$

Also the inversion formula

$$T(x)f = \text{s-lim}_{\lambda\to\infty} e^{-\lambda x} \sum_{j=0}^{\infty} \frac{(\lambda x)^j}{j!} [\lambda R(\lambda; A)]^j f \qquad (f \in X) \tag{5}$$

holds uniformly with respect to x in any finite interval $[0, b]$. Finally, let us formulate the Hille—Yosida theorem: A closed linear operator U with dense domain and range in a Banach space X generates a contraction semi-group $\{T(x); x \geqq 0\}$ of class $(\mathfrak{C}_0)$ in $\mathfrak{E}(X)$ if and only if $\{\lambda; \lambda>0\} \subset \varrho(U)$ and $\lambda\|R(\lambda; U)\| \leqq 1$ for all $\lambda>0$. Moreover, U is the infinitesimal generator of exactly one semi-group given by (5). For a treatment of semi-group theory see E. Hille and R. S. Phillips [9, Part II], K. Yosida [13, Ch. IX] or P. L. Butzer and H. Berens [5, Ch. I].

We now introduce the concepts of integration and differentiation of fractional order. In the following, $f(t)$ will always be a real or complex-valued Lebesgue-measurable function defined on the positive real axis. The integral of f of order

$\gamma>0$ is then defined by the (Laplace) convolution integral

(6) $$[J^\gamma f](t) = \frac{1}{\Gamma(\gamma)} \int_0^t (t-u)^{\gamma-1} f(u)\, au \qquad (t>0).$$

$J^\gamma f$ exists for almost all t in $(0, \infty)$, whenever $f\in L(0, b)$ for every $b>0$, and also belongs to this space. On the other hand, for a $\gamma>0$ with $k-1<\gamma\leqq k$ (k integral) the derivative of f of order γ is defined by

(7) $$[J^{-\gamma} f](t) = \frac{d^k}{dt^k} [J^{k-\gamma} f](t) \qquad (t>0),$$

whenever this expression has a meaning, where $J^0=I$ (for the notations see the remarks in G. DOETSCH [6, vol. III, p. 164]; moreover, we refer to the literature cited in [4]). The following lemma characterizes the fractional integral and derivative of a function f by means of the Laplace transform.

Lemma 1. *Let $\gamma>0$ and f, g be two functions in $L(0, b)$ for every $b>0$.*

(a) *If the Laplace integral of f, i.e.*

$$f^\wedge(s) = \mathfrak{L}[f](s) \equiv \int_0^\infty e^{-st} f(t)\, dt,$$

converges absolutely for every complex number s with $\operatorname{Re} s>0$, *so does* $\mathfrak{L}[J^\gamma f](s)$ *and*

$$[J^\gamma f]^\wedge(s) = s^{-\gamma} f^\wedge(s) \qquad (\operatorname{Re} s>0),$$

where the branch of s^γ is taken such that $\operatorname{Re} s^\gamma>0$, *when* $\operatorname{Re} s>0$.

(b) *If $f^\wedge(s)$ and $g^\wedge(s)$ exist in the absolute sense for* $\operatorname{Re} s>0$ *and if*

$$s^\gamma f^\wedge(s) = g^\wedge(s) \qquad (\operatorname{Re} s>0),$$

then for $k-1<\gamma\leqq k$ (k integral) $J^{k-\gamma}f$ and its derivatives $[J^{k-\gamma-1}f], \dots, [J^{1-\gamma}f]$ up to the order $(k-1)$ are locally absolutely continuous on $[0, \infty)$ with $[J^{k-\gamma}f](0)=\cdots$ $\cdots=[J^{1-\gamma}f](0)=0$ and $[J^{-\gamma}f](t)=g(t)$ a.e.

For a proof see D. V. WIDDER [12, Ch. II, § 8].

Since we are mainly interested in functions f belonging to $L^p(0, \infty)$ $(1\leqq p<\infty)$ (endowed with the usual norm), we restrict the domain of $J^{-\gamma}$ $(\gamma>0)$ to the set

(8) $D(J^{-\gamma}; p)=\{f\in L^p(0, \infty);$ there is a $g\in L^p(0, \infty)$ with $f(t)=[J^\gamma g](t)$ a.e.$\}$.

Then $J^{-\gamma}f$ exists and is equal to g. By Lemma 1 we have equivalently

(9) $D(J^{-\gamma}; p)=\{f\in L^p(0, \infty);$ there is a $g\in L^p(0, \infty)$ with $s^\gamma f^\wedge(s)=g^\wedge(s)$ $(\operatorname{Re} s>0)\}$

and

(10) $$[J^{-\gamma}f]^\wedge(s) = s^\gamma f^\wedge(s) = g^\wedge(s) \qquad (\operatorname{Re} s>0).$$

2. The generalized wave equation

We now formulate and solve the Cauchy problem of first order for the operator $B_\gamma = -J^{-\gamma}$ $(0<\gamma<1)$ with domain $D(B_\gamma; p) = D(J^{-\gamma}; p)$ and range in $L^p(0, \infty)$ $(1 \leqq p < \infty)$.

Cauchy problem II. Given a function $f_0 \in L^p(0, \infty)$, find a function $w_\gamma(x) = w_\gamma(x; f_0)$ on $[0, \infty)$ to $L^p(0, \infty)$ such that

(i) $w_\gamma(x)$ is strongly continuously differentiable on $(0, \infty)$;

(ii) $w_\gamma(x) \in D(B_\gamma; p)$ and $(d/dx)w_\gamma(x) = B_\gamma w_\gamma(x)$ for each $x>0$;

(iii) $\lim_{x\to 0+} \|w_\gamma(x) - f_0\|_p = 0$.

Proposition 2. (a) *B_γ is a closed linear operator with domain dense in $L^p(0, \infty)$.*

(b) *The set $\{\lambda; \lambda>0\}$ belongs to the resolvent set $\varrho(B_\gamma)$ of B_γ, and the resolvent has the representation*

$$(11) \qquad [R(\lambda; B_\gamma)f](t) = \int_0^t f(t-u) r_\gamma(\lambda; u)\, du \qquad (f \in L^p(0, \infty)),$$

where

$$(12) \qquad r_\gamma(\lambda; t) = \frac{\sin \gamma\pi}{\pi} \int_0^\infty e^{-tu} \frac{u^\gamma}{\lambda^2 - 2\lambda u^\gamma \cos \gamma\pi + u^{2\gamma}}\, du.$$

Moreover,

$$(13) \qquad \|R(\lambda; B_\gamma)f\|_p \leqq \frac{\|f\|_p}{\lambda} \qquad (f \in L^p(0, \infty)).$$

Proof. (a) The linearity of B_γ is obvious by definition. To prove that B_γ is closed, suppose there is a sequence $\{f_n\}_{n=1}^\infty$ in $D(B_\gamma; p)$ such that f_n and $B_\gamma f_n$ converge in the L^p-norm to an f_0 and g_0, respectively. Since strong convergence implies weak convergence, we have for each fixed s $(\mathrm{Re}\, s>0)$

$$\lim_{n\to\infty} f_n^\wedge(s) = f_0^\wedge(s) \quad \text{and} \quad \lim_{n\to\infty} -s^\gamma f_n^\wedge(s) = -s^\gamma f_0^\wedge(s) = g_0^\wedge(s),$$

i.e. $f_0 \in D(B_\gamma; p)$ and $B_\gamma f_0 = g_0$. Finally, it is easy to see that $C_{00}^\infty(0, \infty)$, the space of arbitrarily often continuously differentiable functions with compact support in $(0, \infty)$, belongs to $D(B_\gamma; p)$. Since $C_{00}^\infty(0, \infty)$ is dense in $L^p(0, \infty)$, so is $D(B_\gamma; p)$.

(b) At first we prove that $\{\lambda; \lambda>0\} \subset \varrho(B_\gamma)$, i.e. we have to show that for each $\lambda>0$ the operator $\lambda I - B_\gamma$ from $D(B_\gamma; p)$ to $L^p(0, \infty)$ has an inverse $[\lambda I - B_\gamma]^{-1}$ such that its domain is equal to $L^p(0, \infty)$ (since B_γ is a closed operator). Indeed, the equation

$$(14) \qquad [\lambda I - B_\gamma] f = \theta \quad \text{or, equivalently,} \quad \lambda f^\wedge(s) + s^\gamma f^\wedge(s) = 0 \quad (\mathrm{Re}\, s>0)$$

implies that the function $f(t)$ is zero almost everywhere. Thus $[\lambda I - B_\gamma]^{-1}$ exists, and it remains to prove that for any given $g \in L^p(0, \infty)$ there is an $f \in D(B_\gamma; p)$ satisfying

(15) $\lambda f - B_\gamma f = g$ or, equivalently, $\lambda f^\wedge(s) + s^\gamma f^\wedge(s) = g^\wedge(s)$ $\quad$ (Re $s > 0$).

But the function $(\lambda + s^\gamma)^{-1}$ (Re $s > 0$) is the Laplace transform of the function $r_\gamma(\lambda; \cdot)$ defined in (12). $r_\gamma(\lambda; \cdot)$ is non-negative, belongs to $L^1(0, \infty)$ and $\int_0^\infty r_\gamma(\lambda; t)dt = \lambda^{-1}$ (see T. Kato, [10]). Hence the element

$$f(t) = [R_\lambda g](t) = \int_0^t g(t-u) r_\gamma(\lambda; u)\, du$$

belongs to $D(B_\gamma; p)$ and solves the differential equation (15). Thus $\{\lambda;\ \lambda > 0\} \subset \varrho(B_\gamma)$ and the resolvent $R(\lambda; B_\gamma)$ equals the operator R_λ. Finally,

$$\|R_\lambda g\|_p \leqq \|r_\gamma(\lambda;\ \cdot)\|_1 \|g\|_p = \lambda^{-1} \|g\|_p \qquad (\lambda > 0;\ \ g \in L^p(0, \infty))$$

proving the estimate (13).

Proposition 2 shows that the operator B_γ with domain and range in $L^p(0, \infty)$, $(1 \leqq p < \infty)$ satisfies the assumptions of the Hille—Yosida theorem. This leads to

Theorem 3. *The Cauchy problem* II *has a unique solution* $w_\gamma(x; f) = W_\gamma(x) f$ $(x \geqq 0)$ *for any given* $f \in L^p(0, \infty)$. $\{W_\gamma(x);\ x \geqq 0\}$ *is a holomorphic contraction semi-group of class* $(\mathfrak{C}_0)$ *in* $\mathfrak{E}(L^p(0, \infty))$ *generated by* B_γ *and is given by the convolution integral*

(16) $$[w_\gamma(x) f](t) = \int_0^t f(t-u) p_\gamma(x; u)\, du \qquad (f \in L^p(0, \infty))$$

with kernel

(17) $$p_\gamma(x; t) = \frac{1}{\pi} \int_0^\infty \exp(tu \cos\sigma - xu^\gamma \cos\gamma\sigma) \cdot \sin(tu \sin\sigma - xu^\gamma \sin\gamma\sigma + \sigma)\, du$$

$\left(x > 0,\ t \geqq 0;\ \frac{\pi}{2} \leqq \sigma \leqq \pi,\ 0 < \gamma < 1\right)$, *a Lévy stable density function on* $(0, \infty)$.

Proof. By the theorem of Hille—Yosida, the operator B_γ on $D(B_\gamma; p)$ to $L^p(0, \infty)$ generates a unique contraction semi-group $\{W_\gamma(x);\ x \geqq 0\}$ of class $(\mathfrak{C}_0)$ in $\mathfrak{E}(L^p(0, \infty))$. The operator $W_\gamma(x)$ $(x > 0)$ is given via the inversion formula (5) through

$$[W_\gamma(x) f]^\wedge(s) = \lim_{\lambda \to \infty} e^{-\lambda x} \sum_{j=0}^\infty \frac{(\lambda^2 x)^j}{j!} [\{R(\lambda; B_\gamma)\}^j f]^\wedge(s) =$$

$$= \lim_{\lambda \to \infty} e^{-\lambda x} \sum_{j=0}^\infty \frac{(\lambda^2 x)^j}{j!} \frac{f^\wedge(s)}{(\lambda + s^\gamma)^j} =$$

$$= \lim_{\lambda \to \infty} \exp\{-\lambda x + \lambda^2 x/(\lambda + s^\gamma)\} f^\wedge(s) = e^{-xs^\gamma} f^\wedge(s),$$

which holds for each fixed s, $\operatorname{Re} s>0$, and all $f\in L^p(0,\infty)$. Since $\exp(-xs^\gamma)$ is the Laplace transform of the density function $p_\gamma(x;\cdot)$ defined in (17) and since $p_\gamma(x;\cdot)$ is non-negative on $(0,\infty)$, belongs to $L^1(0,\infty)$ with $\int_0^\infty p_\gamma(x;u)du=1$ for each $x>0$ (see G. Doetsch [6, vol. I, p. 263] and K. Yosida [13, p. 259 ff]), the representation (16) of $W_\gamma(x)f$ follows. Moreover, the properties of $p_\gamma(x;\cdot)$ assure that $\{W_\gamma(x);x\geqq 0\}$ is holomorphic. Thus the function $W_\gamma(x)=w_\gamma(x)f$ on $[0,\infty)$ to $L^p(0,\infty)$ solves the Cauchy problem for each $f\in L^p(0,\infty)$. It remains to prove that the solution is unique. To this end suppose there is a non-trivial null solution $w_{\gamma,0}(x)=w_\gamma(x;\theta)$ on $[0,\infty)$, i.e. $w_\gamma(x;\theta)\neq\theta$ for all $x\geqq 0$. Then for each fixed s with $\operatorname{Re} s>0$,

$$\frac{d}{dx}[w_{\gamma,0}(x)]^\wedge(s)=[B_\gamma w_{\gamma,0}(x)]^\wedge(s)=-s^\gamma[w_{\gamma,0}(x)]^\wedge(s)$$

with

$$\lim_{x\to 0+}[w_{\gamma,0}(x)]^\wedge(s)=0.$$

The solution of this ordinary differential equation is given by $c(s)\exp(-xs^\gamma)$. But by the latter limit condition we obtain that $c(s)=0$ for each s, $\operatorname{Re} s>0$, and consequently $[w_{\gamma,0}(x)]^\wedge(s)=0$ ($\operatorname{Re} s>0$) or $w_{\gamma,0}(x)=\theta$ for all $x\geqq 0$. This is a contradiction, proving the theorem.

Here we remark that for $\gamma=1$ the solution of II is given by $W_1(x;f_0)=w(x)f_0$ for any $f_0\in D(J^{-1};p)$, where $\{W(x);x\geqq 0\}$ is the semi-group of right translations (1) on $L^p(0,\infty)$. However, for $\gamma>1$, Laplace transform methods may not be applied to solve the related Cauchy problem, since $\exp(-s^\gamma)$ ($\operatorname{Re} s>0$) is not the Laplace integral of a Lebesgue integrable function (see e.g. G. Doetsch [6, vol. I, p. 163]).

As the solution (16) of the Cauchy problem II is given by a holomorphic semi-group, it is evidently also a solution of the Cauchy problem I of second order for each function f in $L^p(0,\infty)$, where the operator J^{-2} in I is now replaced by $(-B_\gamma)^2=J^{-2\gamma}$ $(0<\gamma<1)$ with domain $D(J^{-2\gamma};p)$ in $L^p(0,\infty)$. Moreover, condition (iii) in I guarantees the uniqueness of the solution (16).

The Cauchy problem I taken in the generalized sense with J^{-2} replaced by $J^{-2\gamma}$ and $\gamma=1/2$ is known to be the formal version of the following boundary-value problem of the heat-conduction equation for a semi-infinite rod $(x\geqq 0)$:

$$\frac{\partial^2 w}{\partial x^2}=\frac{\partial w}{\partial t}\qquad (x,t>0)$$

with initial condition $w(x,0)=0$ $(x>0)$ and boundary conditions

$$w(0,t)=f(t)\quad\text{and}\quad\lim_{x\to\infty}w(x,t)=0\qquad(t>0).$$

Also, for $\gamma=1/2$ the density function is explicitly known, thus

$$p_{1/2}(x;t) = \frac{x}{\sqrt{4\pi}} \frac{\exp(-x^2/4t)}{t^{3/2}} \qquad (x, t > 0), \tag{18}$$

and the solution $w_{1/2}(x;f) = W_{1/2}(x)f$ takes on the form

$$W_{1/2}(x)f](t) = \frac{x}{\sqrt{4\pi}} \int_0^t f(t-u) \frac{\exp(-x^2/4u)}{u^{3/2}} du \qquad (f \in L^p(0,\infty)), \tag{19}$$

the semi-group property of $W_{1/2}(x)$ being reflected in the functional equation satisfied by the kernel (18):

$$\frac{x_1+x_2}{\sqrt{4\pi}} \frac{\exp(-(x_1+x_2)^2/4t)}{t^{3/2}} = \frac{x_1 x_2}{4\pi} \int_0^t \frac{\exp(-x_1^2/4(t-u)) \exp(-x_2^2/4u)}{(t-u)^{3/2} u^{3/2}} du$$

$$(x_1, x_2 > 0, \quad t > 0).$$

The latter relation was already noted in 1902 by E. Cesàro, as D. Doetsch [6, vol. III, p. 81, p. 267] remarks. Moreover, the resolvent operator of B is given by

$$[R(\lambda; B_{1/2})f](t) = \int_0^t f(t-u) \left\{ \frac{1}{\sqrt{\pi u}} - \lambda e^{\lambda^2 u} \operatorname{Erfc}(\lambda \sqrt{u}) \right\} du \tag{20}$$

$(\lambda > 0; f \in L^p(0, \infty))$, where $\operatorname{Erfc} u = (2/\sqrt{\pi}) \int_u^\infty \exp(-v^2) dv$ is the complementary error function.

3. Characterizations of the operator $J^{-\gamma}$, $\gamma > 0$

In the foregoing section we have seen that the characterizations (9) and (10) of the domain $D(J^{-\gamma}; p)$ of $J^{-\gamma}$ in $L^p(0, \infty)$ and of $J^{-\gamma}$ itself ($\gamma > 0$) through the Laplace transform are important auxiliary means for the solution of the Cauchy problem II. However, it is more satisfactory to obtain direct characterizations upon a function $f \in L^p(0, \infty)$ to belong to $D(J^{-\gamma}; p)$. This is the object of this section, generalizing at the same time results obtained in [4]. The following lemma gives an evaluation of the Laplace transform of the integral (2).

Lemma 4. *Let $0 < \gamma < n$ $(n = 1, 2, \cdots)$, and let $f \in L(0, b)$ for every $b > 0$, $\mathfrak{L}[f](s)$ being absolutely convergent for each s, $\operatorname{Re} s > 0$. The Laplace transform of (2) is then given by*

$$\mathfrak{L}\left[\int_\varepsilon^\infty \frac{\Delta_u^n f}{u^{1+\gamma}} du \right](s) = s^\gamma f^\wedge(s) q_{\gamma,n}^\wedge(s\varepsilon) \qquad (\varepsilon > 0; \operatorname{Re} s > 0), \tag{21}$$

where

(22) $$q^{\wedge}_{\gamma,n}(s) = \frac{1}{s^{\gamma}} \int_1^{\infty} \frac{(1-e^{-su})^n}{u^{1+\gamma}}\, du \qquad (\operatorname{Re} s > 0)$$

is the Laplace transform of the function

(23) $$q_{\gamma,n}(t) = \begin{cases} \dfrac{t^{-1}}{\Gamma(1+\gamma)} \displaystyle\sum_{j=0}^{l} (-1)^j \binom{n}{j} (t-j)^{\gamma} & (l < t \leqq l+1;\ l = 0, 1, \cdots, n-1), \\ \dfrac{t^{-1}}{\Gamma(1+\gamma)} \displaystyle\sum_{j=0}^{n} (-1)^j \binom{n}{j} (t-j)^{\gamma} & (t > n) \end{cases}$$

belonging to $L^1(0, \infty)$. *Moreover*

(24) $$C_{\gamma,n} \equiv \lim_{\varepsilon \to 0+} q^{\wedge}_{\gamma,n}(s\varepsilon) = \int_0^{\infty} q_{\gamma,n}(u)\, du = \int_0^{\infty} \frac{(1-e^{-u})^n}{u^{1+\gamma}}\, du$$

and

(25) $$C_{\gamma,n} = \begin{cases} \Gamma(-\gamma) \displaystyle\sum_{j=1}^{n} (-1)^j \binom{n}{j} j^{\gamma} & (0 < \gamma < n,\ \gamma \text{ non-integral}) \\ \dfrac{(-1)^{\gamma+1}}{\gamma!} \displaystyle\sum_{j=1}^{n} (-1)^j \binom{n}{j} j^{\gamma} \log j & (\gamma = 1, 2, \cdots, n-1). \end{cases}$$

Proof. By FUBINI'S theorem we obtain for each fixed s, $\operatorname{Re} s > 0$,

$$\int_0^{\infty} e^{-st}\, dt \int_{\varepsilon}^{\infty} u^{-1-\gamma} [\Delta_u^n f](t)\, du = \int_{\varepsilon}^{\infty} u^{-1-\gamma}\, du \sum_{j=0}^{n} (-1)^j \binom{n}{j} \int_{ju}^{\infty} e^{-st} f(t-ju)\, dt =$$

$$= f^{\wedge}(s) \sum_{j=0}^{n} (-1)^j \binom{n}{j} \int_{\varepsilon}^{\infty} e^{-sju} u^{-1-\gamma}\, du = s^{\gamma} f^{\wedge}(s) \Big\{ \gamma^{-1} (s\varepsilon)^{-\gamma} +$$

$$+ \sum_{j=1}^{n} (-1)^j \binom{n}{j} j^{\gamma} (s\varepsilon)^{-\gamma} \int_j^{\infty} e^{-s\varepsilon u} u^{-1-\gamma}\, du = s^{\gamma} f^{\wedge}(s) q^{\wedge}_{\gamma,n}(s\varepsilon),$$

giving at the same time the representation (22). By the fact that $s^{-\gamma} = \mathfrak{L}[u^{\gamma-1}/\Gamma(\gamma)](s)$ ($\gamma > 0$ and $\operatorname{Re} s > 0$) as well as by the convolution theorem, this leads to

$$q_{\gamma,n}(u) = \frac{u^{\gamma-1}}{\Gamma(\gamma+1)} + \sum_{j=1}^{n} (-1)^j \binom{n}{j} j^{\gamma} h_j(u),$$

where

(26) $$h_j(u) = \begin{cases} 0 & (u < j), \\ \displaystyle\int_j^{u} \frac{(u-v)^{\gamma-1}}{\Gamma(\gamma)} v^{-1-\gamma}\, dv & (u > j). \end{cases}$$

Applying the substitution $(1/v)-(1/u)=t$ in the integral (26), we obtain the representation (23) for $q_{\gamma,n}$.

We now prove that $q_{\gamma,n}$ belongs to $L^1(0,\infty)$. Obviously, $q_{\gamma,n}$ is a continuous function on $(0,\infty)$ and it belongs to $L(0,b)$ for every $b>0$. Moreover, for $\gamma=1,2,\cdots,\cdots,n-1$, $q_{\gamma,n}(t)=0$ for $t>n$. This follows by the fact that the function

$$e^{-\varepsilon t}(1-e^{\varepsilon})^n=\sum_{j=0}^{n}(-1)^j\binom{n}{j}e^{-\varepsilon(t-j)}$$

and its $(n-1)$ derivatives with respect to ε vanish at $\varepsilon=0$. So we may restrict the discussion to non-integral γ, $0<\gamma<n$. By a lengthy calculation one obtains for $t>n$ the representation

$$q_{\gamma,n}(t)=-\frac{\sin\gamma\pi}{t\pi}\int_0^\infty e^{-tv}(1-e^v)^n v^{-\gamma-1}\,dv \qquad (t>n).$$

Hence, for $t>n$ $q_{\gamma,n}(t)$ has a uniquely determined sign: $\operatorname{sgn} q_{\gamma,n}(t)=(-1)^{n-k}$ $(k-1<\gamma<k,\ k=1,2,\cdots,n;\ t>n)$. So it suffices to prove that the limit $\int_a^b q_{\gamma,n}(u)du$ $(b>a>n)$ exists as $b\to\infty$. For $k-1<\gamma<k$, partial integration (k-times) gives

$$\int_a^b u^{\gamma-1}\sum_{j=0}^{n}(-1)^j\binom{n}{j}\left(1-\frac{j}{u}\right)^{\gamma}du=$$

$$=\sum_{l=0}^{k-1}(-1)^l\frac{u^{\gamma-l}}{\gamma-l}\sum_{j=0}^{n}(-1)^j\binom{n}{j}j^l\left(1-\frac{j}{u}\right)^{\gamma-l}\Bigg|_{u=a}^{b}+$$

$$+\sum_{j=0}^{n}(-1)^j\binom{n}{j}(-1)^k j^k\int_a^b u^{\gamma-k-1}\left(1-\frac{j}{u}\right)^{\gamma-k}du.$$

The first sum on the right-hand side of this equation for $u=b$ tends to zero as $b\to\infty$, while the absolute value of the second sum is majorized by

$$\sum_{j=0}^{n}\binom{n}{j}j^k(a-j)^{\gamma-k}/(k-\gamma).$$

Clearly, since $q_{\gamma,n}\in L^1(0,\infty)$ for each fixed s $(\operatorname{Re} s>0)$

$$C_{\gamma,n}\equiv\lim_{\varepsilon\to0+}q^{\wedge}_{\gamma,n}(s\varepsilon)=\int_0^\infty q_{\gamma,n}(u)\,du,$$

which by the representation (22) with $s=1$ leads to the equation (24). From (24) one may determine the explicit form (25) of $C_{\gamma,n}$.

As a consequence of Lemma 4 we have

Theorem 5. *Let f and g belong to $L^p(0,\infty)$ $(1\leqq p<\infty)$. Then the following are equivalent for $0<\gamma<n$:*

(i) $f\in D(J^{-\gamma};p)$ *with* $J^{-\gamma}f=g$; *i.e.* $s^\gamma f^\wedge(s)=g^\wedge(s)$ $(\operatorname{Re} s>0)$;

(ii) $$\lim_{\varepsilon\to 0+}\left\|\frac{1}{C_{\gamma,n}}\int_\varepsilon^\infty \frac{\Delta_u^n f}{u^{1+\gamma}}\,du-g\right\|_p=0.$$

Proof. If (i) holds, then by Lemma 4 (21)

(27) $$\frac{1}{C_{\gamma,n}}\int_\varepsilon^\infty \frac{[\Delta_u^n f](t)}{u^{1+\gamma}}\,du=\frac{1}{C_{\gamma,n}}\int_0^t g(t-u)\,q_{\gamma,n}\left(\frac{u}{\varepsilon}\right)\frac{du}{\varepsilon}\qquad(\varepsilon>0)$$

for almost all $t>0$. Since $q_{\gamma,n}\in L^1(0,\infty)$ with

$$\int_0^\infty q_{\gamma,n}(u)\,du=C_{\gamma,n},$$

the right-hand side of (27) converges to g in the L^p-norm as $\varepsilon\to 0+$, giving (ii). On the other hand, for each fixed s, $\operatorname{Re} s>0$, by (21)

$$\left|\frac{1}{C_{\gamma,n}}s^\gamma f^\wedge(s)\,q_{\gamma,n}^\wedge(s\varepsilon)-g^\wedge(s)\right|\leqq\frac{1}{(\sigma p')^{1/p'}}\left\|\frac{1}{C_{\gamma,n}}\int_\varepsilon^\infty\frac{\Delta_u^n f}{u^{1+\gamma}}\,du-g\right\|_p\qquad(\varepsilon>0),$$

where $\sigma=\operatorname{Re} s$ and $p^{-1}+p'^{-1}=1$. Letting $\varepsilon\to 0+$, (i) follows by (24).

We remark that Theorem 5 generalizes Theorem 2 of [4] to arbitrary $\gamma>0$.

Proposition 6. *If a function $f\in L^p(0,\infty)$ belongs to $D(J^{-\gamma_0};p)$ $(\gamma_0>0)$ then f belongs to $D(J^{-\gamma};p)$ for each $0<\gamma<\gamma_0$.*

Proof. Let n be an integer such that $n>\gamma_0$. By Theorem 5 we have to prove that the limit in the L^p-norm of

$$\int_\varepsilon^\infty u^{-\gamma-1}\Delta_u^n f\,du$$

exists as $\varepsilon\to 0+$. Using

$$u^{\gamma_0-\gamma}=\varepsilon^{\gamma_0-\gamma}+(\gamma_0-\gamma)\int_\varepsilon^u v^{\gamma_0-\gamma-1}\,dv,$$

one obtains by a change of integration

$$\int_{\varepsilon}^{\infty} \frac{\Delta_u^n f}{u^{1+\gamma}}\, du = \int_{\varepsilon}^{\infty} \frac{\Delta_u^n f}{u^{1+\gamma_0}} u^{\gamma_0-\gamma}\, du =$$

$$= (\gamma_0-\gamma)\int_{\varepsilon}^{\infty} v^{\gamma_0-\gamma-1}\, dv \int_{v}^{\infty} \frac{\Delta_u^n f}{u^{1+\gamma_0}}\, du + \varepsilon^{\gamma_0-\gamma}\int_{\varepsilon}^{\infty} \frac{\Delta_u^n f}{u^{1+\gamma_0}}\, du \qquad (\varepsilon>0).$$

But

$$\left\| \int_{v}^{\infty} u^{-\gamma_0-1} \Delta_u^n f\, du \right\|_p \leqq 2^n \gamma_0^{-1} v^{-\gamma_0} \|f\|_p$$

for all $f \in L^p(0,\infty)$ and $\leqq \|q_{\gamma,n}\|_1 \|J^{-\gamma_0} f\|_p$ for all $f \in D(J^{-\gamma_0}; p)$ by (27). The desired result now follows immediately. As a consequence of Theorem 5 and Proposition 6 we have

Theorem 7. *Let* $0<\gamma<n$ $(n=1,2,\cdots)$, $\gamma=k+\sigma$ $(k=0,1,\cdots,n-1$ *and* $0<\sigma\leqq 1)$. *An element* $f \in L^p(0,\infty)$ *belongs to* $D(J^{-\gamma}; p)$ *with* $J^{-\gamma} f = g$ *if and only if* $f \in D(J^{-k}_{-}; p)$ *and*

$$\lim_{\varepsilon\to 0+} \left\| \frac{1}{C_{\sigma,1}} \int_{\varepsilon}^{\infty} \frac{\Delta_u f^{(k)}}{u^{1+\sigma}}\, du - g \right\|_p = 0 \ (0<\sigma<1),$$

$$\lim_{\varepsilon\to 0+} \left\| \frac{1}{C_{1,2}} \int_{\varepsilon}^{\infty} \frac{\Delta_u^2 f^{(k)}}{u^2}\, du - g \right\|_p = 0 \quad (\sigma=1).$$

There are further characterizations of $J^{-\gamma}$ $(\gamma>0)$ and its domain in $L^p(0,\infty)$ by the semi-groups of operators defined in (1) and (16), respectively.

Theorem 8. *Let* $0<\gamma\leqq n$ $(n=1,2,\cdots)$. *An element* $f \in L^p(0,\infty)$ *belongs to* $D(J^{-\gamma}; p)$ *with* $J^{-\gamma} f = g$ *if and only if*

$$\lim_{x\to 0+} \left\| \frac{[I-W_{\gamma/n}(x)]^n f}{x^n} - g \right\|_p = 0 \quad (0<\gamma<n), \qquad \lim_{x\to 0+} \left\| \frac{\Delta_x^n f}{x^n} - g \right\|_p = 0 \quad (\gamma=n).$$

The theorem can be proved directly via Laplace transform methods, thus by the results of Sec. 2. On the other hand, it is also a consequence of a general theorem on powers of generators of semi-groups of operators on Banach spaces given in [1] (see also [5, Sec. 2. 2]). Let us finally state one further more general result.

Theorem 9. *Let* $0<\gamma<n\,(n=1, 2, \cdots)$. *For an element* $f\in L^p(0, \infty)$ $(1\leqq p<\infty)$ *the following assertions are equivalent:*

(i) *for* $p=1$: *there is a function* μ *of bounded variation on* $[0, \infty)$ *such that*

$$s^\gamma f^\wedge(s) = \mu^\vee(s) \equiv \int_0^\infty e^{-st}d\mu(t) \qquad (\operatorname{Re} s>0),$$

for $1<p<\infty$: *there is a function* $g\in L^p(0, \infty)$ *with* $s^\gamma f^\wedge(s)=g^\wedge(s)$ $(\operatorname{Re} s>0)$;

(ii) $$\left\| \frac{1}{C_{\gamma, n}} \int_\varepsilon^\infty \frac{\Delta_u^n f}{u^{1+\gamma}}\, du \right\|_p = O(1) \qquad (\varepsilon \to 0+);$$

(iii) $$\|[I-W_{\gamma/n}(x)]^n f\|_p = O(1) \qquad (x \to 0+).$$

This theorem solves the saturation problem connected with the operators $J^{-\gamma}$ $(\gamma>0)$ in the function space $L^p(0, \infty)$ $(1\leqq p<\infty)$ posed in [2] for $0<\gamma<1$. For a proof of Theorem 9 in case $n=1$ see Theorems 5 and 6 in [4]. Using Lemma 4 these methods may then be easily generalized to arbitrary $n=1, 2, \cdots$ (cf. Theorem 5 above).

References

[1] H. Berens, Equivalent representations for the infinitesimal generator of higher orders in semi-group theory, *Nederl. Akad. Wetensch. Indag. Math.*, **27** (1965), 497—512.

[2] H. Berens and P. L. Butzer, *On the best approximation for singular integrals by Laplace transform methods* (Proceedings of a Conference in Oberwolfach, 1963). In: *On Approximation Theory*, edited by P. L. Butzer and J. Korevaar, ISNM 5, 12—24 (Basel, 1964).

[3] H. Berens, P. L. Butzer and U. Westphal, Representation of fractional powers of infinitesimal generators of semigroups, *Bull. Amer. Math. Soc.*, **74** (1968), 191—196.

[4] H. Berens und U. Westphal, Zur Charakterisierung von Ableitungen nichtganzer Ordnung im Rahmen der Laplacetransformation, *Math. Nachr.* (in print).

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

[6] G. Doetsch, *Handbuch der Laplace-Transformation*, Bd. I, III (Basel, 1950, 1956).

[7] E. Hille, Une généralisation du problème de Cauchy, *Annales Institut Fourier*, **4** (1952), 31—48.

[8] E. Hille, An abstract formulation of Cauchy's problem, *Comptes rendus du Douzième Congrès des Mathématiciens Scandinaves, Lund*, 10—15 *août* 1953.

[9] E. Hille and R. S. Phillips, *Functional Analysis and Semi-Groups* (Providence, 1957).

[10] T. Kato, Note on fractional powers of linear operators. *Proc. Japan Acad.*, **36** (1960), 94—96.

[11] Yu. I. Lyubich, The classical and local Laplace transformation in an abstract Cauchy problem, *Russian Math. Surveys*, **21** (1966), 1—52.

[12] D. V. Widder, *The Laplace Transform* (Princeton, 1946).

[13] K. Yosida, *Functional Analysis* (Berlin—Göttingen—Heidelberg, 1965).

AACHEN, GERMANY, THE TECHNICAL UNIVERSITY

(Received September 20, 1967)

On the characterization of classes of functions by their best linear approximation

By G. ALEXITS in Budapest

1. In the theory of approximation it is an attractive problem to characterize different classes of functions by means of their best approximation. Denote by $\{E_n\}$ a positive non-increasing number sequence. We say that the class of functions $\mathscr{C}$ is characterized by $\{E_n\}$, if the nth best approximation of every $f \in \mathscr{C}$ by polynomials, or another by appropriate system of functions, remains $\leqq E_n$ for every n (so-called direct theorem) and there exists a constant $K > 0$ such that every $f(x)$ for which the nth best approximation is $\leqq KE_n$ for every n belongs to $\mathscr{C}$ (so-called inverse theorem).

It is well known that many important classes of functions are characterized by the sequence of their best trigonometric polynomial approximation. Such are e.g. the classes $\mathrm{Lip}_M\, \alpha$ for $0 < \alpha < 1$ and a fixed $M > 0$ (Lipschitz-constant), the characteristic sequence $\{E_n\}$ being $\{CM \cdot n^{-\alpha}\}$ where C is an absolute constant. But, for $\alpha = 1$, the class $\mathrm{Lip}_M\, 1$ is no more characterized by $\{CM \cdot n^{-1}\}$, this sequence being characteristic for the Zygmund class Z_M containing all 2π-periodic functions for which

$$\max_{0 \leqq x \leqq 2\pi} |f(x+h) + f(x-h) - 2f(x)| \leqq M|h|.$$

In the following, we intend to investigate the nature of the classes $\mathscr{C}$ characterizable by the sequence of their best approximation with linear combinations of an arbitrary system of functions $\{f_\nu(x)\}$. Our main result can be expressed, roughly speaking, about so: if a class $\mathscr{C}$ is characterizable by a sufficiently regular sequence $\{E_n\}$ of any best approximation, then E_n grants the order of magnitude of the absolutely best possible nth linear approximation of $\mathscr{C}$.

One may be tempted to think that the order of the absolutely best approximation could be essentially improved if we turned to non-linear methods of approximation. But, it is not easy to find out what kind of a non-linear method could accomplish this task. If we confront, for instance, the linear method of approximation by polynomials with the non-linear method of approximation by rational functions, we shall show that, for the classes characterizable by polynomial approximation, the best approximation by rational functions gives no essentially better result. Hence, only the polynomially non-characterizable classes are worth to be investigated concerning their best approximation by rational functions.

. Approximations in a Banach space

2. Let B be a Banach space, $\|x\|_B$ the norm of $x \in B$ and $\{y_\nu\}$ a sequence of elements of B set in a fixed order. We form linear combinations $\sum_{k=1}^{n} a_k y_k$ of the first n elements where the a_k's are real numbers. The non-negative number

$$E_n^{(B)}(x, \{y_\nu\}) = \inf_{a_k} \|x - \sum_{k=1}^{n} a_k y_k\|_B$$

is the *nth best* $\{y_\nu\}$*-approximation* of $x \in B$. We call $\{y_\nu\}$ the *basis* of this approximation. If $\mathscr{C}$ is any subset of B, then

$$E_n^{(B)}(\mathscr{C}, \{y_\nu\}) = \sup_{x \in \mathscr{C}} E_n^{(B)}(x, \{y_\nu\})$$

is the nth best $\{y_\nu\}$-approximation *of the set* $\mathscr{C}$.

Denote by $\mathscr{C}(\{E_n\}, \{y_\nu\})$ the set of all those elements $x \in B$ for which

$$E_n^{(B)}(x, \{y_\nu\}) \leqq E_n \qquad (n = 1, 2, \cdots).$$

We call $\mathscr{C}(\{E_n\}, \{y_\nu\})$ the $\{E_n\}$*-saturation set* of the $\{y_\nu\}$-approximation. This is the set of all elements x of B for which a "direct theorem" with the best approximation sequence $\{E_n\}$ exists (referring to $\{y_\nu\}$-approximation). Saturation sets are closely connected with characterizable classes: a set $\mathscr{C} \subset B$ is called $\{E_n\}$*-characterizable,* if there exists a sequence $\{y_\nu\}$ and a positive absolute constant K_1 such that

$$\mathscr{C}(\{K_1 E_n\}, \{y_\nu\}) \subset \mathscr{C} \subset \mathscr{C}(\{E_n\}, \{y_\nu\}).$$

($K_1, K_2, \cdots$ will denote always positive absolute constants.) In the following, we shall suppose that $\{E_n\}$ is a positive non-decreasing number sequence tending to zero.

3. Theorem 1. *A saturation set* $\mathscr{C}(\{E_n\}, \{y_\nu\})$ *is a closed and convex subset of* B, *provided that the elements* $\{y_\nu\}$ *are independent.*

First of all, $\mathscr{C}(\{E_n\}, \{y_\nu\})$ contains infinitely many elements because of a theorem of Bernstein (cf. [3], p. 332) according to which, for every $p \geqq 1$, there exists an element x_p such that

$$E_n^{(B)}(x_p, \{y_\nu\}) = \frac{E_n}{p} \qquad (n=1, 2, \cdots),$$

hence every x_p belongs to $\mathscr{C}(\{E_n\}, \{y_\nu\})$. After this, let x, y be two different elements of $\mathscr{C}(\{E_n\}, \{y_\nu\})$ and consider the element $z=(1-\lambda)x+\lambda y$ where $0<\lambda<1$. Since, for every n,

$$E_n^{(B)}(z, \{y_\nu\}) \leqq (1-\lambda)E_n^{(B)}(x, \{y_\nu\})+\lambda E_n^{(B)}(y, \{y_\nu\}) \leqq ((1-\lambda)+\lambda)\cdot E_n = E_n,$$

the element z belongs to $\mathscr{C}(\{E_n\}, \{y_\nu\})$, this set is therefore convex. — As for the closure, let x be an element of accumulation and $\{x_j\}$ a sequence of different elements of $\mathscr{C}(\{E_n\}, \{y_\nu\})$ converging to x. By assumption, to every x_j belongs at least one linear combination $\sum_{k=1}^{n} a_k^{(j)} y_k$ such that

$$\| x_j - \sum_{k=1}^{n} a_k^{(j)} y_k \|_B = E_n^{(B)}(x_j, \{y_\nu\}) \leqq E_n \qquad (n=1, 2, \cdots).$$

Then we have

$$E_n^{(B)}(x, \{y_\nu\}) \leqq \| x - \sum_{k=1}^{n} a_k^{(j)} y_k \|_B \leqq \|x-x_j\|_B + E_n,$$

hence, going over to the limit $j\to\infty$,

$$E_n^{(B)}(x, \{y_\nu\}) \leqq E_n \qquad (n=1, 2, \cdots),$$

i.e. $x\in\mathscr{C}$, what we had to prove.

Corollary 1. *A set $\mathscr{C}$ is characterizable only if it contains a convex continuum, namely $\mathscr{C}(\{K_1 E_n\}, \{y_\nu\})$.*

4. The class $\mathscr{C}$ of convex functions having bounded nth derivatives was, in the last time, subject of investigations ([8], I and II) concerning its best approximation by rational functions. This class $\mathscr{C}$ can serve as an example for a class which is absolutely not characterizable by linear approximations in any Banach space. Because if, on the contrary, there were a characteristic sequence $\{E_n\}$ to a linear $\{y_\nu\}$-approximation, then $\mathscr{C}$ would contain $\mathscr{C}(\{K_1 E_n\}, \{y_\nu\})$. So choose two convex functions $f(x)$ and $g(x)$ contained in $\mathscr{C}$ and a number $0<\lambda<1$ such that $h(x)=(1-\lambda)f(x)+\lambda g(x)$ should not be convex. Then $K_1 f(x)$ and $K_1 g(x)$ would be contained in $\mathscr{C}(\{K_1 E_n\}, \{y_\nu\})$ but $K_1 h(x)$ not, contrary to Corollary 1.

5. We call the number sequence $\{E_n\}$ *slowly decreasing*, if it is positive, non-increasing and tends to zero such that $E_{2n}\geqq K_2 E_n$ for every n.

Theorem 2. *If $\{E_n\}$ is slowly decreasing and the basis of approximation bounded* ($\|y_\nu\|_B \leqq K_3$), *then there exists a* $K_4 > 0$ *such that*

$$K_4 E_n \cdot y_k \in \mathscr{C}(\{E_n\}, \{y_\nu\}) \qquad (k = 1, 2, \cdots, 2n;\ n = 1, 2, \cdots).$$

The boundedness of the basis $\{y_\nu\}$ implies

$$\left\| \frac{E_{2n}}{2K_3} \cdot y_k - \frac{E_{2n}}{2K_3} \cdot y_m \right\|_B \leqq \frac{E_{2n}}{2K_3} (\|y_k\|_B + \|y_m\|_B) \leqq E_{2n}.$$

Therefore, if $k \leqq 2n$ and $m \leqq k$,

$$E_m^{(B)} \left(\frac{E_{2n}}{2K_3} \cdot y_k, \{y_\nu\} \right) \leqq E_{2n} \leqq E_m.$$

This estimation is the more satisfied when $m > k$, because then we have

$$E_m^{(B)} \left(\frac{E_{2n}}{2K_3} \cdot y_k, \{y_\nu\} \right) = 0.$$

So it follows

$$\frac{E_{2n}}{2K_3} \cdot y_k \in \mathscr{C}(\{E_n\}, \{y_\nu\}) \qquad (k = 1, 2 \cdots, 2n;\ n = 1, 2, \cdots).$$

Put $K_4 = K_2/2K_3$, then, by the slow decrease of $\{E_n\}$, we obtain

$$K_4 E_n \leqq \frac{E_{2n}}{2K_3},$$

hence the more we have $K_4 E_n \cdot y_k \in \mathscr{C}(\{E_n\}, \{y_\nu\})$ for $k = 1, 2, \cdots, 2n$ and $n = 1, 2, \cdots$.

Corollary 2. *If $\{E_n\}$ is slowly decreasing and $\{y_\nu\}$ bounded, then a set $\mathscr{C}$ can be $\{E_n\}$-characterisable by $\{y_\nu\}$-approximation only if there exists a K_4 such that $K_4 E_n \cdot y_k \in \mathscr{C}$ for $k = 1, 2, \cdots, 2n$ and $n = 1, 2, \cdots$.*

6. It is known [4] that the best linear approximation in the space C of continuous functions with the basis $\{w_\nu(x)\}$ of Walsh functions provides, for the classes $\mathrm{Lip}\,\alpha$ with $0 < \alpha < 1$ and Lipschitz constant 1, about the same order of magnitude as the best polynomial approximation, namely $\{n^{-\alpha}\}$. Although, for the $\{w_\nu(x)\}$-approximation of $\mathrm{Lip}\,\alpha$ only direct theorems can be obtained. Because if $\mathscr{C} = \mathrm{Lip}\,\alpha$ were $\{E_n\}$-characterizable by $\{w_\nu(x)\}$-approximation, then by Corollary 2 the functions $K_4 n^{-\alpha} w_k(x)$ $(k = 2n)$ would belong to $\mathscr{C}$ and this is impossible, since $w_k(x)$ is not continuous.

Approximation in Banach spaces contained in a Hilbert space

7. Let H be a Hilbert space, (x, y) the inner product of the elements x, y of H and $\|x\|_H = \sqrt{(x, x)}$ the norm of $x \in H$. By B we denote now a Banach space $B \subset H$ for which the relation $\|x\|_B \geqq K_5 \|x\|_H$ is valid for all $x \in B$.

We prove now the counterpart of Theorem 2, a statement proved previously, in a somewhat less general form, by KNAPOWSKI and myself [2].

Theorem 3. *Assume that for a set $\mathscr{C} \subset B$ and for an orthonormal system $\{\xi_\nu\} \subset B$ we have*

(1) $$E_n \cdot \xi_k \in \mathscr{C} \qquad (k = 1, 2, \cdots, 2n;\ n = 1, 2, \cdots),$$

then, for any system $\{y_\nu\} \subset B$, $E_n(\mathscr{C}, \{y_\nu\}) \geqq K_6 E_n \qquad (n = 1, 2, \cdots)$.

First, let $\{\eta_\nu\} \subset B$ be an arbitrary orthonormal system and

$$s_n(\xi_k, \{\eta_\nu\}) = \sum_{\nu=1}^{n} (\xi_k, \eta_\nu) \cdot \eta_\nu .$$

By the orthonormality, we have

$$1 = \|\xi_k\|_H \leqq \|\xi_k - s_n(\xi_k, \{\eta_\nu\})\|_H + \|s_n(\xi_k, \{\eta_\nu\})\|_H .$$

Hence, owing attention to $E_n^{(H)}(\xi_k, \{\eta_\nu\}) = \|\xi_k - s_n(\xi_k, \{\eta_\nu\})\|_H$, and summing for $k = 1, 2, \cdots, 2n$, we get

(2) $$2n \leqq \sum_{k=1}^{2n} E_n^{(H)}(\xi_k, \{\eta_\nu\}) + \sum_{k=1}^{2n} \|s_n(\xi_k, \{\eta_\nu\})\|_H .$$

Applying CAUCHY's inequality, we see that

$$\sum_{k=1}^{2n} \|s_n(\xi_k, \{\eta_\nu\})\|_H \leqq \sqrt{2n} \left\{ \sum_{k=1}^{2n} \|s_n(\xi_k, \{\eta_\nu\})\|_H^2 \right\}^{\frac{1}{2}} = \sqrt{2n} \left\{ \sum_{k=1}^{2n} \sum_{j=1}^{n} (\xi_k, \eta_j)^2 \right\}^{\frac{1}{2}} .$$

Both systems $\{\xi_\nu\}$ and $\{\eta_\nu\}$ being orthonormed, by BESSEL's inequality,

$$\sum_{k=1}^{2n} (\xi_k, \eta_j)^2 \leqq \|\eta_j\|_H^2 = 1,$$

hence we obtain the estimation $\sum_{k=1}^{2n} \|s_n(\xi_k, \{\eta_\nu\})\|_H \leqq \sqrt{2}\, n$.

Therefore (2) leads to $\sum_{k=1}^{2n} E_n^{(H)}(\xi_k, \{\eta_\nu\}) \geqq (2 - \sqrt{2})n$, and consequently to

(3) $$\max_{1 \leqq k \leqq 2n} E_n^{(H)}(\xi_k, \{\eta_\nu\}) \geqq \frac{2 - \sqrt{2}}{2} .$$

By assumption, $\|x\|_B \geqq K_5 \|x\|_H$ for all $x \in B$, hence also $E_n^{(B)}(x, \{\eta_\nu\}) \geqq K_5 E_n^{(H)}(x, \{\eta_\nu\})$. Thus, if we apply (3) to the elements $E_n \cdot \xi_k$ instead of ξ_k, we obtain

$$\max_{1 \leqq k \leqq 2n} E_n^{(B)}(E_n \cdot \xi_k, \{\eta_\nu\}) \geqq K_5 \frac{2-\sqrt{2}}{2} E_n.$$

Since we assumed that, for $k \leqq 2n$, $E_n \cdot \xi_k$ is an element of $\mathscr{C}$, hence

$$E_n^{(B)}(\mathscr{C}, \{\eta_\nu\}) \geqq K_6 E_n. \tag{4}$$

The estimation (4) shows that our theorem holds good for the best approximations with an orthonormal basis $\{\eta_\nu\} \subset B$. Consider, now, an arbitrary approximation basis $\{y_\nu\} \subset B$. Delete from $\{y_\nu\}$ the linearly dependent elements and denote by $\{y_\nu^*\}$ the remainder system. Since there are less linear combinations of $y_1, y_2, \cdots, y_\nu$ than those of $y_1^*, y_2^*, \cdots, y_n^*$, we have $E_n^{(B)}(x, \{y_\nu\}) \geqq E_n^{(B)}(x, \{y_\nu^*\})$. But $\{y_\nu^*\}$ can be orthonormalized such that the νth element η_ν of the orthonormalized system should be a linear combination of the elements $y_1^*, y_2^*, \cdots, y_\nu^*$. Thus the linear combinations of $y_1^*, y_2^*, \cdots, y_n^*$ are the same as those of $\eta_1, \eta_2, \cdots, \eta_n$, therefore

$$E_n^{(B)}(x, \{y_\nu\}) \geqq E_n^{(B)}(x, \{y_\nu^*\}) = E_n^{(B)}(x, \{\eta_\nu\})$$

for all $x \in B$ and $n = 1, 2, \cdots$. Hence, by (4), $E_n^{(B)}(\mathscr{C}, \{y_\nu\}) \geqq K_6 E_n$, as we asserted.

8. For two positive number sequences $\{a_n\}$ and $\{b_n\}$, we write $\{a_n\} \approx \{b_n\}$, if $a_n \leqq K_7 b_n$ and $b_n \leqq K_8 a_n$ for $n = 1, 2, \cdots$, i.e. if $\{a_n\}$ and $\{b_n\}$ have the same order of magnitude. If there exists an orthonormal system satisfying (1) for a set $\mathscr{C} \subset B$ then, by Theorem 3, it follows that the numbers

$$E_n^{(B)}(\mathscr{C}) = \inf E_n^{(B)}(\mathscr{C}, \{y_\nu\}) \qquad (n = 1, 2, \cdots)$$

are positive where the inf has to be taken for all possible basises $\{y_\nu\} \subset B$. We call $E_n^{(B)}(\mathscr{C})$ the nth *absolutely best* linear approximation of $\mathscr{C}$. (We are not concerned with the question whether $E_n^{(B)}(\mathscr{C})$ is attained or not by a system $\{y_\nu\} \subset B$.)

If $\{\xi_\nu\}$ is an orthonormal system, set $E_n^* = \sup e_n$, where the 'sup' has to be taken for all $e_n > 0$ for which

$$e_n \cdot \xi_\kappa \in \mathscr{C} \qquad (k = 1, 2, \cdots, 2n;\ n = 1, 2, \cdots).$$

We shall show that, in many cases, the sequence $\{E_n^*\}$ is equivalent to $\{E_n\}$ and $\{E_n^{(B)}(\mathscr{C})\}$.

Theorem 4. *Let $\{\xi_\nu\}$ be a bounded orthonormal system and $\mathscr{C}$ a closed set $\{E_n\}$-characterizable by best $\{\xi_\nu\}$-approximation. If $\{E_n\}$ is slowly decreasing, then*

$$\{E_n\} \approx \{E_n^{(B)}(\mathscr{C})\} \approx \{E_n^*\}.$$

Since $\mathscr{C}$ is closed, we have

$$E_n^* \cdot \xi_k \in \mathscr{C} \qquad (k = 1, 2, \cdots, 2n;\ n = 1, 2, \cdots).$$

By Theorem 3 it follows then $E_n \geqq K_6 E_n^*$. By Theorem 2, there is a K_4 such that $K_4 E_n \cdot \xi_k \in \mathscr{C}(\{E_n\}, \{\xi_\nu\})$ for $k=1, 2, \cdots, 2n$, hence

$$(5) \qquad K_1 K_4 E_n \cdot \xi_k \in \mathscr{C}(\{K_1 E_n\}, \{\xi_\nu\}) \subset \mathscr{C} \qquad (k=1, 2, \cdots, 2n;\ n=1, 2, \cdots)$$

and therefore $K_1 K_4 E_n \leqq E_n^*$, i.e. $E_n \approx E_n^*$. We have still to prove $E_n^* \approx E_n^{(B)}(\mathscr{C})$. First, we have $E_n^{(B)}(\mathscr{C}) \geqq K_6 E_n^*$ by Theorem 3. But $E_n^{(B)}(\mathscr{C}) \leqq E_n$ and (5) implies

$$K_1 K_4 E_n^{(B)}(\mathscr{C}) \cdot \xi_k \in \mathscr{C} \qquad (k=1, 2, \cdots, 2n;\ n=1, 2, \cdots),$$

hence $K_1 K_4 E_n^{(B)}(\mathscr{C}) \leqq E_n^*$, thus also $E_n^* \approx E_n^{(B)}(\mathscr{C})$ is proved.

Comparison of the best approximations by polynomials and by rational functions

9. The mostly used linear method is the approximation by polynomials; its simplest extension to a non-linear approximation method consists in the substitution of the polynomials by corresponding rational functions. We intend to compare efficacy of these two methods and shall see that, for classes characterizable by best polynomial approximation, the non-linear method provides no better results than the linear method does.

Denote by $r_n(x)$ a rational function of degree n, i.e. $r_n(x) = P_n(x)/Q_n(x)$ where $P_n(x)$ and $Q_n(x)$ are polynomials of degree $\leqq n$. Then, $\mathscr{C}$ being a given class of continuous functions, we call

$$\varrho_n(\mathscr{C}) = \sup_{f \in \mathscr{C}} \inf_{r_n} \|f - r_n\|_C$$

the nth best rational approximation of $\mathscr{C}$ in the space C.

SZÜSZ—TURÁN [8], FREUD [5], and SZABADOS [7] have proved that, for some classes, the best rational approximation in the space C may be essentially better than the best polynomial approximation. The classes considered by these autors are not characterizable by polynomial approximation. But, for the classical polynomially characterizable classes, the best rational approximation is equivalent to the best polynomial approximation. (NEWMAN [6], SZABADOS [7].) We shall see that this phenomenon occurs for all polynomially characterizable classes.

Theorem 5. *Let $\mathscr{C}$ be a class of continuous functions $\{E_n\}$-characterizable by polynomial approximation in the space C where $\{E_n\}$ is slowly decreasing. Then, for the best rational approximation, we have $\{\varrho_n(\mathscr{C})\} \approx \{E_n\}$.*

Consider the function

$$f(x) = K_{11} \cdot \sum_{k=0}^{\infty} (E_{3^k} - E_{3^{k+1}}) T_{3^k}(x)$$

where $T_\nu(x)$ denotes the ν-th normed Chebysheff polynomial and K_{11} an appropriate constant. Let $s_n(x)$ be the nth partial sum of this series. Since $|T_\nu(x)| \leqq \sqrt{2/\pi}$, we have

$$\max_{-1 \leqq x \leqq 1} |f(x) - s_{3^m}(x)| \leqq K_{11} E_{3^m}.$$

A result of BERNSTEIN (cf. ACHYESER [1], p. 79) states that $s_{3^m}(x)$ represents, for $3^m \leqq n < 3^{m+1}$, the best approximating polynomial of degree n and even the best approximating rational function of degree n. Thus, denoting by $\varrho_n(f)$ the best approximation of the function $f(x)$ by rational functions of degree n, we obtain

$$\varrho_n(f) = E_n(f, \{T_n\}) = K_{11} E_n \qquad (3^m \leqq n < 3^{m+1};\ m = 0, 1, \cdots). \tag{7}$$

Because of the characterisability of $\mathscr{C}$ by polynomial approximation, there is a K_1 such that

$$\mathscr{C}(\{K_1 E_n\}, \{T_n\}) \subset \mathscr{C},$$

while $\{E_n\}$ is slowly decreasing; therefore, by appropriate choice of K_{11}, it follows that

$$K_{11} E_{3^m} \leqq K_1 E_{3^{m+1}} \leqq K_1 E_n \qquad (3^m \leqq n < 3^{m+1};\ m = 0, 1, \cdots).$$

Thus, in accordance with (7), we see that $f \in \mathscr{C}$ and so $\varrho_n(\mathscr{C}) \geqq \varrho_n(f) = K_{11} E_n$. This is just our assertion.

References

[1] Н. И. Ахиезер, *Лекции по теории аппроксимации* (Москва—Ленинград, 1947).

[2] G. ALEXITS—S. KNAPOWSKI, A note on the best approximation by linear forms of functions, *Acta Sci. Math.*, **27** (1966), 49—54.

[3] PH. J. DAVIS, *Interpolation and approximation* (New York—Toronto—London, 1963).

[4] N. J. FINE, On the Walsh functions, *Trans. Amer. Math. Soc.*, **63** (1949), 372—414.

[5] G. FREUD, Über die Approximation reeller Funktionen durch gebrochene rationale Funktionen, *Acta Math. Acad. Sci. Hung.*, **17** (1966), 313—324.

[6] D. J. NEWMAN, Rational approximation to $|x|$, *Michigan Math. Journ.*, **11** (1964), 11—14.

[7] J. SZABADOS, Generalization of two theorems of G. Freud concerning the rational approximation, *Studia Sci. Math. Hung.*, **2** (1967), 73—80.

[8] P. SZÜSZ—P. TURÁN, On the constructive theory of functions. I, *Publ. Math. Inst. Hung. Acad. Sci.*, **9** (1964), 495—502; II, *Studia Sci. Math. Hung.*, **1** (1966), 65—69; III, *Studia Sci. Math. Hung.*, **1** (1966), 315—322.

(Received May 11, 1967)

Inequalities and theorems concerning strongly multiplicative systems

By FERENC MÓRICZ in Szeged

Introduction

ALEXITS introduced the following definitions (see [1], p. 186).

The sequence of real measurable functions $\varphi_1(t), \varphi_2(t), \cdots$ defined in the interval [0, 1], is called a multiplicative system if all their finite products are Lebesgue-integrable with

$$(1)\qquad \int_0^1 \varphi_{n_1}(t)\varphi_{n_2}(t)\cdots\varphi_{n_k}(t)\,dt = 0 \qquad (n_1 < n_2 < \cdots < n_k;\ k=1,2,\cdots).$$

The sequence $\{\varphi_n(t)\}$ is called a strongly multiplicative system (SMS) if the system $\{\varphi_{n_1}(t)\varphi_{n_2}(t)\cdots\varphi_{n_k}(t)\}$ is an orthogonal system, i.e.

$$(2)\qquad \int_0^1 \varphi_{n_1}^{\alpha_1}(t)\varphi_{n_2}^{\alpha_2}(t)\cdots\varphi_{n_k}^{\alpha_k}(t)\,dt = 0 \qquad (n_1 < n_2 < \cdots < n_k;\ k = 1, 2, \cdots),$$

where $\alpha_1, \alpha_2, \cdots, \alpha_k$ can be equal to 1 or 2 but at least one element of the sequence $\alpha_1, \alpha_2, \cdots, \alpha_k$ is equal to 1.

The sequence $\{\varphi_n(t)\}$ is called an equinormed strongly multiplicative system (ESMS) if the system $\{\varphi_{n_1}(t)\varphi_{n_2}(t)\cdots\varphi_{n_k}(t)\}$ is an orthogonal and normal system, i.e.

$$\int_0^1 \varphi_n(t)\,dt = 0, \quad \int_0^1 \varphi_n^2(t)\,dt = 1 \qquad (n=1,2,\cdots);$$

$$(3)\qquad \int_0^1 \varphi_{n_1}^{\alpha_1}(t)\varphi_{n_2}^{\alpha_2}(t)\ldots\varphi_{n_k}^{\alpha_k}(t)dt =$$

$$= \int_0^1 \varphi_{n_1}^{\alpha_1}(t)\,dt \int_0^1 \varphi_{n_2}^{\alpha_2}(t)\,dt \cdots \int_0^1 \varphi_{n_k}^{\alpha_k}(t)\,dt \quad (n_1 < n_2 < \cdots < n_k;\ k=1,2,\ldots),$$

where $\alpha_1, \alpha_2, \cdots, \alpha_k$ can be equal to 1 or 2.

Evidently a sequence of independent functions (with mean value 0 and dispersion 1) is an ESMS. Another example is a strongly lacunary sequence of trigonometric functions, i.e. $\{\sqrt{2}\sin 2\pi n_k t\}$ if $n_{k+1}/n_k \geqq 3$ $(k=1, 2, \cdots)$.

ALEXITS proved that an ESMS has the property of the independent functions, i. e. $\sum_{n=1}^{\infty} a_n\varphi_n(t)$ is convergent almost everywhere if and only if $\sum_{n=1}^{\infty} a_n^2 < \infty$. More exactly he proved

Theorem A. *If $\{\varphi_n(t)\}$ in a uniformly bounded* ESMS *then under the condition $\sum_{n=1}^{\infty} a_n^2 < \infty$ the series $\sum_{n=1}^{\infty} a_n\varphi_n(t)$ is convergent almost everywhere. Furthermore, if for every measurable set $E \subset [0, 1]$ and for sufficiently large n the relation*

$$\int_E \varphi_n^2(t)\,dt \geqq C \operatorname{mes}(E) \quad [1]$$

holds (where C is a positive constant depending only on E), and if the series $\sum_{n=1}^{\infty} a_n\varphi_n(t)$ is convergent in a set of positive measure then $\sum_{n=1}^{\infty} a_n^2 < \infty$.

(In [1] this theorem is given in a more general form.)

The aim of the present paper is to study what other properties of the independent functions remain valid for an ESMS. Namely we prove the inequality due to BERNSTEIN and other exponential bounds, furthermore, the central limit theorem and a weaker form of the law of iterated logarithm for ESMS. Let me recall here the well-known forms of these theorems.

We shall use, for any sequence $\{\varphi_n(t)\}$ of functions, the following notations:

$$S_N(t) = \sum_{n=1}^{N} a_n\varphi_n(t), \quad A_N^2 = \sum_{n=1}^{N} a_n^2, \quad M_N = \max_{1 \leqq n \leqq N} |a_n| \qquad (N=1, 2, \cdots).$$

The following inequality is due to BERNSTEIN [2]:

Theorem B. *Let $\{\varphi_n(t)\}$ be a system of independent functions on* [0, 1] *with mean value* 0 *and dispersion* 1, *and uniformly bounded by the constant K, furthermore, let x be a positive real number such that*

$$\vartheta = \frac{xM_N K}{A_N^2} \leqq 1.$$

Then

$$\operatorname{mes}(\{S_N(t) \geqq x\}) \leqq \exp\left\{-\frac{x^2}{2A_N^2}(1-\vartheta)\right\}.$$

---

[1]) mes (E) denotes the Lebesgue measure of the set E.

Here we prove the following analogous form:

Theorem 1. *Let* $\{\varphi_n(t)\}$ *be a uniformly bounded* ESMS, *with bound* K, *and let* x *be a positive real number. Then*

$$\text{(4)} \qquad \operatorname{mes}(\{S_N(t) \geqq x\}) \leqq \exp\left\{-\frac{x^2}{2A_N^2}(1-\theta)\right\} \quad \textit{with} \quad \theta = \frac{xM_NK^3}{A_N^2}.$$

Remark 1. We observe if $S_N(t)$ is replaced by $-S_N(t)$ the conclusion yields

$$\operatorname{mes}(\{|S_N(t)| \geqq x\}) \leqq 2\exp\left\{-\frac{x^2}{2A_N^2}(1-\theta)\right\} \quad \text{with} \quad \theta = \frac{xM_NK^3}{A_N^2}.$$

We show that the reverse inequality also holds if xM_N/A_N^2 is sufficiently small and x^2/A_N^2 is sufficiently large, the analogous form of which can be found in the quoted paper of KOLMOGOROFF [2].

Theorem 2. *Let* $\{\varphi_n(t)\}$ *be a uniformly bounded* ESMS, *with bound* K, *and let* x *be a positive real number. If the inequalities*

$$\text{(5)} \qquad \text{(i)}\ \frac{xM_NK^3}{A_N^2} = \alpha \leqq \frac{1}{2^{13}} \quad \textit{and} \quad \text{(ii)}\ \frac{x^2}{A_N^2} = \beta \geqq 2^{14}$$

are satisfied, then

$$\text{(6)} \qquad \operatorname{mes}(\{S_N(t) \geqq x\}) \geqq \exp\left\{-\frac{x^2}{2A_N^2}(1+\varepsilon)\right\},$$

where

$$\varepsilon = \max\left\{64\sqrt{2\alpha},\ 32\sqrt{\frac{\log\beta}{\beta}}\right\}.$$

MARCINKIEWICZ and ZYGMUND [3] proved the following [2])

Theorem C. *Let* $\{\varphi_n(t)\}$ *be a system of independent functions on* $[0, 1]$, *with mean value* 0 *and dispersion* 1. *Then, for all positive real numbers* $p\ (>1)$, *we have*

$$\text{(7)} \qquad \bar{C}_p A_N \leqq \left\{\int_0^1 \left(\max_{1\leqq n\leqq N} |S_n(t)|^p\, dt\right)\right\}^{\frac{1}{p}} \leqq \bar{D}_p A_N,$$

where $\bar{C}_p$ *and* $\bar{D}_p$ *are positive constants depending only on* p.

An essentially similar result holds for lacunary trigonometric series [3]), too

---

[2]) Here we give the original theorem with a little modification.

[3]) $\sum_{k=1}^{\infty}(a_k \cos n_k t + b_k \sin n_k t)$ in called lacunary if $n_{k+1}/n_k \geqq q > 1$ $(k=1, 2, \ldots)$.

(see [4], v. 1, p. 203). Unfortunately we cannot assert the analogous result for ESMS, but the following result is valid:

Theorem 3. *Let* $\{\varphi_n(t)\}$ *be a uniformly bounded normed* SMS. *Then, for all positive real numbers p, we have*

$$C_p A_N \leqq \left\{\int_0^1 |S_N(t)|^p\,dt\right\}^{\frac{1}{p}} \leqq D_p A_N, \tag{8}$$

where C_p *and* D_p *are positive constants depending only on p. Furthermore, if for* A_n *and a positive real number* λ *we have*

$$A_N \leqq 1 \quad and \quad \lambda \leqq \frac{1}{8eK^6}, \tag{9}$$

then

$$\int_0^1 \exp\{\lambda S_N^2(t)\}\,dt \leqq 2. \tag{10}$$

Moreover, we succeeded in proving the following theorem (for the case Rademacher functions, see [4], v. II, p. 235.):

Theorem 4. *Let* $\{\varphi_n(t)\}$ *be a uniformly bounded normed* SMS. *Then the following estimations are valid:*

$$CA_N \log^+ A_N - C' \leqq \int_0^1 |S_N(t)| \log^+ |S_N(t)|\,dt \leqq CA_N \log^+ A_N + C', \tag{11}$$ [4])

where C and C′ are positive absolute constants.

Remark 2. It will be clear from the proofs that both Theorem 3 and Theorem 4 remain valid if $S_N(t)$ and A_N are replaced by $\sum_{n=1}^{\infty} a_n\varphi_n(t)$ and $A^2 = \sum_{n=1}^{\infty} a_n^2$ in them supposing that $A<\infty$ or $A\leqq 1$, respectively. In particular, if $\sum_{n=1}^{\infty} a_n^2 < \infty$ then the sum of $\sum_{n=1}^{\infty} a_n\varphi_n(t)$ belongs to L^p for every positive real number p.

Concerning the law of iterated logarithm, the basic result, obtained by KOLMOGOROFF [2], reads as follows:

Theorem D. *Let* $\{\varphi_n(t)\}$ *be a system of bounded independent functions on* $[0, 1]$, *with mean value* 0 *and dispersion* 1. *If*

$$\text{(i) } A_N \to \infty, \qquad \text{(ii) } |a_N\varphi_N(t)| \leqq m_N = o\left(\sqrt{\frac{A_N^2}{\log\log A_N^2}}\right), \tag{12}$$

[4]) By $\log^+|u|$ we mean $\log|u|$ wherever $|u|\geqq 1$, and 0 otherwise.

then

$$\text{mes}\left(\left\{\limsup_{N\to\infty} \frac{S_N(t)}{\sqrt{2A_N^2 \log\log A_N^2}} = 1\right\}\right) = 1. \tag{13}$$

For lacunary trigonometric series SALEM and ZYGMUND [5] have shown that under the hypotheses (12) we have (13) with "$\leqq$" instead of "$=$". In this case a complete proof of (13) was given later by ERDŐS and GÁL [6]. Recently, RÉVÉSZ [7] obtained the following result:

Theorem E. *If $\{\varphi_n(t)\}$ is a uniformly bounded* ESMS, *then*

$$\text{mes}\left(\left\{\limsup_{N\to\infty} \frac{\varphi_1(t)+\varphi_2(t)+\ldots+\varphi_N(t)}{\sqrt{N\log\log N}} \leqq 6\right\}\right) = 1.$$

We managed to prove the following result which can be roughly formulated as follows: if the sequence of indices $m_1 < m_2 < \cdots$ is rare enough, then the law of iterated logarithm will be valid for the subsequence $\{S_{m_k}(t)\}$ with "$\leqq$" instead of "$=$". More exactly, we prove

Theorem 5. *Let $\{\varphi_n(t)\}$ be a uniformly bounded* ESMS. *Under the conditions*

$$\text{(i)}\ A_N \to \infty \quad \textit{and} \quad \text{(ii)}\ M_N = o\left(\sqrt{\frac{A_N^2}{\log\log A_N^2}}\right), \tag{14}$$

for every positive real number ε there exists a sequence of natural numbers $N_1 < N_2 < \cdots$ having the following property: if $m_1 < m_2 < \cdots$ is an arbitrary sequence of natural numbers for which $N_k \leqq m_k < N_{k+1}$ $(k=1, 2, \cdots)$, then we have

$$\text{mes}\left(\left\{\limsup_{k\to\infty} \frac{S_{m_k}(t)}{\sqrt{2A_{m_k}^2 \log\log A_{m_k}^2}} \leqq 1+\varepsilon\right\}\right) = 1. \tag{15}$$

Remark 3. It will be clear from the proof that if we had the stronger inequality (7) for a uniformly bounded ESMS too, then under the hypotheses (14) we could assert also (13) with "$\leqq$" instead of "$=$". Unfortunately, we only have the weaker inequality (8) for a uniformly bounded normed SMS.

A number of authors have generalized the central limit theorem for the lacunary trigonometric series. The most general result is due to SALEM and ZYGMUND [8], who state the following

Theorem F. *Let $S_N(t)$ denote the Nth partial sum of the lacunary trigonometric series $\sum_{k=1}^{\infty} (a_k \cos n_k t + b_k \sin n_k t)$, $n_{k+1}/n_k \geqq q > 1$ $(k=1, 2, \cdots)$, and let $a_1, a_2, \cdots; b_1, b_2, \cdots$ be arbitrary sequences of real numbers for which*

$$C_N = \left\{\frac{1}{2}\sum_{k=1}^{N}(a_k^2+b_k^2)\right\}^{\frac{1}{2}} \to \infty \quad \text{and} \quad \{a_N^2+b_N^2\}^{\frac{1}{2}} = o(C_N).$$

Then, for any set $E \subset [0, 2\pi]$ of positive measure, the distribution functions

$$F_N(y; E) = \frac{\operatorname{mes}(\{t \in E: S_N(t)/C_N \leqq y\})}{\operatorname{mes}(E)} \qquad (N = 1, 2, \cdots)$$

tend to the Gaussian distribution with mean value 0 *and dispersion* 1.

We obtained the following result:

Theorem 6. *Let $\{\varphi_n(t)\}$ be a uniformly bounded* ESMS. *If*

(16) $$A_N \to \infty \quad \text{and} \quad a_N = o(A_N);$$

then the distribution functions

(17) $$F_N(y) = \operatorname{mes}\left(\left\{\frac{S_N(t)}{A_N} \leqq y\right\}\right) \qquad (N = 1, 2, \cdots)$$

tend pointwise to the Gaussian distribution function

$$G(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{u^2}{2}} du.$$

This theorem contains a result of Révész [7] (case $a_n = 1$ for every n).

§ 1. The proof of Theorem 1 and Theorem 2

The following lemma has a fundamental significance in the proof of Theorem 1 and Theorem 2.

Lemma 1. *Let λ be an arbitrary non-negative real number. Then*

(18) $$\exp\left\{\frac{\lambda^2 A_N^2}{2}\left(1 - \frac{\lambda^2 M_N^2}{2} - \lambda M_N K^3\right)\right\} \leqq \int_0^1 \exp\{\lambda S_N(t)\}\, dt \leqq$$
$$\leqq \exp\left\{\frac{\lambda^2 A_N^2}{2}(1 + \lambda M_N K^3)\right\}.$$

Proof. For every real number u, we have that

(19) $$\left|\log\left(1 + u + \frac{u^2}{2}\right) - u\right| \leqq \frac{|u|^3}{2}.$$ [5]

[5]) (19) follows from the sharper estimates: $0 \leqq u - \log(1 + u + u^2/2) \leqq u^3/3$ for $u \geqq 0$ and $u^3/3 \leqq$ $\leqq u - \log(1 + u + u^2/2) \leqq 0$ for $u \leqq 0$. We only have to remark that the function $\varkappa(u) =$ $= u - \log(1 + u + u^2/2)$ is non-decreasing $-\infty < u < \infty$ and $\varkappa(0) = 0$, and that the function $\mu(u) =$ $= u - \log(1 + u + u^2/2) - u^3/3$ is non-increasing and $\mu(0) = 0$.

Applying this inequality, we get that

$$\exp\{\lambda a_n \varphi_n(t)\} = \left(1+\lambda a_n \varphi_n(t) + \frac{\lambda^2 a_n^2 \varphi_n^2(t)}{2}\right) \exp\{R_n(t)\},$$

where

$$|R_n(t)| \leqq \frac{\lambda^3 K^3 M_N a_n^2}{2} \qquad (n=1,2,\ldots,N).$$

Hence

$$(20)\qquad \int_0^1 \exp\{\lambda S_N(t)\}\,dt \leqq \prod_{n=1}^{N} \exp\left\{\frac{\lambda^3 K^3 M_N a_n^2}{2}\right\} \int_0^1 \prod_{n=1}^{N} \left(1+\lambda a_n \varphi_n(t) + \frac{\lambda^2 a_n^2 \varphi_n^2(t)}{2}\right) dt.$$

By a simple calculation we get that

$$(21)\qquad \prod_{n=1}^{N} \exp\left\{\frac{\lambda^3 K^3 M_N a_n^2}{2}\right\} = \exp\left\{\frac{\lambda^3 K^3 M_N A_N^2}{2}\right\},$$

furthermore,

$$\int_0^1 \prod_{n=1}^{N} \left(1+\lambda a_n \varphi_n(t) + \frac{\lambda^2 a_n^2 \varphi_n^2(t)}{2}\right) dt =$$

$$= 1 + \sum{}' \lambda^k a_{n_1} \cdots a_{n_k} \int_0^1 \varphi_{n_1}(t) \cdots \varphi_{n_k}(t)\,dt + \sum{}' \frac{\lambda^{2k}}{2^k} a_{n_1}^2 \cdots a_{n_k}^2 \int_0^1 \varphi_{n_1}^2(t) \cdots \varphi_{n_k}^2(t)\,dt +$$

$$+ \sum{}'' \frac{\lambda^{k+2l}}{2^l} a_{n_1} \cdots a_{n_k} a_{m_1}^2 \cdots a_{m_l}^2 \int_0^1 \varphi_{n_1}(t) \cdots \varphi_{n_k}(t) \varphi_{m_1}^2(t) \ldots \varphi_{m_l}^2(t)\,dt = 1 + I + J + K,$$

where the sum Σ' is extended for all systems of integer values $(1 \leqq)\, n_1 < \cdots < n_k (\leqq N)$ $(1 \leqq k \leqq N)$, the sum Σ'' is extended for all systems of integer values $(1 \leqq)\, n_1 < \cdots < n_k\, (\leqq N)$ and $(1 \leqq)\, m_1 < \cdots < m_l\, (\leqq N)$ for which $n_i \neq m_j$ $(1 \leqq i \leqq k,\ 1 \leqq j \leqq l)$; $1 \leqq k$, $1 \leqq l$ and $k+l \leqq N$. It follows from (3) that $I = K = 0$ and

$$J = \sum{}' \frac{\lambda^2 a_{n_1}^2}{2} \cdots \frac{\lambda^2 a_{n_k}^2}{2}.$$

So we obtain that

$$(22)\qquad \int_0^1 \prod_{n=1}^{N} \left(1+\lambda a_n \varphi_n(t) + \frac{\lambda^2 a_n^2 \varphi_n^2(t)}{2}\right) dt = \prod_{n=1}^{N} \left(1+\frac{\lambda^2 a_n^2}{2}\right).$$

Applying the well-known inequality

$$1+u \leqq e^u \quad \text{if} \quad u \geqq 0,$$

from (20), (21) and (22) we get that

$$\int_0^1 \exp\{\lambda S_N(t)\}\,dt \leqq \exp\left\{\frac{\lambda^3 K^3 M_N A_N^2}{2}\right\} \prod_{n=1}^{N} \exp\left\{\frac{\lambda^2 a_n^2}{2}\right\} =$$

$$= \exp\left\{\frac{\lambda^3 K^3 M_N A_N^2}{2}\right\} \exp\left\{\frac{\lambda^2 A_N^2}{2}\right\} = \exp\left\{\frac{\lambda^2 A_N^2}{2}(1+\lambda M_N K^3)\right\}.$$

This shows that the right-hand inequality of (18) is true.

We get similarly to (20) that

(23)
$$\int_0^1 \exp\{\lambda S_N(t)\}\,dt \geqq$$
$$\geqq \prod_{n=1}^{N} \exp\left\{-\frac{\lambda^3 K^3 M_N a_n^2}{2}\right\} \int_0^1 \prod_{n=1}^{N} \left(1+\lambda a_n \varphi_n(t) + \frac{\lambda^2 a_n^2 \varphi_n^2(t)}{2}\right) dt.$$

Applying the simple inequality

$$e^{u(1-u)} \leqq 1+u \quad \text{if} \quad u \geqq 0, \text{ [6]}$$

we get that

$$\prod_{n=1}^{N}\left(1+\frac{\lambda^2 a_n^2}{2}\right) \geqq \prod_{n=1}^{N} \exp\left\{\frac{\lambda^2 a_n^2}{2}\left(1-\frac{\lambda^2 M_N^2}{2}\right)\right\} = \exp\left\{\frac{\lambda^2 A_N^2}{2}\left(1-\frac{\lambda^2 M_N^2}{2}\right)\right\}.$$

This and (21), (22), (23) show that the left-hand inequality of (18) is true. This completes the proof of Lemma 1.

In the proof which follows we use some ideas from the classical paper of Kolmogoroff [2]. First we introduce the notation

$$W_N(x) = \operatorname{mes}(\{S_N(t) \geqq x\}) \quad \text{for} \quad x > 0.$$

Proof of Theorem 1. Let λ be a positive real number determined later on. It is obvious that

$$W_N(x) e^{\lambda x} \leqq \int_0^1 \exp\{\lambda S_N(t)\}\,dt,$$

and it follows from (18) that

(24)
$$W_N(x) \leqq \exp\left\{-\lambda x + \frac{\lambda^2 A_N^2}{2}(1+\lambda M_N K^3)\right\}.$$

[6]) This sharper inequality $u-u^2/2 \leqq \log(1+u)$ $(u \geqq 0)$ is also true, as the function $\varkappa(u) = \log(1+u)-u+u^2/2$ is non-decreasing and $\varkappa(0)=0$.

Setting $\lambda = x/A_N^2$ we get

$$W_N(x) \leqq \exp\left\{-\frac{x^2}{A_N^2}+\frac{x^2}{2A_N^2}\left(1+\frac{xM_NK^3}{A_N^2}\right)\right\} = \exp\left\{-\frac{x^2}{2A_N^2}\left(1-\frac{xM_NK^3}{A_N^2}\right)\right\}.$$

This proves (4) and finishes the proof of Theorem 1.

We need the next two lemmas only for the proof of Theorem 2.

Lemma 2. *If*

(25) $$\frac{xM_NK^3}{A_N^2} \leqq \frac{1}{2},$$

then

(26) $$W_N(x) \leqq \exp\left\{-\frac{x^2}{4A_N^2}\right\}.$$

Proof. As $\theta \leqq 1/2$ by (25), on the ground of Theorem 1, (26) holds obviously.

Lemma 3. *If*

(27) $$\frac{xM_NK^3}{A_N^2} \geqq \frac{1}{2},$$

then

(28) $$W_N(x) \leqq \exp\left\{-\frac{x}{8M_NK^3}\right\}.$$

Proof. In the proof of Theorem 1 we obtained (24), where λ is an arbitrary positive real number. Now we set

$$\lambda = \frac{1}{2M_NK^3}.$$

From (24) and (27) we get that

$$W_N(x) \leqq \exp\left\{-\frac{x}{2M_NK^3}+\frac{A_N^2}{8M_N^2K^6}\left(1+\frac{1}{2}\right)\right\} \leqq$$

$$\leqq \exp\left\{-\frac{x}{2M_NK^3}+\frac{3x}{8M_NK^3}\right\} = \exp\left\{-\frac{x}{8M_NK^3}\right\}.$$

So the proof of Lemma 3 is ready.

The proof of the inequality (6) is much more involved. The following argument follows closely that of a similar theorem in the paper of KOLMOGOROFF [2].

Proof of Theorem 2. Let $\delta = \varepsilon/8$. Then

(29) $$\delta^2 = \max(128\alpha, 16(\log\beta)/\beta).$$

Hence it follows that

(30) $$\delta^2 \leqq 1/64, \quad \delta \leqq 1/8 \quad \text{and} \quad \delta > 2\delta^2.$$

We set now

$$\lambda = x/[A_N^2(1-\delta)]$$

so that, by (30),

$$x/A_N^2 < \lambda < 2x/A_N^2,$$

furthermore, in virtue of (5) we have

(31) $$\lambda M_N K^3 < 2\alpha \leqq 2^{-12}$$

and

(32) $$\lambda^2 A_N^2 > \beta \geqq 2^{14}.$$

On account of Lemma 1

$$\int_0^1 \exp\{\lambda S_N(t)\}\,dt \geqq \exp\{\tfrac{1}{2}\lambda^2 A_N^2(1-\tfrac{1}{2}\lambda^2 M_N^2 - \lambda M_N K^3)\}.$$

By (29) and (31), we get

$$\tfrac{1}{2}\lambda^2 M_N^2 + \lambda M_N K^3 < \tfrac{1}{2}(2\alpha)^2 + 2\alpha \leqq 4\alpha \leqq \delta^2/4.$$

Hence

(33) $$\int_0^1 \exp\{\lambda S_N(t)\}\,dt \geqq \exp\{\tfrac{1}{2}\lambda^2 A_N^2(1-\delta^2/4)\}.$$

On the other hand, integrating by parts, we obtain

$$\int_0^1 \exp\{\lambda S_N(t)\}\,dt = -\int_{-\infty}^{+\infty} e^{\lambda y}\,dW_N(y) = \lambda \int_{-\infty}^{+\infty} e^{\lambda y} W_N(y)\,dy.$$

We decompose the interval $(-\infty, +\infty)$ of integration into the five intervals $I_1 = (-\infty, 0]$, $I_2 = (0, \lambda A_N^2(1-\delta)]$, $I_3 = (\lambda A_N^2(1-\delta), \lambda A_N^2(1+\delta)]$, $I_4 = (\lambda A_N^2(1+\delta), 8\lambda A_N^2]$ and $I_5 = (8\lambda A_N^2, +\infty)$ and search for upper bounds of the integral over I_1 and I_5 and over I_2 and I_4.

We have

(34) $$J_1 = \lambda \int_{-\infty}^{0} e^{\lambda y} W_N(y)\,dy \leqq \lambda \int_{-\infty}^{0} e^{\lambda y}\,dy = 1$$

because $W_N(y) \leqq 1$ for all y. According to (31), Lemma 3, and Lemma 2, we have on I_5

$$W_N(y) \leqq \exp\left\{-\frac{y}{8M_N K^3}\right\} \leqq e^{-2\lambda y} \quad \text{for} \quad y \geqq \frac{A_N^2}{2M_N K^3},$$

and

$$W_N(y) \leqq \exp\left\{-\frac{y^2}{4A_N^2}\right\} \leqq e^{-2\lambda y} \quad \text{for} \quad 8\lambda A_N^2 \leqq y \leqq \frac{A_N^2}{2M_N K^3}.$$

Therefore

(35) $$J_5 = \lambda \int_{8\lambda A_N^2}^{+\infty} e^{\lambda y} W_N(y)\,dy \leqq \lambda \int_{8\lambda A_N^2}^{+\infty} e^{-\lambda y}\,dy < 1.$$

It follows, by (30), (32), and (33), that

$$\int_0^1 \exp\{\lambda S_N(t)\}\,dt > 8.$$

Hence, on account of (34) and (35), we can see that

(36) $$J_1+J_5 < \frac{1}{4}\int_0^1 \exp\{\lambda S_N(t)\}\,dt.$$

On the intervals I_2 and I_4, applying Theorem 1, we have

$$e^{\lambda y}W_N(y) \leqq \exp\left\{\lambda y-\frac{y^2}{2A_N^2}\left(1-\frac{\delta^2}{8}\right)\right\} = e^{\mu(y)}$$

because, by (29) and (31), we obtain that

$$\theta = \frac{yM_NK^3}{A_N^2} \leqq 8\lambda M_NK^3 < 16\alpha \leqq \frac{\delta^2}{8}.$$

The quadratic expression $\mu(y)$ attains its maximum for $y=\lambda A_N^2(1-\delta^2/8)^{-1}$ which lies in I_3. Hence, in the intervals I_2 and I_4 $\mu(y)$ is majorized by its value at $y=\lambda A_N^2(1+\delta)$ (as $\lambda A_N^2(1+\delta)$ lies closer to the right endpoint of the interval I_3 than to the left one). This value does not exceed

$$\lambda^2A_N^2(1+\delta)-\frac{\lambda^2A_N^2}{2}(1+\delta)^2\left(1-\frac{\delta^2}{8}\right) =$$

$$= \frac{\lambda^2A_N^2}{2}\left(1-\delta^2+\frac{\delta^2}{8}(1+\delta)^2\right) < \frac{\lambda^2A_N^2}{2}\left(1-\frac{\delta^2}{2}\right).$$

Therefore

$$J_2+J_4 = \lambda\left\{\int_0^{\lambda A_N^2(1-\delta)} + \int_{\lambda A_N^2(1+\delta)}^{8\lambda A_N^2}\right\}e^{\lambda y}W_N(y)\,dy <$$

$$< \lambda\int_0^{8\lambda A_N^2}\exp\left\{\frac{\lambda^2A_N^2}{2}\left(1-\frac{\delta^2}{2}\right)\right\}dy = 8\lambda^2A_N^2\exp\left\{\frac{\lambda^2A_N^2}{2}\left(1-\frac{\delta^2}{2}\right)\right\}.$$

From (5), (29) and (32), we get the following estimates:

(37)
$$\log 2^7\beta < 2\log\beta \leqq \frac{\beta\delta^2}{8},$$
$$\log 2^5\lambda^2A_N^2 < 2\log\lambda^2A_N^2 \leqq \frac{\lambda^2A_N^2}{8}\delta^2$$

because $\lambda^2 A_N^2 > \beta$ and $\log u/u$ is a decreasing function if $u \geqq e$. So we have from (33)

$$(38) \qquad J_2 + J_4 < \frac{1}{4} \exp\left\{\frac{\lambda^2 A_N^2}{2}\left(1 - \frac{\delta^2}{4}\right)\right\} < \frac{1}{4}\int_0^1 \exp\{\lambda S_N(t)\}\, dt.$$

It follows, from (36) and (38)

$$(39) \qquad J_3 = \int_{\lambda A_N^2(1-\delta)}^{\lambda A_N^2(1+\delta)} e^{\lambda y} w_N(y)\, dy > \frac{1}{2}\int_0^1 \exp\{\lambda S_N(t)\}\, dt > \frac{1}{2}\exp\left\{\frac{\lambda^2 A_N^2}{2}(1-\delta)\right\}$$

because $\delta > \delta^2/4$. Since $W_N(y)$ is a decreasing function, on account of the definition of λ, we have that

$$(40) \qquad J_3 < 2\lambda^2 A_N^2 \delta \exp\{\lambda^2 A_N^2 (1+\delta)\} W_N(x).$$

From (39) and (40) we obtain that

$$W_N(x) > \frac{1}{4\lambda^2 A_N^2 \delta} \exp\left\{-\frac{\lambda^2 A_N^2}{2}(1+3\delta)\right\}.$$

Similarly to (37), we have

$$\log 4\lambda^2 A_N^2 \delta < \tfrac{1}{2}\lambda^2 A_N^2 \delta$$

as $4\lambda^2 A_N^2 \delta > 4\beta\delta \geqq 16\sqrt{\beta \log \beta} \geqq 2^{12}$, and $\log u/u \leqq 1/8$ if $u \geqq 2^{12}$. So we get that

$$W_N(x) \exp\left\{-\frac{\lambda^2 A_N^2}{2}(1+4\delta)\right\} = \exp\left\{-\frac{x^2}{2A_N^2(1-\delta)^2}(1+4\delta)\right\} >$$

$$> \exp\left\{-\frac{x^2}{2A_N^2}(1+8\delta)\right\} = \exp\left\{-\frac{x^2}{2A_N^2}(1+\varepsilon)\right\}$$

because $\delta = \varepsilon/8$ and, by (30), $\delta \leqq 1/8$. This yields (6) with a suitable ε, by (29). And this is what we wished to prove.

§ 2. The proof of Theorem 3 and Theorem 4

We need a result concerning series with RADEMACHER's functions defined as follows

$$r_n(x) = \operatorname{sign} \sin 2^{n+1}\pi x \qquad (0 \leqq x \leqq 1;\ n = 1, 2, \cdots).$$

The following assertion holds (see [4], v. 1, p. 213):

Lemma 4. *If p is an arbitrary positive real number then*

$$(41) \qquad \left\{\int_0^1 \left|\sum_{n=1}^N a_n r_n(x)\right|^p dx\right\}^{\frac{1}{p}} \leqq 2p^{\frac{1}{2}} \left\{\sum_{n=1}^N a_n^2\right\}^{\frac{1}{2}}.$$

Proof of Theorem 3. This argument will follow closely that on page 215 of [4]. First we show (10), hence then the second inequality (8) immediately follows. The first inequality (8) follows from the second one by a simple argument.

Let K denote a common bound for the system $\{\varphi_n(t)\}$, i.e.

$$|\varphi_n(t)| \leqq K \qquad (0 \leqq t \leqq 1;\ n=1, 2, \cdots).$$

Furthermore, let μ denote a sufficiently small positive real number. We set

$$S_N(t; x) = \sum_{n=1}^{N} a_n \varphi_n(t) r_n(x).$$

Applying (41), with a simple calculation we get

$$\begin{aligned}\int_0^1 \exp\{\mu S_N^2(t;x)\}\,dx &= \sum_{k=0}^{\infty} \frac{\mu^k}{k!} \int_0^1 S_N^{2k}(t;x)\,dx \leqq \\ \leqq \sum_{k=0}^{\infty} \frac{k^k}{k!} \left\{4\mu \sum_{n=1}^{N} a_n^2 \varphi_n^2(t)\right\}^k &\leqq \sum_{k=0}^{\infty} \left\{4\mu e \sum_{n=1}^{N} a_n^2 \varphi_n^2(t)\right\}^k\end{aligned} \tag{42}$$

since $k^k/k! < \sum_{n=0}^{\infty} k^n/n! = e^k$. On the basis of (9i)

$$4e\mu \sum_{n=1}^{N} a_n^2 \varphi_n^2(t) \leqq 4e\mu K^2 A_N^2 \leqq \frac{1}{2}$$

if

$$\mu \leqq \frac{1}{8eK^2}. \tag{43}$$

Thus, the series on the right of (42) uniformly converges in t $(0 \leqq t \leqq 1)$, and its sum does not exceed 2.

Integrate (42) over $0 \leqq t \leqq 1$ and interchange the order of integration; then

$$\int_0^1 dx \int_0^1 \exp\{\mu S_N^2(t;x)\}\,dt \leqq 2.$$

It follows that there is a dyadic irrational [7]) number x_0 $(0 < x_0 < 1)$ for which

$$\int_0^1 \exp\{\mu S_N^2(t;x_0)\}\,dt \leqq 2. \tag{44}$$

Consider the following representation of $S_N(t)$

$$S_N(t) = K^2 \int_0^1 S_N(u;x_0) P_N(t,u;x_0)\,du, \tag{45}$$

[7]) x_0 is dyadic irrational number if $x_0 \neq p/2^q$ where p and q are positive natural numbers.

where

$$P_N(t,u;x_0) = \prod_{n=1}^{N}\left(1+\frac{\varphi_n(t)\varphi_n(u)r_n(x_0)}{K^2}\right).$$

First of all, $P_N(t, u; x_0)$ is non-negative. Furthermore, $P_N(t, u; x_0)$ is symmetric in t and u, and

$$\int_0^1 P_N(t,u;x_0)\,du = 1+\sum{}'\frac{1}{K^{2k}}\varphi_{n_1}(t)\cdots\varphi_{n_k}(t)r_{n_1}(x_0)\cdots r_{n_k}(x_0)\int_0^1\varphi_{n_1}(u)\cdots\varphi_{n_k}(u)\,du,$$

where the sum Σ' is extended for all systems of integer values $(1\leqq)n_1<\cdots<n_k(\leqq N)$ $(1\leqq k\leqq N)$. It follows from (2) that

$$\int_0^1 P_N(t,u;x_0)\,du = 1. \tag{46}$$

As to the representation (45), after carrying out the multiplications and integrating term by term, the right-hand side can be written as follows:

$$K^2\sum_{n=1}^{N}a_n r_n(x_0)\int_0^1\varphi_n(u)\,du+\sum_{n=1}^{N}\sum_{m=1}^{N}a_n r_n(x_0)r_m(x_0)\varphi_m(t)\int_0^1\varphi_n(u)\varphi_m(u)\,du+$$

$$+\sum_{n=1}^{N}a_n r_n(x_0)\sum{}''\frac{1}{K^{2k-2}}\varphi_{n_1}(t)\cdots\varphi_{n_k}(t)r_{n_1}(x_0)\cdots r_{n_k}(x_0)\int_0^1\varphi_n(u)\varphi_{n_1}(u)\cdots$$

$$\cdots\varphi_{n_k}(u)\,du = I+J+K,$$

where the sum Σ'' is extended for all systems of integer values $(1\leqq)n_1<\cdots<n_k(\leqq N)$ $(2\leqq k\leqq N)$. Taking into account that the functions $\varphi_n(t)$ are normed, it follows from (2) that $I=K=0$ and

$$J=\sum_{n=1}^{N}a_n r_n^2(x_0)\varphi_n(t) = S_N(t)$$

because $r_n^2(x_0)=1$ $(1\leqq n\leqq N)$. This proves (45).

The function $\chi(u)=\exp(\mu u^2)$ is increasing and convex for $u\geqq 0$. On account of (46), JENSEN's inequality (see [4], v. I, p. 24) gives

$$\chi\left(\frac{|S_N(t)|}{K^2}\right)=\chi\left(\int_0^1|S_N(t;x_0)|\cdot P_N(t,u;x_0)\,du\right)\leqq$$

$$\leqq\int_0^1\chi(|S_N(t;x_0)|)P_N(t,u;x_0)\,du.$$

Integrate this over $0 \leqq t \leqq 1$ and interchange the order of integration, from (44) and (46), we get that

$$\int_0^1 \chi\left(\frac{|S_N(t)|}{K^2}\right) dt \leqq \int_0^1 \chi(|S_N(t; x_0)|)\, du \int_0^1 P_N(t, u; x_0)\, dt =$$

$$= \int_0^1 \exp\{\mu S_N^2(u; x_0)\}\, du \leqq 2.$$

Now we set $\mu = K^4\lambda$ then, it follows from (9ii), this μ satisfies (43). We finished the proof of (10).

As to the second inequality (8), we set

$$S_N^*(t) = S_N(t)/A_N.$$

The condition (9i) is satisfied by the coefficients of $S_N^*(t)$. Thus, if λ is sufficiently small then, on account of (10),

$$\int_0^1 \exp\{\lambda S_N^{*2}(t)\}\, dt = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \int_0^1 |S_N^*(t)|^{2k}\, dt \leqq 2.$$

Hence it follows for every k

$$\int_0^1 |S_N^*(t)|^{2k}\, dt \leqq \frac{2k!}{\lambda^k},$$

that is

$$\left\{\int_0^1 |S_N(t)|^{2k}\, dt\right\}^{\frac{1}{2k}} \leqq D_{2k} A_N,$$

where, choosing λ equal to $(8eK^6)^{-1}$,

$$D_{2k} = \{2 \cdot 8^k e^k K^{6k} k!\}^{\frac{1}{2k}} \leqq 8K^3 k^{\frac{1}{2}} \qquad (k = 1, 2, \cdots).$$

If now for the positive real number p we have $2k-2 \leqq p < 2k$ with a suitable natural number k then it is sufficient to remark that

$$\left\{\int_0^1 |S_N(t)|^p\, dt\right\}^{\frac{1}{p}} \leqq \left\{\int_0^1 |S_N(t)|^{2k} dt\right\}^{\frac{1}{2k}}$$

(see [4], v. I, p. 25).

It still remains to prove the first inequality (8). This is immediate for $p \geqq 2$, for then

$$\left\{\int_0^1 |S_N(t)|^p\, dt\right\}^{\frac{1}{p}} \geqq \left\{\int_0^1 S_N^2(t)\, dt\right\}^{\frac{1}{2}} = A_N.$$

If $0<p<2$, let α_1 and α_2 be positive and such that $\alpha_1+\alpha_2=1$, $2=p\alpha_1+4\alpha_2$. The function

$$\int_0^1 |S_N(t)|^\alpha \, dt$$

being logarithmically convex in α (see [4], v. I, p. 25),

$$A_N^2 = \int_0^1 S_N^2(t)\,dt \leqq \left\{\int_0^1 |S_N(t)|^p\,dt\right\}^{\alpha_1} \left\{\int_0^1 S_N^4(t)\,dt\right\}^{\alpha_2} \leqq \left\{\int_0^1 |S_N(t)|^p\,dt\right\}^{\alpha_1} (D_4 A_N)^{\alpha_2},$$

which gives

$$\left\{\int_0^1 |S_N(t)|^p\,dt\right\}^{\frac{1}{p}} \geqq D_4^{-(4-2p)/p} A_N.$$

This completes the proof of Theorem 3.

The following lemma needs in the proof of Theorem 4.

Lemma 5. *There exist positive absolute constants η ($\geqq 1$) and ε such that*

$$\text{mes}\left(\{|S_N(t)| \geqq \eta A_N\}\right) \geqq \varepsilon.$$

The proof is based on a lemma which can find in [4], Chapter V, (8. 26), and it goes, applying (8), word by word as there.

As to the proof of Theorem 4, it can be proved in the same way as the analogous assertion for Rademacher functions, see [4], Chapter XV., (5. 14), applying (8) to prove the second inequality (11) and Lemma 5 the first one.

§ 3. The proof of Theorem 5

We are going to apply the following well-known assertion: if the sequence $\{E_k\}$ of measurable subsets of the interval [0, 1] is such that

$$\sum_{k=1}^{\infty} \text{mes}(E_k) < \infty,$$

then

$$\text{mes}\left(\limsup_{k\to\infty} E_k\right) = 0. \text{ [8]}$$

For the arbitrary fixed positive real number $\varepsilon(<1)$, we choose the real number $\eta(<1)$ such that

$$\eta(1+\varepsilon) > 1, \quad \text{e.g.} \quad \eta = 1-\frac{\varepsilon}{2}. \tag{47}$$

[8]) $\limsup_{k\to\infty} E_k$ is the set of all those points which belong to infinitely many E_k.

Now we define the sequence of indices $n_1 \leqq n_2 \leqq \cdots$ in the following manner:

(48) $$A^2_{n_k-1} \leqq e^{k^\eta} < A^2_{n_k} \qquad (k=1,2,\cdots).$$

This is possible in virtue of (14i).

We set

$$E_k = \left\{ \frac{S_{n_k}(t)}{\sqrt{2A^2_{n_k}\log\log A^2_{n_k}}} \geqq 1+\varepsilon \right\}.$$

On account of Theorem 1, we get that

(49) $$\operatorname{mes}(E_k) = W_{n_k}\big((1+\varepsilon)\sqrt{2A^2_{n_k}\log\log A^2_{n_k}}\big) \leqq \exp\{-(1+\varepsilon)^2(1-\theta)\log\log A^2_{n_k}\},$$

where

$$\theta = (1+\varepsilon)K^3 M_{n_k}\sqrt{\frac{2\log\log A^2_{n_k}}{A^2_{n_k}}}.$$

Here K denotes a common bound of the system $\{\varphi_n(t)\}$. Taking into account (14ii), this θ tends to 0 if k tends to ∞. Thus, θ is not greater than $\varepsilon/2$ if k is sufficiently large. Continuing the estimation (49) we obtain

$$\operatorname{mes}(E_k) \leqq \exp\left\{-(1+\varepsilon)^2\left(1-\frac{\varepsilon}{2}\right)\log\log A^2_{n_k}\right\} \leqq$$

$$\leqq \exp\{-(1+\varepsilon)\log\log A^2_{n_k}\} = (\log A^2_{n_k})^{-(1+\varepsilon)}.$$

By (48), hence we get

$$\sum_{k=1}^{\infty}\operatorname{mes}(E_k) \leqq \sum_{k=1}^{\infty}\frac{1}{k^{\eta(1+\varepsilon)}} < \infty$$

in virtue of (47). So, we have shown that in the case of sequence of indices defined by (48), we have that

$$\limsup_{k\to\infty}\frac{S_{n_k}(t)}{\sqrt{2A^2_{n_k}\log\log A^2_{n_k}}} \leqq 1+\varepsilon$$

holds almost everywhere.

Let $m_1 \leqq m_2 \leqq \cdots$ be an arbitrary sequence of indices for which

(50) $$\begin{aligned} n_k \leqq m_k < n_{k+1} \quad &\text{if} \quad n_k \neq n_{k+1}, \quad \text{and} \\ n_k = m_k \quad &\text{if} \quad n_k = n_{k+1} \qquad (k=1,2,\cdots). \end{aligned}$$

It is sufficient to show that

$$\Delta_k(t) = \frac{S_{m_k}(t) - S_{n_k}(t)}{\sqrt{2A^2_{n_k}\log\log A^2_{n_k}}}$$

tends to 0 for almost every t $(0 \leqq t \leqq 1)$.

It is obvious that $\Delta_k(t)=0$ if $n_k = n_{k+1}$. Therefore, in the following we assume that $n_k < n_{k+1}$. Let p be a positive real number to be determined later on. Applying Theorem 3, we get

$$\int_0^1 \Delta_k^{2p}(t)\,dt \leqq D_{2p}^{2p}\left(\frac{A_{n_{k+1}-1}^2 - A_{n_k}^2}{2A_{n_k}^2 \log\log A_{n_k}^2}\right)^p \leqq \tag{51}$$

$$\leqq D_{2p}^{2p}\left(\frac{e^{(k+1)^\eta} - e^{k^\eta}}{2e^{k^\eta}\eta\log k}\right)^p \leqq D_{2p}^{2p}\left(\frac{e^{(k+1)^\eta - k^\eta} - 1}{2\eta}\right)^p.$$

We apply the following inequalities:

$$e^u \leqq 1+3u \quad \text{if} \quad 0 \leqq u \leqq 1,$$

and

$$(u+1)^\eta - u^\eta \leqq \frac{\eta}{u^{1-\eta}} \quad \text{if} \quad u \geqq 0 \qquad (0<\eta<1).$$ [9]

On account of these and (51), we obtain

$$\int_0^1 \Delta_k^{2p}(t)\,dt \leqq D_{2p}^{2p}\left(\frac{3((k+1)^\eta - k^\eta)}{2\eta}\right)^p \leqq \left(\frac{3}{2}\right)^p D_{2p}^{2p}\frac{1}{k^{p(1-\eta)}}.$$

If we fix the real number p so large that $p(1-\eta)>1$ is satisfied then

$$\sum_{k=1}^\infty \int_0^1 \Delta_k^{2p}(t)\,dt \leqq \left(\frac{3}{2}\right)^p D_{2p}^{2p}\sum_{k=1}^\infty \frac{1}{k^{p(1-\eta)}} < \infty.$$

It follows from the theorem of Beppo Levi that

$$\Delta_k^{2p}(t) \to 0 \qquad (k\to\infty)$$

almost everywhere. As p is fixed, therefore we have proved the assertion (15).

Now, we set $N_1 = n_1$, and let N_l $(l \geqq 2)$ be equal to the first index n_k for which $n_k > N_{l-1}$. It is obvious that the sequence $N_1 < N_2 < \cdots$ has the property as asserted in Theorem 5.

§ 4. The proof of Theorem 6

Lemma 6. *Let $\{b_n\}$ be a sequence of non-negative real numbers. If*

$$\text{(i)}\ s_N = \sum_{n=1}^N b_n \to \infty \quad \textit{and} \quad \text{(ii)}\ b_N = o(s_N), \tag{52}$$

then for arbitrary positive real number $\alpha(>1)$, we have

$$\sum_{n=1}^N b_n^\alpha = o(s_N^\alpha). \tag{53}$$

[9]) This inequality follows from the fact that the function $u^\eta (0<\eta<1)$ is concave for $u \geqq 0$.

Proof. Let ε be an arbitrary positive real number. We choose the natural number n_0 in such a manner that

$$\frac{b_n}{s_n} \leqq \left(\frac{\varepsilon}{2}\right)^{\frac{1}{\alpha-1}} \quad \text{if} \quad n \geqq n_0.$$

This is possible in virtue of (52ii). Next we choose the natural number N_0 such that

$$\frac{1}{s_{N_0}} \sum_{n=1}^{n_0-1} b_n^\alpha \leqq \frac{\varepsilon}{2},$$

that is also possible in virtue of (52i). Then

$$\frac{1}{s_N^\alpha} \sum_{n=1}^{N} b_n^\alpha = \frac{1}{s_N^\alpha} \left\{\sum_{n=1}^{n_0-1} + \sum_{n=n_0}^{N}\right\} \leqq \frac{\varepsilon}{2} + \frac{1}{s_N^\alpha} \sum_{n=n_0}^{N} \frac{\varepsilon}{2} s_N^{\alpha-1} b_n \leqq \varepsilon$$

whenever $N \geqq N_0$, and assertion (53) is proved.

Proof of Theorem 6. [10]) In the proof we apply the following elementary inequality: for every real number u and every natural number n, we have (see [10], p. 365)

(54) $$\left| e^{iu} - \sum_{k=0}^{n-1} \frac{(iu)^k}{k!} \right| \leqq \frac{|u|^n}{n!}.$$

We make use of the classical method of characteristic functions. Let us introduce the following notation:

$$\psi_N(\lambda) = \int_{-\infty}^{+\infty} e^{i\lambda y}\, dF_N(y),$$

where $F_N(y)$ is defined by (17). It is enough to prove that for any fixed λ the characteristic function $\psi_N(\lambda)$ tends to the characteristic function of the normal distribution, i.e.

(55) $$\psi_N(\lambda) \to e^{-\frac{\lambda^2}{2}} \qquad (N \to \infty).$$

It is obvious that

(56) $$\psi_N(\lambda) = \int_0^1 \exp\left\{\frac{i\lambda S_N(t)}{A_N}\right\} dt.$$

Applying (54) with $n=3$, we get that

(57) $$\psi_N(\lambda) = \int_0^1 \prod_{n=1}^{N} \left\{\left(1 + \frac{i\lambda a_n \varphi_n(t)}{A_N} - \frac{\lambda^2 a_n^2 \varphi_n^2(t)}{2A_N^2}\right) + \theta_n \frac{\lambda^3 a_n^3 \varphi_n^3(t)}{6A_N^3}\right\} dt,$$

where θ_n also depends on N, and $|\theta_n| \leqq 1$ $(n = 1, 2, \cdots, N)$.

---

[10]) The proof follows that of LINDEBERG's theorem which is due to FELLER [9]. See also [10], pp. 365—368.

We show that the integral on the right-hand side of (57) can be replaced by the following simpler integral

$$(58)\qquad \int_0^1 \prod_{n=1}^{N}\left(1+\frac{i\lambda a_n\varphi_n(t)}{A_N}-\frac{\lambda^2 a_n^2\varphi_n^2(t)}{2A_N^2}\right)dt,$$

in the sense that for every fixed λ the difference of (57) and (58) tends to 0 if N tends to ∞. For the sake of brevity we denote

$$P_n(t)=1+\frac{i\lambda a_n\varphi_n(t)}{A_N}-\frac{\lambda^2 a_n^2\varphi_n^2(t)}{2A_N^2}\quad\text{and}\quad R_n(t)=\theta_n\frac{\lambda^3 a_n^3\varphi_n^3(t)}{6A_N^3},$$

where we do not indicate the dependence on N. Applying the following identity (see [10], p. 367.)

$$\prod_{n=1}^{N}(p_n+r_n)-\prod_{n=1}^{N}p_n=\sum_{n=1}^{N}r_n\left(\prod_{k=1}^{n-1}p_k\right)\left(\prod_{k=n+1}^{N}(p_k+r_k)\right)$$

(the empty product equals 1), we obtain

$$(59)\qquad \begin{aligned}\left|\exp\left\{\frac{i\lambda S_N(t)}{A_N}\right\}-\prod_{n=1}^{N}P_n(t)\right| &= \left|\prod_{n=1}^{N}(P_n(t)+R_n(t))-\prod_{n=1}^{N}P_n(t)\right|\leqq\\ &\leqq\sum_{n=1}^{N}|R_n(t)|\cdot\left(\prod_{k=1}^{n-1}|P_k(t)|\right)\cdot\left(\prod_{k=n+1}^{N}(|P_k(t)|+|R_k(t)|)\right).\end{aligned}$$

By a simple calculation, we get that

$$(60)\qquad \begin{aligned}|P_n(t)| &= \left\{\left(1-\frac{\lambda^2 a_n^2\varphi_n^2(t)}{2A_N^2}\right)^2+\frac{\lambda^2 a_n^2\varphi_n^2(t)}{A_N^2}\right\}^{\frac{1}{2}}=\\ &=\left\{1+\frac{\lambda^4 a_n^4\varphi_n^4(t)}{4A_N^4}\right\}^{\frac{1}{2}}\leqq 1+\frac{\lambda^2 a_n^2 K^2}{2A_N^2},\end{aligned}$$

furthermore,

$$(61)\qquad |R_n(t)|\leqq\frac{|\lambda|^3|a_n|^3K^3}{6A_N^3}\qquad(n=1,2,\cdots,N),$$

where K denotes a common bound of the system $\{\varphi_n(t)\}$.

From (60) and (61) we obtain that the right-hand side of (59) does not exceed

$$\sum_{n=1}^{N}\frac{|\lambda|^3K^3|a_n|^3}{6A_N^3}\left\{\prod_{k=1}^{n-1}\left(1+\frac{\lambda^2a_k^2K^2}{2A_N^2}\right)\prod_{k=n+1}^{N}\left(1+\frac{\lambda^2a_k^2K^2}{2A_N^2}+\frac{|\lambda|^3K^3|a_k|^3}{6A_N^3}\right)\right\}.$$

Applying the inequality $1+u\leqq e^u$ $(u\geqq 0)$, the last sum is not greater than

$$\sum_{n=1}^{N}\frac{|\lambda|^3K^3|a_n|^3}{6A_N^3}\exp\left\{\sum_{k=1}^{N}\frac{\lambda^2K^2a_k^2}{2A_N^2}+\sum_{k=n+1}^{N}\frac{|\lambda|^3K^3|a_k|^3}{6A_N^3}\right\}\leqq$$

$$\leqq\frac{|\lambda|^3K^3}{6}\cdot\exp\left\{\frac{\lambda^2K^2}{2}+\frac{|\lambda|^3K^3}{6}\right\}\cdot\frac{1}{A_N^3}\sum_{n=1}^{N}|a_n|^3,$$

as it is clear that $A_N^{-3}\sum_{n=1}^{\infty}|a_n|^3\leqq 1$. It follows from (16) that the conditions (52) of Lemma 6 are satisfied by the sequence $\{a_n^2\}$. Therefore, applying Lemma 6 with $\alpha=3/2$, on the basis of (53), we get that the difference of the integrand of (57) and (58) tends to 0 $(N\to\infty)$ uniformly in t $(0\leqq t\leqq 1)$ if λ is fixed.

To prove (55) for any fixed λ, we need the following inequalities:

(62)
$$1-u\leqq e^{-u}\quad\text{if}\quad u\geqq 0,$$
$$e^{-u(1+u)}\leqq 1-u\quad\text{if}\quad 0\leqq u\leqq\frac{\sqrt{2}-1}{2}.\ ^{11)}$$

Now carry out the multiplication in the integrand of (58) and integrate term by term

$$\int_0^1\prod_{n=1}^{N}\left(1+\frac{i\lambda a_n\varphi_n(t)}{A_N}-\frac{\lambda^2a_n^2\varphi_n^2(t)}{2A_N^2}\right)dt=$$

$$=1+\sum{}'\frac{i^k\lambda^k}{A_N^k}a_{n_1}\cdots a_{n_k}\int_0^1\varphi_{n_1}(t)\cdots\varphi_{n_k}(t)\,dt+$$

$$+\sum{}'\frac{(-1)^k\lambda^{2k}}{2^kA_N^{2k}}a_{n_1}^2\cdots a_{n_k}^2\int_0^1\varphi_{n_1}^2(t)\cdots\varphi_{n_k}^2(t)\,dt+$$

$$+\sum{}''\frac{i^k(-1)^l\lambda^{k+2l}}{2^kA_N^{k+2l}}a_{n_1}\cdots a_{n_k}a_{m_1}^2\cdots a_{m_l}^2\int_0^1\varphi_{n_1}(t)\cdots\varphi_{n_k}(t)\varphi_{m_1}^2(t)\cdots\varphi_{m_l}^2(t)\,dt,$$

where the sum Σ' is extended for all systems of integer values $(1\leqq)\ n_1<\cdots<n_k(\leqq N)$ $(1\leqq k\leqq N)$, the sum Σ'' is extended for all systems of integer values $(1\leqq)\ n_1<\cdots<n_k\ (\leqq N)$ and $(1\leqq)\ m_1<\cdots<m_l(\leqq N)$ for which $n_i\neq m_j$ $(1\leqq i\leqq k,\ 1\leqq j\leqq l)$; $1\leqq k$, $1\leqq l$ and $k+l\leqq N$. It follows from (3) that the integral (58) equals

$$1+\sum{}'\frac{(-1)^k\lambda^{2k}}{2^kA_N^{2k}}a_{n_1}^2\cdots a_{n_k}^2=\prod_{n=1}^{N}\left(1-\frac{\lambda^2a_n^2}{2A_N^2}\right).$$

11) This inequality follows from the fact that the curve $v=e^{-u(1+u)}$ $\left((-1-\sqrt{2})/2\leqq u\leqq\right.$ $\left.\leqq(-1+\sqrt{2})/2\right)$, which is concave, lies below its tangent at the point $u=0$, $v=1$.

Taking into account (62), on the one hand

$$(63)\qquad \prod_{n=1}^{N}\left(1-\frac{\lambda^2 a_n^2}{2A_N^2}\right) \leqq \exp\left\{-\sum_{n=1}^{N}\frac{\lambda^2 a_n^2}{2A_N^2}\right\} = e^{-\frac{\lambda^2}{2}}$$

holds for every N, on the other hand

$$(64)\quad \prod_{n=1}^{N}\left(1-\frac{\lambda^2 a_n^2}{2A_N^2}\right) \geqq \exp\left\{-\sum_{n=1}^{N}\frac{\lambda^2 a_n^2}{2A_N^2}\left(1+\frac{\lambda^2 a_n^2}{2A_N^2}\right)\right\} = \exp\left\{-\frac{\lambda^2}{2}-\sum_{n=1}^{N}\frac{\lambda^4 a_n^4}{4A_N^4}\right\}$$

holds if

$$\frac{\lambda^2 a_n^2}{2A_N^2} \leqq \frac{\sqrt{2}-1}{2} \qquad (1 \leqq n \leqq N).$$

But, in virtue of (16ii), this is satisfied for every sufficiently large N. Applying again Lemma 6 with $\alpha=2$, we get

$$\sum_{n=1}^{N}\frac{\lambda^4 a_n^4}{4A_N^4} \to 0 \qquad (N\to\infty).$$

According to (63) and (64)

$$\lim_{N\to\infty}\prod_{n=1}^{N}\left(1-\frac{\lambda^2 a_n^2}{2A_N^2}\right) = e^{-\frac{\lambda^2}{2}}$$

holds for every fixed λ. This completes the proof of Theorem 6.

References

[1] G. Alexits, *Convergence problems of orthogonal series* (Budapest, 1961).

[2] A. Kolmogoroff, Über das Gesetz des iterierten Logarithmus, *Math. Annalen*, **101** (1929), 126—135.

[3] J. Marcinkiewicz and A. Zygmund, Sur les fonctions indépendantes, *Fundamenta Math.*, **29** (1937), 60—90.

[4] A. Zygmund, *Trigonometric series* (Cambridge, 1959).

[5] R. Salem and A. Zygmund, La loi du logarithme itéré pour les séries trigonométriques lacunaires, *Bull. Sci. Math.*, **74** (1950), 209—224.

[6] P. Erdős and I. S. Gál, On the law of the iterated logarithm, *Nederl. Akad. Wet., Proc.*, Ser. A, **58** (1955), 65—84.

[7] P. Révész, Some remarks on strongly multiplicative systems, *Acta Math. Acad. Sci. Hung.*, **16** (1965), 441—446.

[8] R. Salem and R. Zygmund, On lacunary trigonometric series, *Proc. Nat. Acad. Sci. U.S.A.*, **33** (1947), 333—338.

[9] W. Feller, Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung, *Math. Zeitschr.*, **40** (1935), 531—559.

[10] A. Rényi, *Wahrscheinlichkeitsrechnung* (Berlin, 1962).

(Received March 18, 1967)

Inequalities for polynomials and their derivatives

By A. K. VARMA in Edmonton (Alberta, Canada) *)

Introduction

1. Recently J. BALÁZS and P. TURÁN [2] have obtained certain interesting inequalities which arise from their consideration of (0, 2) interpolation on π-abscissas ($\pi_n(x)=(1-x^2)P'_{n-1}(x)$, $P_n(x)$ being the Legendre polynomial of degree $\leqq n$). By (0, 2) interpolation they mean the problem of finding interpolatory polynomials $R_n(x)$ of degree $\leqq 2n-1$ for which

$$R_n(x_k) = \alpha_k, \quad R''_n(x_k) = \beta_k \qquad (k=1, 2, \cdots, n) \tag{1.1.1}$$

are prescribed. From this consideration they proved the following

Theorem 1. 1. 1. *Let n be even and further if we are given for a polynomial $Q_{2n-1}(x)$ of degree $\leqq 2n-1$*

$$|Q_{2n-1}(x_k)| \leqq A, \quad |Q''_{2n-1}(x_k)| \leqq B \qquad (k=1, 2, \cdots, n) \tag{1.1.2}$$

then for $-1 \leqq x \leqq +1$ we have

$$|Q_{2n-1}(x)| \leqq \pi^6 nA + \frac{\pi^5 B}{n} \tag{1.1.3}$$

and

$$|Q'_{2n-1}(x)| \leqq \pi^8 n^{5/2} A + \pi^5 B n^{1/2}. \tag{1.1.4}$$

2. The appearence of the exponent 5/2 in (1. 1. 4) is unusual. They proved that the results (1. 1. 3) and (1. 1. 4) are also best possible in a certain sense. The object of this note is to obtain analogous results when the x_k's are taken to be the zeros of $(1-x^2)T_n(x)$, $T_n(x)$ being the Tchebycheff polynomials of the first kind.

---

*) The author acknowledges financial support for this work from the University of Alberta Post-Doctoral Fellowship (1966—67).

He is presently at the University of Florida, Gainesville.

In an earlier work [3] we proved that (for n even) there exists a unique polynomial $R_n(x)$ of degree $\leqq 2n+1$ for which

$$(1.2.1) \qquad R_n(x_k) = a_k \qquad (k = 1, 2, \cdots, n+2),$$

$$(1.2.2) \qquad R_n''(x_k) = b_k \qquad (k = 2, 3, \cdots, n+1)$$

are prescribed in advance. Let

$$(1.2.3) \qquad 1 = x_1 > x_2 > \cdots > x_{n+2} = -1$$

be the zeros of the polynomial $(1-x^2)T_n(x)$, $T_n(x) = \cos(n \arccos x)$. From our earlier work [3] we have

$$(1.2.4) \qquad R_n(x) = \sum_{k=1}^{n+2} a_k r_k(x) + \sum_{k=2}^{n+1} b_k \varrho_k(x) \quad (n \text{ even}),$$

where fundamental polynomials $r_k(x)$ and $\varrho_k(x)$ are mentioned in the next section. From the uniqueness theorem [3] it follows that if $Q_{2n+1}(x)$ is an arbitrary polynomial of degree $\leqq 2n+1$, then

$$(1.2.5) \qquad Q_{2n+1}(x) = \sum_{k=1}^{n+2} Q_{2n+1}(x_k) r_k(x) + \sum_{k=2}^{n+1} Q_{2n+1}''(x_k) \varrho_k(x).$$

Based on this we shall prove the following main theorem:

Theorem 1.2.1. *Suppose the polynomial $Q_{2n+1}(x)$ of degree $\leqq 2n+1$ (n even) satisfies:*

$$(1.2.6) \qquad |Q_{2n+1}(x_k)| \leqq A \qquad (k = 1, 2, \cdots, n+2),$$

$$(1.2.7) \qquad |Q_{2n+1}''(x_k)| \leqq \frac{B}{1-x_k^2} \qquad (k = 2, 3, \cdots, n+1).$$

Then for $-1 \leqq x \leqq +1$ *we have*

$$(1.2.8) \qquad |Q_{2n+1}(x)| \leqq c_1(n^{3/2}A + Bn^{-1/2}) \quad \text{with} \quad (c_1 = 54)$$

and

$$(1.2.9) \qquad |Q_{2n+1}'(x)| \leqq c_2(n^{5/2}A + Bn^{1/2}) \quad \text{with} \quad c_2 = 251.$$

First we remark that the result (1.2.8) is essentially best possible, i.e. we can find a suitable polynomial $f_0(x)$ of degree $\leqq 2n+1$ which satisfies (1.2.6) and (1.2.7) and for a numerical positive c_3

$$(1.2.10) \qquad |f_0(d_n)| > c_3(An^{3/2} + Bn^{-1/2})$$

where $d_n = \cos\chi_n$, $\chi_n = \frac{\pi}{2} - \frac{\pi}{4n}$. Thus comparing the results on these two abscissas we find that (1.2.8) is not so good as (1.1.3) although (1.2.8) is best possible as explained above. Nevertheless, the estimation of the derivative in both cases are equally good. If we apply MARKOV's inequality on (1. 2. 8) in the closed interval $-1 \leqq x \leqq +1$ we get

$$|Q'_{2n+1}(x)| \leqq c_1(An^{7/2}+Bn^{3/2}). \tag{1.2.11}$$

The result stated in (1. 2. 9) is much better than (1. 2. 11). If, however, we consider only closed subintervals of $(-1, 1)$, S. BERNSTEIN's inequality gives from (1. 2. 8) that

$$|Q'_{2n+1}(x)| \leqq \frac{2c_1}{\sqrt{\varepsilon}}[An^{5/2}+Bn^{1/2}] \quad \text{for} \quad -1+\varepsilon \leqq x \leqq 1-\varepsilon. \tag{1.2.12}$$

Comparing (1. 2. 9) with (1. 2. 12) we observe that both inequalities assert in $-1+\varepsilon \leqq x \leqq 1-\varepsilon$ essentially the same thing.

2. Preliminaries

1. The explicit forms of the fundamental functions $r_k(x)$ and $\varrho_k(x)$ $(k=2, 3, \cdots, n+1)$ that we have obtained in [3] are the following:

$$\varrho_k(x) = \frac{(1-x^2)^{1/4}T_n(x)}{2T'_n(x_k)}\left[A_k\int_{-1}^{x}\frac{T_n(t)}{(1-t^2)^{1/4}}dt + \int_{-1}^{x}\frac{l_k(t)}{(1-t^2)^{1/4}}dt\right] \tag{2.1.1}$$

where

$$A_k\int_{-1}^{+1}\frac{T_n(t)}{(1-t^2)^{1/4}}dt = -\int_{-1}^{+1}\frac{l_k(t)}{(1-t^2)^{1/4}}dt \tag{2.1.2}$$

and $l_k(t)$ is the fundamental polynomial of Lagrange interpolation

$$l_k(t) = \frac{T_n(t)}{(t-x_k)T'_n(x_k)} \qquad (k = 2, 3, \cdots, n+1), \tag{2.1.3}$$

$$\begin{aligned} r_k(x) = {} & \frac{(1-x^2)}{2(1-x_k^2)}l_k^2(x) + \frac{(1-x^2)l_k(x)T'_n(x)}{2(1-x_k^2)T'_n(x_k)} + b_k\varrho_k(x) + \\ & + \frac{(1-x^2)^{1/4}T_n(x)}{4(1-x_k^2)T'_n(x_k)}\left[A'_k\int_{-1}^{x}\frac{T_n(t)}{(1-t^2)^{1/4}}dt + \int_{-1}^{x}\frac{tl'_k(t)}{(1-t^2)^{1/4}}dt\right], \end{aligned} \tag{2.1.4}$$

where

$$b_k = \frac{n^2}{1-x_k^2} + \frac{1}{(1-x_k^2)^2}, \tag{2.1.5}$$

$$A'_k\int_{-1}^{+1}\frac{T_n(t)}{(1-t^2)^{1/4}}dt = -\int_{-1}^{+1}\frac{tl'_k(t)}{(1-t^2)^{1/4}}dt. \tag{2.1.6}$$

For $k=1$ and $k=n+2$ we have

$$(2.1.7)\quad r_1(x) = \frac{1+x}{2} T_n^2(x) + (1-x^2)T_n(x)T_n'(x) - \frac{(1-x^2)^{1/4} T_n(x)}{2} \int_{-1}^{x} \frac{T_n'(t)}{(1-t^2)^{1/4}}\, dt$$

and

$$(2.1.8)\quad r_{n+2}(x) = \frac{1-x}{2} T_n^2(x) - (1-x^2)T_n(x)T_n'(x) - \frac{(1-x^2)^{1/4} T_n(x)}{2} \int_{-1}^{x} \frac{T_n'(t)}{(1-t^2)^{1/4}}\, dt.$$

2. We wish to express these fundamental polynomials in another form, suitable to our purpose. For this we denote

$$(2.2.1)\quad \begin{aligned} P_{2r}(x) &= \frac{-\Gamma\left(r-\frac{1}{4}\right)}{\Gamma\left(r+\frac{5}{4}\right)} \sum_{i=0}^{r-1} \frac{\Gamma\left(i+\frac{5}{4}\right)}{\Gamma\left(i+\frac{3}{4}\right)} T_{2i+1}(x) = \\ &= \text{Polynomial part of } (1-x^2)^{-3/4} \int_{-1}^{x} \frac{T_{2r}(t)}{(1-t^2)^{1/4}} \end{aligned}$$

and

$$(2.2.2)\quad \begin{aligned} V_{2r-1}(x) &= \frac{-\Gamma\left(r-\frac{3}{4}\right)}{\Gamma\left(r+\frac{3}{4}\right)} \left[\frac{1}{2}\frac{\Gamma\left(\frac{3}{4}\right)}{\Gamma\left(\frac{1}{4}\right)} + \sum_{i=1}^{r-1} \frac{\Gamma\left(i+\frac{3}{4}\right)}{\Gamma\left(i+\frac{1}{4}\right)} T_{2i}(x)\right] = \\ &= (1-x^2)^{-3/4} \int_{-1}^{x} \frac{T_{2r}(t)\, dt}{(1-t^2)^{1/4}}. \end{aligned}$$

Thus $\varrho_k(x)$ and $r_k(x)$ can be written, for $k=2, 3, \cdots, n+1$, in the following forms

$$(2.2.3)\quad \varrho_k(x) = \frac{(1-x^2)T_n(x)}{2T_n'(x_k)} q_{n-1}(x)$$

where

$$(2.2.4)\quad q_{n-1}(x) = A_k P_n(x) + \frac{2}{n} \sum_{r=1}^{\frac{n}{2}} [T_{2r}(x_k)P_{2r}(x) + T_{2r-1}(x_k)V_{2r-1}(x)]$$

and A_k is defined by (2. 1. 2), further

$$(2.2.5)\quad r_k(x) = \frac{(1-x^2)l_k^2(x)}{1-x_k^2} + b_k\varrho_k(x) + \frac{(1-x^2)T_n(x)s_{n-1}(x)}{4T_n'(x_k)(1-x_k^2)}$$

where $s_{n-1}(x)$ is given by

(2.2.6) $$s_{n-1}(x) = A_k' P_n(x) - \frac{4}{n} \sum_{r=1}^{\frac{n}{2}} \left(4r^2 T_{2r}(x_k) P_{2r}(x) + (2r-1)^2 T_{2r-1}(x_k) V_{2r-1}(x)\right).$$

Here b_k and A_k' are stated in (2. 1. 5) and (2. 1. 6).

3. We will prove some results which we require in the estimation of the fundamental polynomials.

Lemma 2. 3. 1. *For* $k=2, 3, \cdots, n+1$ *we have*

(2. 3. 1) $$\int_{-1}^{+1} \frac{l_k(t)}{(1-t^2)^{1/4}} dt \geqq 0,$$

(2. 3. 2) $$\sum_{k=2}^{n+1} \frac{1}{\sin \theta_k} \int_{-1}^{+1} \frac{l_k(t)}{(1-t^2)^{1/4}} dt \leqq 12.$$

This lemma is established in our earlier work [see formula (5. 8) and (4. 7) in [4]].

Lemma 2. 3. 2. *For* $k=2, 3, \cdots, n+1$ *we have*

(2. 3. 3) $$|d_k| \equiv \left| \int_{-1}^{+1} \frac{t\, l_k'(t)}{(1-t^2)^{1/4}} dt \right| \leqq \frac{4}{\sqrt{n}} \frac{1}{\sqrt{1-x_k^2}}.$$

Proof. From a result of L. Fejér we have

$$l_k'(t) = \frac{2}{n} \sum_{r=1}^{n-1} T_r(x_k) T_r'(t).$$

Integration by parts and using the differential equation for $T_r(t)$ yields

$$\int_{-1}^{+1} \frac{t T_r'(t)}{(1-t^2)^{1/4}} dt = \frac{\sqrt{\pi}}{2} r^2 \frac{\Gamma\left(\frac{r}{2} - \frac{1}{4}\right)}{\Gamma\left(\frac{r}{2} + \frac{5}{4}\right)} \text{ for } r \text{ even, } = 0 \text{ for } r \text{ odd.}$$

From this the result follows by using Abel's inequality.

Lemma 2. 3. 3. *For* $-1 \leqq x \leqq +1$ *we have*

(2. 3. 4) $$|P_{2r}(x)| \leqq 1, \quad |V_{2r-1}(x)| \leqq 1,$$

(2. 3. 5) $$|(1-x^2)^{1/2} P_{2r}(x)| \leqq \frac{1}{r}, \quad |(1-x^2)^{1/2} V_{2r-1}(x)| \leqq \frac{1}{r} \qquad (r \geqq 1),$$

(2. 3. 6) $$|(1-x^2) P_{2r}'(x)| \leqq 2, \quad |(1-x^2) V_{2r-1}'(x)| \leqq 2,$$

where $P_{2r}(x)$ *and* $V_{2r-1}(x)$ *are defined in* (2. 2. 1) *and* (2. 2. 2), *respectively.*

Proof. We will prove results for $P_{2r}(x)$, the corresponding results for $V_{2r-1}(x)$ are similar. From (2. 2. 1) we have

$$|P_{2r}(x)| \leqq r \frac{\Gamma\left(r-\frac{1}{4}\right)}{\Gamma\left(r+\frac{5}{4}\right)} \frac{\Gamma\left(r+\frac{1}{4}\right)}{\Gamma\left(r-\frac{1}{4}\right)} \leqq 1.$$

In order to prove (2. 3. 5) we observe that $\dfrac{\Gamma\left(i+\frac{5}{4}\right)}{\Gamma\left(i+\frac{3}{4}\right)}$ is a monotonically increasing function of i. Using ABEL's inequality we get

$$|(1-x^2)^{1/2} P_{2r}(x)| \leqq \frac{\Gamma\left(r-\frac{1}{4}\right)}{\Gamma\left(r+\frac{5}{4}\right)} \frac{\Gamma\left(r+\frac{1}{4}\right)}{\Gamma\left(r-\frac{1}{4}\right)} \max_{1 \leqq p \leqq r-1} \left| \sum_{i=1}^{p} \cos(2i+1)\theta \sin\theta \right| \leqq \frac{1}{r}.$$

Again, using ABEL's inequality, we have

$$|(1-x^2) P'_{2r}(x)| \leqq \left| \frac{(2r-1)}{r+\frac{1}{4}} \max_{1 \leqq p \leqq r-1} \sum_{i=1}^{p} \sin(2i+1)\theta \sin\theta \right| \leqq 2.$$

This completes the proof of the above Lemma by using the representation of $q_{n-1}(x)$ as given in (2. 2. 4). From the above lemma we get

Lemma 2. 3. 4. *For* $-1 \leqq x \leqq +1$ *we have*

(2. 3. 7) $$|q_{n-1}(x)| \leqq 4 + 2C_k n^{3/2},$$

(2. 3. 8) $$|(1-x^2)^{1/2} q_{n-1}(x)| \leqq \frac{4}{n} \log n + 2n^{1/2} C_k,$$

(2. 3. 9) $$|(1-x^2) q'_{n-1}(x)| \leqq 8 + 4n^{3/2} C_k.$$

Here C_k *is given by*

(2. 3. 10) $$C_k = \int_{-1}^{+1} \frac{l_k(t)}{(1-t^2)^{1/4}} dt.$$

Let us denote

$$t_1(x) = (1-x^2) T_n(x) q_{n-1}(x).$$

Then from Lemma 2. 3. 4 we have at once

(2. 3. 11) $$|t_1(x)| \leqq \frac{4}{n}\log n + 2n^{1/2} C_k \qquad (-1 \leqq x \leqq +1),$$

(2. 3. 12) $$|t_1'(x)| \leqq 20\log n + 10n^{3/2} C_k \qquad (-1 \leqq x \leqq +1).$$

3. Estimates of the fundamental polynomials

1. The above Lemmas lead us to formulate:

Lemma 3. 1. 1. *For* $-1 \leqq x \leqq +1$ *we have*

(3. 1. 1) $$\sum_{k=2}^{n+1} |\varrho_k(x)| \leqq \sum_{k=2}^{n+1} \frac{|\varrho_k(x)|}{(1-x_k^2)} \leqq \frac{14}{n^{1/2}},$$

(3. 1. 2) $$\sum_{k=2}^{n+1} |\varrho_k'(x)| \leqq \sum_{k=2}^{n+1} \frac{|\varrho_k'(x)|}{1-x_k^2} \leqq 80n^{1/2}.$$

From (2. 2. 3), (2. 3. 10), (2. 3. 1), (2. 3. 2) and (2. 3. 11) we have

$$\sum_{k=2}^{n+1} \frac{|\varrho_k(x)|}{1-x_k^2} \leqq \frac{1}{2n}\left[\frac{4}{n}\log n \sum_{k=2}^{n+1} \frac{1}{\sqrt{1-x_k^2}} + 2\cdot n^{1/2}\cdot 12\right],$$

Using again the above relations and (2. 3. 12) we have

$$\sum_{k=2}^{n+1} \frac{|\varrho_k'(x)|}{1-x_k^2} \leqq \frac{1}{2n}\left[20\log n \sum_{k=2}^{n+1} \frac{1}{\sqrt{1-x_k^2}} + 10n^{3/2}\cdot 12\right] \leqq 80n^{1/2}.$$

2. In order to determine the estimate of the fundamental polynomials of the first kind we need the following Lemmas:

Lemma 3. 2. 1. *For* $-1 \leqq x \leqq +1$ *we have*

(3. 2. 1) $$\sum_{k=2}^{n+1} \frac{(1-x^2)l_k^2(x)}{1-x_k^2} \leqq 8,$$

(3. 2. 2) $$\sum_{k=2}^{n+1} \left[\frac{(1-x^2)l_k^2(x)}{1-x_k^2}\right]' \leqq 36n^2,$$

where dash denotes differentiation with respect to x.

A proof of (3. 2. 1) is given in our earlier work [4], and (3. 2. 2) follows very easily by using the inequalities:

(3. 2. 3) $$|l_k(x)| \leqq 2 \qquad (-1 \leqq x \leqq +1),$$

(3. 2. 4) $$|(1-x^2)^{1/2} l_k'(x)| \leqq 2n \qquad (-1 \leqq x \leqq +1).$$

Lemma 3. 2. 2. *For* $-1 \leqq x+1$ *we have*

(3. 2. 5) $$|s_{n-1}(x)| \leqq 3n^2+2n^{3/2}|d_k|,$$

(3. 2. 6) $$|(1-x^2)^{1/2}s_{n-1}(x)| \leqq 16n+2n^{1/2}|d_k|,$$

(3. 2. 7) $$|(1-x^2)s'_{n-1}(x)| \leqq 1n^2+4n^{3/2}|d_k|,$$

where d_k *is defined by* (2. 3. 3), *and* $s_{n-1}(x)$ *is a polynomial in* x *of degree* $\leqq n-1$ *given by* (2. 2. 6).

The proof of this lemma is clear from Lemma 2. 3. 3 and (2. 1. 6), and so we omit the details. Let us denote:

(3. 2. 8) $$t_2(x) = (1-x^2)T_n(x)s_{n-1}(x);$$

then by the above Lemma 3. 2. 2 it follows that

(3. 2. 9) $$|t_2(x)| \leqq 16n+2n^{1/2}|d_k| \qquad (-1 \leqq x \leqq +1),$$

(3. 2. 10) $$|t'_2(x)| \leqq 33n^2+10n^{3/2}|d_k| \qquad (-1 \leqq x \leqq +1).$$

3. Next we state:

Lemma 3. 3. 1. *For* $-1 \leqq x \leqq +1$ *we have*

(3. 3. 1) $$|r_1(x)| \leqq 3n, \quad |r_{n+2}(x)| \leqq 3n,$$

(3. 3. 2) $$|r'_1(x)| \leqq 13n^2, \quad |r'_{n+2}(x)| \leqq 13n^2.$$

A proof of (3. 3. 1) is given in our earlier work [formula 6. 10, [4]] and (3. 3. 2) can be obtained easily by a simple computation using similar ideas as in Lemma 2. 3. 3.

Lemma 3. 3. 2. *For* $-1 \leqq x \leqq +1$, *we have*

(3. 3. 3) $$\sum_{k=1}^{n+2} |r_k(x)| \leqq C_5 n^{3/2} \quad \textit{with} \quad C_5 = 54$$

and

(3. 3. 4) $$\sum_{k=1}^{n+2} |r'_k(x)| \leqq C_6 n^{5/2} \quad \textit{with} \quad C_6 = 251.$$

Proof. Using the representation of $r_k(x)$ as given in (2. 2. 5) we have

$$\sum_{k=2}^{n+1} |r_k(x)| \leqq 8+2n^2 \sum_{k=2}^{n+1} \frac{|\varrho_k(x)|}{(1-x_k^2)} + \sum_{k=2}^{n+1} \frac{2n^{1/2}|d_k|+16n}{4n\sqrt{1-x_k^2}} \leqq$$

$$\leqq 8+2n^2\frac{14}{n^{1/2}}+4n\log n+4n\sum_{k=2}^{n+1}\frac{1}{k^2} \leqq 48n^{3/2}.$$

Here we have used (2. 3. 3), (3. 2. 1), (3. 1. 1). Combining with (3. 3. 1) we obtain (3. 3. 3) with $C_5 = 48$. Again,

$$\sum_{k=2}^{n+1}{}' |r_k'(x)| \leqq 36n^2 + 2n^2 \sum_{k=2}^{n+1}{}' \frac{|\varrho_k'(x)|}{(1-x_k^2)} + \sum_{k=2}^{n+1}{}' \frac{33n^2 + 10n^{3/2}|d_k|}{4n\sqrt{1-x_k^2}} \leqq$$

$$\leqq 36n^2 + 2n^2 \cdot 80n^{1/2} + 9n^2 \log n + 20n^2 \leqq 225n^{5/2}.$$

Here we have used (2. 3. 3), (3. 2. 2) and (3. 1. 2). Combining with (3. 3. 2) we obtain 3. 3. 4.

Proof of Theorem 1. 2. 1. From the representation of $Q_{2n+1}(x)$ as given in (1. 2. 5) we have on using (1. 2. 6), (1. 2. 7):

$$|Q_{2n+1}(x)| \leqq A \sum_{k=1}^{n+2} |r_k(x)| + B \sum_{k=2}^{n+1}{}' \frac{|\varrho_k(x)|}{1-x_k^2} \leqq C_1(An^{3/2} + Bn^{1/2}).$$

Here we have used only Lemma 3. 1. 1 and Lemma 3. 3. 2. Similarly using the same Lemmas

$$|Q_{2n+1}'(x)| \leqq A \sum_{k=1}^{n+2} |r_k'(x)| + B \sum_{k=2}^{n+1}{}' \frac{|\varrho_k'(x)|}{1-x_k^2} \leqq C_2(An^{5/2} + Bn^{1/2}).$$

Now it remains to prove (1. 2. 10). From our earlier work [5. 3, 6. 9 [4]] we know

(3. 3. 5) $$\sum_{k=2}^{n+1}{}' \frac{|\varrho_k(d_n)|}{1-x_k^2} \geqq n^{-1/2},$$

(3. 3. 6) $$\sum_{k=1}^{n+2} |r_k(d_n)| \geqq 2^{-10} n^{3/2},$$

where $d_n = \cos \chi_n$, $\chi_n = \frac{\pi}{2} - \frac{\pi}{4n}$. The polynomial $f_0(x)$ stated in (1. 2. 10) has the following representation:

$$f_0(x) = \sum_{k=1}^{n+2} A \operatorname{sign} r_k(d_n) r_k(x) + \sum_{k=2}^{n+1} B\varrho_k(x) \cdot (1-x_k^2)^{-1} \operatorname{sign} \varrho_k(d_n).$$

Obviously,

$$f_0(x_k) = A \operatorname{sign} r_k(d_n), \quad f_0''(x_k) = B(1-x_k^2)^{-1} \cdot \operatorname{sign} \varrho_k(d_n).$$

Therefore

$$f_0(d_n) = A \sum_{k=1}^{n+2} |r_k(d_n)| + B \sum_{k=2}^{n+1}{}' \frac{|\varrho_k(d_n)|}{1-x_k^2} \geqq C_3(An^{3/2} + Bn^{-1/2})$$

from (3. 3. 5) and (3. 3. 6). This completes the proof of the theorem.

Note. It is rather easy to prove that

$$\sum_{k=0}^{n+2} |r_k'(0)| \geqq C_5 n^{5/2} \quad \text{and} \quad \sum_{k=2}^{n+1}{}' \frac{|\varrho_k'(0)|}{1-x_k^2} \geqq C_6 n^{1/2},$$

from which it follows that (1. 2. 9) is also best possible, i.e. we can find a polynomial $f_1(x)$ of degree $\leqq 2n+1$ which satisfies (1. 2. 6) and (1. 2. 7) and for a numerical positive C_7

$$f_1'(0) \geqq C_7(An^{5/2}+Bn^{1/2}).$$

The author is thankful to Professors P. TURÁN, A. SHARMA, and A. MEIR for some valuable suggestions.

References

[1] J. BALÁZS and P. TURÁN, Notes on interpolation. III, *Acta. Math. Acad. Sci. Hung.*, **9** 1958), 195—214.

[2] J. BALÁZS and P. TURÁN, Notes on interpolation. IV, *ibid.*, **9** (1958), 243—258.

[3] A. K. VARMA and A. SHARMA, Some interpolatory properties of Tchebycheff polynomials (0, 2) case modified, *Publicationes Math. Debrecen*, (1961), 336—349.

[4] A. K. VARMA, On a problem of P. Turán on lacunary interpolation, *Canadian Math. Bull.* **10** (1967), 531—557.

(Received March 22, 1967)

On summability of Fourier series

By LÁSZLÓ LEINDLER in Szeged*)

Introduction

Let Σc_n be a given infinite series and let s_n denote its n-th partial sum. Let $\lambda=\{\lambda_n\}$ be a monotone non-decreasing sequence of integers such that $\lambda_1=1$ and $\lambda_{n+1}-\lambda_n\leqq 1$.

The mean

$$V_n(\lambda)=\frac{1}{\lambda_n}\sum_{\nu=n-\lambda_n}^{n-1} s_\nu \qquad (n\geqq 1)$$

defines the n-th generalized de la Vallée Poussin mean of the sequence $\{s_n\}$ generated by the sequence $\{\lambda_n\}$. The series Σc_n is said to be (V, λ)-summable if $V_n(\lambda)$ converges, and absolutely (V, λ)-summable or, in brief, $|V, \lambda|$-summable if the series

$$\sum_{n=1}^{\infty}|V_{n+1}(\lambda)-V_n(\lambda)|$$

converges.

In the previous papers [7] and [8] we have dealt with (V, λ)-summability of general orthogonal series and $|V, \lambda|$-summability of Fourier series, multiplied by a factor sequence.

The main purpose of the present paper is to unite, in terms of (V, λ)-summation, some classical theorems on the partial sums, the $(\mathscr{C}, 1)$-means, and the proper de la Vallée Poussin means, of Fourier series. Indeed, it is easy to see that, by suitable choice of $\lambda=\{\lambda_n\}$, the $V_n(\lambda)$ means include the partial sums $(\lambda_n\equiv 1, V_{n+1}(\lambda)=s_n)$, the $(\mathscr{C}, 1)$-means $(\lambda_n=n, V_{n+1}(\lambda)=\sigma_n)$ and the proper de la Vallée Poussin means $\left(\lambda_n=\left[\frac{n}{2}\right],^{1)} V_{2n}(\lambda)=V_n\right)$, as special cases.

---

*) This research was supported by the National Research Council of Canada while the author stood at the University of Toronto.

1) $[y]$ denotes the integral part of y.

Analogous problems will be investigated for the so-called "strong (V, λ)-summability" (Theorems 5—7).

Furthermore, the method introduced in the present paper shall be used also to give a very simple proof of a theorem on strong $(\mathscr{C}, \alpha, k)$-summability of Fourier series which generalizes ZYGMUND's theorem [13] concerning strong (H, k)-summability (Theorem 8).

Finally we prove some theorems concerning absolute (V, λ)-summability.

Let $f(x)$ be a function integrable in the sense of Lebesgue and periodic with period 2π, and let

$$\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx) \tag{1}$$

be its Fourier series. $s_n(x)=s_n(f; x)$ and $V_n(\lambda; x)=V_n(f, \lambda; x)$ will denote the n-th partial sum of (1) and the n-th generalized de la Vallée Poussin mean of (1), respectively.

We prove the following theorems:

Theorem 1. *If the function $f(x)$ is bounded, $|f(x)|\leqq M$, then the means $V_n(\lambda; x)$ satisfy the inequality*

$$|V_n(\lambda; x)| \leqq M\left(3+\log\frac{2n-\lambda_n}{\lambda_n}\right). \tag{2}$$

If $\lambda_n=1$ or $\lambda_n=n$, this theorem reduces to classical results of LEBESGUE and FEJÉR, respectively. [2])

We will write as usual $\varphi_x(t) = f(x+2t)+f(x-2t)-2f(x)$ and

$$\Phi_x(h) = \int_0^h |\varphi_x(t)|\, dt.$$

Theorem 2. *If the sequence $\{\lambda_n\}$ tends to infinity and the conditions*

$$\int_{\frac{1}{n}}^{\frac{1}{\lambda_n}} \frac{|\varphi_x(t)|}{t}\, dt = o(1) \qquad (n\to\infty), \tag{3}$$

$$n\Phi_x\left(\frac{1}{n}\right) = o(1) \qquad (n\to\infty) \tag{4}$$

are fulfilled, then the means $V_n(\lambda; x)$ converge to $f(x)$.

[2]) In this and in similar statements it will be understood that the constant factors occurring in the new and the known estimates should not be necessarily the same.

If $n=O(\lambda_n)$, then (3) and (4) are fulfilled at any Lebesgue point of $f(x)$ and consequently almost everywhere, hence this theorem includes the classical Fejér—Lebesgue theorem and the following

Corollary 1. *If the sequence $\{n_n/\lambda\}$ is bounded, then for any integrable $f(x)$ the means $V_n(\lambda; x)$ converge almost everywhere to $f(x)$.*

Let $E_n=E_n(f)$ denote the best approximation to $f(x)$ in the space $\mathscr{C}(0, 2\pi)$ of continuous functions in $(0, 2\pi)$ by a trigonometric polynomial of order not higher than n.

Theorem 3. *If $f(x)$ is continuous, then the estimate*

$$|V_n(\lambda; x)-f(x)| \leqq \left(4+\log\frac{2n-\lambda_n}{\lambda_n}\right) E_{n-\lambda_n}$$

holds true for all x uniformly.

If $\lambda_n\equiv 1$, this theorem gives the well-known result of LEBESGUE [6]. If $\lambda_n=\left[\frac{n}{2}\right]$, then we get the classical theorem of DE LA VALLÉE POUSSIN [12].

Corollary 2. *If there exist two positive numbers, K_1 and K_2, such that $1<K_1\leqq \leqq\frac{n}{\lambda_n}\leqq K_2$, then for any function $f(x)\in \text{Lip}\,\alpha$ $(0<\alpha\leqq 1)$ the estimate*

$$|V_n(\lambda; x)-f(x)| = O(1/n^\alpha)$$

is valid for all x uniformly.

For a sequence $\{\lambda_n\}$ of general type we have

Theorem 4. *If $f(x)\in\text{Lip}\,\alpha$, then*

$$|V_n(\lambda; x)-f(x)| = \begin{cases} O\left(\frac{1}{\lambda_n^\alpha}\right) & \text{for} \quad \alpha<1, \\ O\left(\frac{1+\log\lambda_n}{\lambda_n}\right) & \text{for} \quad \alpha=1, \end{cases} \tag{5}$$

for all x uniformly.

This theorem covers BERNSTEIN's results [3] concerning $(\mathscr{C}, 1)$-means.
From Theorem 3 and Lemma 1 we immediately obtain

Corollary 3. *If $n=O(\lambda_n)$, then, for any function $f(x)\in\mathscr{C}(0, 2\pi)$, the means $V_n(\lambda; x)$ converge uniformly.*

We remark that Corollary 2 and 3 can also be deduced from a theorem of EFIMOV [4] (see p. 770).

The strong (V, λ)-summability, i.e., the means of the form

$$T_n(f, \lambda, k; x) \equiv \left\{\frac{1}{\lambda_n} \sum_{\nu=n-\lambda_n}^{n-1} |s_\nu(x) - f(x)|^k\right\}^{\frac{1}{k}}$$

can also be investigated. We can however prove the strong analogues of the previous theorems only under the restriction $n = O(\lambda_n)$.

Theorem 5. *Suppose that $n = O(\lambda_n)$. Then, for any $f(x) \in \mathscr{C}(0, 2\pi)$ and $k > 0$,*

$$\left\{\frac{1}{\lambda_n} \sum_{\nu=n-\lambda_n}^{n-1} |s_\nu(x) - f(x)|^k\right\}^{\frac{1}{k}} = O(E_{n-\lambda_n}) \tag{6}$$

holds; and if $|f(x)| \leqq M$, then we have

$$\left\{\frac{1}{\lambda_n} \sum_{\nu=n-\lambda_n}^{n-1} |s_\nu(x)|^k\right\}^{\frac{1}{k}} = O(M). \tag{7}$$

See also Alexits—Králik [1], Satz 1, and [9], Satz 1.

Theorem 6. *Suppose that $f(x)$ can be differentiated r times and $f^{(r)}(x) \in \in \operatorname{Lip} \alpha$ $(0 < \alpha \leqq 1)$, and that $n = O(\lambda_n)$. Then for any $k > 0$*

$$T_n(f, \lambda, k; x) = \begin{cases} O\left(\dfrac{1}{n^{r+\alpha}}\right) & \text{for } (r+\alpha)k < 1, \\ O\left(\dfrac{1}{n^{r+\alpha}}\left(1 + \log \dfrac{n}{n - \lambda_n + 1}\right)^{\frac{1}{k}}\right) & \text{for } (r+\alpha)k = 1, \end{cases} \tag{8}$$

uniformly. The same estimate is also valid for the conjugate function $\tilde{f}(x)$.

Furthermore, if $(r+\alpha)k = 1$ $(0 < \alpha \leqq 1)$, then there exist functions $f_1(x)$ and $f_2(x)$ such that their r-th derivatives exist and belong to Lip α, *moreover*

$$\overline{\lim_{n\to\infty}}\, T_n(f_1, \lambda, k; 0) \leqq \frac{c}{n^{r+\alpha}}\left(1 + \log \frac{n}{n - \lambda_n + 1}\right)^{\frac{1}{k}} \tag{9}$$

and

$$\overline{\lim_{n\to\infty}}\, T_n(\tilde{f}_2, \lambda, k; 0) \geqq \frac{c}{n^{r+\alpha}}\left(1 + \log \frac{n}{n - \lambda_n + 1}\right)^{\frac{1}{k}}, \tag{10}$$

where $c > 0$ is independent of n.

MARCINKIEWICZ [11] ($k=2$) and ZYGMUND [13] ($k>0$), proved for any integrable function $f(x)$ that

(11) $$\frac{1}{n}\sum_{\nu=0}^{n}|s_\nu(x)-f(x)|^k = o(1) \quad \text{a.e.}$$

From this result and Lemma 3 we obtain the following theorems.

Theorem 7. *If $f(x)$ is integrable, then, for any positive k and δ,*

(12) $$\frac{1}{\lambda_n}\sum_{\nu=n-\lambda_n}^{n}|s_\nu(x)-f(x)|^k = o\left(\left(\frac{n}{\lambda_n}\right)^{\delta}\right) \quad \text{a.e.}$$

Theorem 8. *If $f(x)$ is an integrable function and k is any positive value, then for any positive α*

(13) $$\frac{1}{A_n^{(\alpha)}}\sum_{\nu+0}^{n}A_{n-\nu}^{(\alpha-1)}|s_\nu(x)-f(x)|^k = o(1) \quad \text{a.e.},$$

with $A_k^{(\alpha)} = \binom{k+\alpha}{k}$.

It is obvious that similar theorems can be proved on the conjugate function too.

Finally we prove two theorems concerning absolute (V,λ)-summability of Fourier series.

Let $E_n^{(2)} = E_n^{(2)}(f)$ denote the best approximation of $f(x)$ in $L^2(0, 2\pi)$.

Theorem 9. *If*

(14) $$\sum_{n=1}^{\infty}\frac{1}{\sqrt{n\lambda_n}}E_n^{(2)}(f) < \infty,$$

then the Fourier series of $f(x)$ is $|V,\lambda|$-summable a.e.

Theorem 10. *Let $\lambda(x)$ ($x\geqq 1$) be a continuous function, linear between n and $n+1$, furthermore let $\lambda(n)=\sqrt{n\lambda_n}$. If*

(15) $$\int_0^1 \frac{1}{t^2\lambda\left(\frac{1}{t}\right)}\left(\int_0^{2\pi}[f(x+t)-f(x-t)]^2\,dx\right)^{\frac{1}{2}}dt < \infty,$$

then the Fourier series of $f(x)$ is $|V,\lambda|$-summable a.e.

It may be worth while to remark that, under the conditions (14) and (15), the same conclusions are valid for the functions: $|f(x)|$, $\tilde{f}(x)$ and $1/f(x)$ if $|f(x)|\geqq c>0$.

§ 1. Lemmas

Lemma 1. *If* $g(t)\in L\left(0,\frac{\pi}{2}\right)$ *and* $|g(t)|\leqq M$ *in* $(0,\delta)$, $\delta\leqq\frac{\pi}{2}$, *then*

$$(1.1)\qquad \begin{aligned} I(g,\lambda_n) &\equiv \int_0^{\frac{\pi}{2}} g(t)\frac{|\sin\lambda_n t\sin(2n-\lambda_n)t|}{\lambda_n\sin^2 t}\,dt \leqq \\ &\leqq \frac{\pi}{2}M\left(3+\log\frac{2n-\lambda_n}{\lambda_n}\right)+\frac{1}{\lambda_n}J(g,\delta), \end{aligned}$$

where $J(g,\delta)$ *is independent of* λ_n *and* $J\left(g,\frac{\pi}{2}\right)=0$.

Proof. *Let* $\alpha_1=\alpha_1(n)=\min\left(\frac{2}{\pi(2n-\lambda_n)},\,\delta\right)$ and $\alpha_2=\alpha_2(n)=\min\left(\frac{2}{\pi\lambda_n},\,\delta\right)$. It is evident that

$$I(g,\lambda_n)=\int_0^{\alpha_1}+\int_{\alpha_1}^{\alpha_2}+\int_{\alpha_2}^{\delta}+\int_{\delta}^{\frac{\pi}{2}}\equiv I_1+I_2+I_3+I_4.$$

Each term I_i can easily be estimated by the use of the inequalities $|\sin nt|\leqq n|\sin t|$ and $|\sin t|\geqq\frac{2}{\pi}t$:

$$I_1\leqq\frac{1}{\lambda_n}M\lambda_n(2n-\lambda_n)\int_0^{\alpha_1}dt\leqq\frac{2}{\pi}M,$$

$$I_2\leqq\frac{1}{\lambda_n}\frac{\pi}{2}M\lambda_n\int_{\alpha_1}^{\alpha_2}\frac{1}{t}\,dt\leqq\frac{\pi}{2}M\log\frac{2n-\lambda_n}{\lambda_n},$$

$$I_3\leqq\frac{1}{\lambda_n}\frac{\pi^2}{4}M\int_{\alpha_2}^{\delta}\frac{1}{t^2}\,dt\leqq\frac{\pi^3}{8}M,$$

$$I_4\leqq\frac{1}{\lambda_n}\frac{\pi^2}{4}\int_{\delta}^{\frac{\pi}{2}}\frac{g(t)}{t^2}\,dt\equiv\frac{1}{\lambda_n}J(g,\delta).$$

Summing up, we obtain (1. 1).

Lemma 2. *If* $g(t)\in L(0,2\pi)$ *and* $|g(t)|\leqq M$ *for all* t, *then, for any* $q>0$, *we have*

$$(1.2)\qquad \left\{\frac{1}{m}\sum_{k=1}^{m}|s_k(g;x)|^q\right\}^{\frac{1}{q}}\leqq C_qM.$$

Proof. We can assume with no loss in generality that $q \geqq 2$. Indeed, if (1. 2) holds for a certain q, then it remains valid for any $q^* > 0$ not greater than q.

Using DIRICHLET's formula for the partial sums we obtain

$$\sum_{k=1}^{m} |s_k|^q \leqq C_1(q) \sum_{k=0}^{m} \Big\{ \Big(\int_{-\frac{1}{m}}^{\frac{1}{m}} |g(x+t)|\, |D_k(t)|\, dt \Big)^q + $$

$$+ \Big| \int_I g(x+t) D_k(t)\, dt \Big|^q \Big\} \equiv \sum\nolimits_1 + \sum\nolimits_2,$$

where $D_k(t)$ is DIRICHLET's kernel and I denotes the set $[-\pi, \pi] \setminus \left(-\frac{1}{m}, \frac{1}{m}\right)$.

It is obvious that

$$\Sigma_1 \leqq C_2(q) M^q m$$

and

$$\sum\nolimits_2 \leqq C_3(q) \sum_{k=0}^{m} \Big\{ \Big| \int_I g(x+t) \operatorname{cotg} \frac{t}{2} \sin kt\, dt \Big|^q + \Big| \int_I g(x+t) \cos kt\, dt \Big|^q \Big\} \equiv$$

$$\equiv \sum\nolimits_3 + \sum\nolimits_4.$$

By using the Hausdorff—Young inequality and the notation $r = \dfrac{q}{q-1}$ $(q \geqq 2)$, we get

$$\sum\nolimits_3 \leqq C_4(q) \Big(\int_I \frac{|g(x+t)|^r}{|t|^r}\, dt \Big)^{\frac{q}{r}} \leqq C_5(q) M^q m$$

and

$$\sum\nolimits_4 \leqq C_6(q) \Big(\int_I |g(x+t)|^r\, dt \Big)^{\frac{q}{r}} \leqq C_7(q) M^q.$$

Collecting our estimates, we obtain (1. 2).

Lemma 3. *Let $\{c_n\}$ be a given sequence. If for any positive β*

$$\frac{1}{n} \sum_{\nu=0}^{n} |c_\nu|^\beta = o(1), \tag{1.3}$$

then, for any triangular matrix $\|\alpha_{n\nu}\|$ and for any positive γ and $p > 1$, we have the estimate

$$\sum_{\nu=0}^{n} \alpha_{n\nu} |c_\nu|^\gamma = o\left(n^{1-\frac{1}{p}} \Big\{ \sum_{\nu=0}^{n} |\alpha_{n\nu}|^p \Big\}^{\frac{1}{p}} \right). \tag{1.4}$$

Proof. By HÖLDER's inequality, we get

$$\sum_{\nu=0}^{n} \alpha_{n\nu}|c_\nu|^\gamma \leqq \left\{\sum_{\nu=0}^{n} |c_\nu|^{\frac{\gamma p}{p-1}}\right\}^{\frac{p-1}{p}} \left\{\sum_{\nu=0}^{n} |\alpha_{n\nu}|^p\right\}^{\frac{1}{p}} = o\left(n^{1-\frac{1}{p}}\right)\left\{\sum_{\nu=0}^{n} |\alpha_{n\nu}|^p\right\}^{\frac{1}{p}},$$

which is the required statement.

§ 2. Proof of the theorems

Ad Theorem 1. A standard computation gives that

$$V_n(\lambda; x) = \frac{1}{\lambda_n \pi}\int_0^{\frac{\pi}{2}} [f(x+2t)+f(x-2t)]\frac{\sin \lambda_n t \sin(2n-\lambda_n)t}{\sin^2 t}\,dt. \tag{2. 1}$$

Applying Lemma 1 with $g(t) = f(x+2t)-f(x-2t)$ and $\delta = \dfrac{2}{\pi}$, we get the inequality (2).

Ad Theorem 2. Since the (V, λ)-means of a constant function $f(x) \equiv c$ equal to c, from (2. 1) it follows that

$$V_n(\lambda; x)-f(x) = \frac{1}{\lambda_n \pi}\int_0^{\frac{\pi}{2}} \varphi_x(t)\frac{\sin \lambda_n t \sin(2n-\lambda_n)t}{\sin^2 t}\,dt. \tag{2. 2}$$

Divide the integral into three parts:

$$\int_0^{\frac{\pi}{2}} = \int_0^{\frac{1}{n}} + \int_{\frac{1}{n}}^{\frac{1}{\lambda_n}} + \int_{\frac{1}{\lambda_n}}^{\frac{\pi}{2}} \equiv I_1+I_2+I_3.$$

These integrals can easily be estimated by standard methods:

$$I_1 \leqq O(1)\int_0^{\frac{1}{n}} |\varphi_x(t)|\lambda_n n\,dt = O\left(\lambda_n n \Phi_x\left(\frac{1}{n}\right)\right),$$

$$I_2 \leqq O(1)\int_{\frac{1}{n}}^{\frac{1}{\lambda_n}} |\varphi_x(t)|\frac{\lambda_n t}{t^2}\,dt = O\left(\lambda_n \int_{\frac{1}{n}}^{\frac{1}{\lambda_n}} \frac{|\varphi_x(t)|}{t}\,dt\right),$$

$$I_3 \leqq O(1)\int_{\frac{1}{\lambda_n}}^{\frac{\pi}{2}} \frac{|\varphi_x(t)|}{t^2}\,dt \leqq O(1)[\Phi_x(t)t^{-2}]_{\frac{1}{\lambda_n}}^{\frac{\pi}{2}} + O(1)\int_{\frac{1}{\lambda_n}}^{\frac{\pi}{2}} \frac{\Phi_x(t)}{t^3}\,dt \leqq$$

$$\leqq O(1)+O\left(\lambda_n^2 \Phi_x\left(\frac{1}{\lambda_n}\right)\right)+o(\lambda_n)+O(1) \leqq O\left(1+\lambda_n^2 \Phi_x\left(\frac{1}{\lambda_n}\right)\right)+o(\lambda_n).$$

From these estimates we get that

$$V_n(\lambda; x)-f(x) = O\left(n\Phi_x\left(\frac{1}{n}\right)+\int_{\frac{1}{n}}^{\frac{1}{\lambda_n}} \frac{|\varphi_x(t)|}{t}\,dt+\lambda_n\Phi_x\left(\frac{1}{\lambda_n}\right)\right)+o(1),$$

which gives the conclusion of Theorem 2 by (3) and (4).

Ad Theorem 3. Let $T_m^*(x)$ denote the trigonometric polynomial of best approximation of order not higher than m. From the definitions of $s_n(f; x)$ and $V_n(f, \lambda; x)$ it is clear that if $n-\lambda_n \geqq m$, then

$$V_n(f-T_m^*, \lambda; x) = V_n(f, \lambda; x)-T_m^*(x).$$

Hence we have

$$\begin{aligned}|V_n(f, \lambda; x)-f(x)| &\leqq |V_n(f, \lambda; x)-T_{n-\lambda_n}^*(x)|+|T_{n-\lambda_n}^*(x)-f(x)| \leqq \\ &\leqq |V_n(f-T_{n-\lambda_n}^*, \lambda; x)|+E_{n-\lambda_n}.\end{aligned} \tag{2. 3}$$

According to (2. 1) we get

$$|V_n(f-T_{n-\lambda_n}^*, \lambda; x)| \leqq \frac{1}{\lambda_n\pi}\int_0^{\frac{\pi}{2}} 2E_{n-\lambda_n}\frac{|\sin\lambda_n t\sin(2n-\lambda_n)t|}{\sin^2 t}\,dt.$$

Using Lemma 1 with $g(t)\equiv 2E_{n-\lambda_n}$ and $\delta=\frac{\pi}{2}$, we obtain

$$|V_n(f-T_{n-\lambda_n}^*, \lambda; x)| \leqq E_{n-\lambda_n}\left(3+\log\frac{2n-\lambda_n}{\lambda_n}\right). \tag{2. 4}$$

The statement of Theorem 3 follows from (2. 3) and (2. 4).

Ad Theorem 4. *If* $f(x)\in\mathrm{Lip}\,\alpha$, we obtain, using (2. 2),

$$|V_n(\lambda; x)-f(x)| \leqq O(1)\frac{1}{\lambda_n}\int_0^{\frac{\pi}{2}} t^\alpha\frac{|\sin\lambda_n t\sin(2n-\lambda_n)t|}{t^2}\,dt.$$

Let us split the integral into three parts:

$$\int_0^{\frac{\pi}{2}} = \int_0^{\frac{1}{2n-\lambda_n}} + \int_{\frac{1}{2n-\lambda_n}}^{\frac{1}{\lambda_n}} + \int_{\frac{1}{\lambda_n}}^{\frac{\pi}{2}} \equiv I_1+I_2+I_3.$$

I_1 and I_2 can easily be estimated for any $\alpha \leqq 1$ as follows.

$$\frac{1}{\lambda_n} I_1 \leqq \frac{1}{\lambda_n} \int_0^{\frac{1}{2n-\lambda_n}} t^\alpha \lambda_n (2n-\lambda_n)\, dt \leqq \frac{1}{(2n-\lambda_n)^\alpha} \leqq \frac{1}{\lambda_n^\alpha},$$

$$\frac{1}{\lambda_n} I_2 \leqq \frac{1}{\lambda_n} \int_{\frac{1}{2n-\lambda_n}}^{\frac{1}{\lambda_n}} t^{\alpha-1} \lambda_n\, dt \leqq \frac{1}{\alpha} \frac{1}{\lambda_n^\alpha}.$$

The estimate of I_3 differs according to whether $\alpha < 1$ or $\alpha = 1$:

$$\frac{1}{\lambda_n} I_3 \leqq \frac{1}{\lambda_n} \int_{\frac{1}{\lambda n}}^{\frac{\pi}{2}} t^{\alpha-2}\, dt \leqq \begin{cases} \dfrac{1}{1-\alpha} \dfrac{1}{\lambda_n^\alpha}, & \text{if} \quad \alpha < 1 \\[2ex] \dfrac{1+\log \lambda_n}{\lambda_n}, & \text{if} \quad \alpha = 1. \end{cases}$$

Collecting our estimates, we obtain (5).

Ad Theorem 5. Using the notations introduced in the proof of Theorem 3, it is obvious that if $\nu \geqq m$, then

$$s_\nu(f - T_m^*; x) = s_\nu(f; x) - T_m^*(x).$$

From this it follows that

$$\left\{ \frac{1}{\lambda_n} \sum_{\nu=n-\lambda_n}^{n-1} |s_\nu(x) - f(x)|^k \right\}^{\frac{1}{k}} \leqq \left\{ \frac{2^k}{\lambda_n} \sum_{\nu=n-\lambda_n}^{n-1} \left(|s_\nu(f - T_{n-\lambda_n}^*; x)|^k + \right. \right. \tag{2. 5}$$

$$\left. \left. + |T_{n-\lambda_n}^*(x) - f(x)|^k \right) \right\}^{\frac{1}{k}} \leqq 2^{1+\frac{1}{k}} \left(\left\{ \frac{1}{\lambda_n} \sum_{\nu=n-\lambda_n}^{n-1} |s_\nu(f - T_{n-\lambda_n}^*; x)|^k \right\}^{\frac{1}{k}} + E_{n-\lambda_n} \right).$$

Applying Lemma 2 (with $g(t) = f(t) - T_{n-\lambda_n}^*(t)$ and $q = k$), we get the required estimate (6) by (2. 5).

The statement can immediately be derived from Lemma 2.

The proof of Theorem 5 is thus completed.

Ad Theorem 6. The assumption $f^{(r)}(x) \in \text{Lip}\,\alpha$ implies that $E_n(f) = O\left(\frac{1}{n^{r+\alpha}}\right)$ and $E_n(\tilde{f}) = O\left(\frac{1}{n^{r+\alpha}}\right)$ (see [14], (13. 14) Theorem, p. 117., and (13. 29) Theorem, p. 121). From these and (6) we get, with $\lambda_n = \left[\frac{n}{2}\right]$ and $n = 2m$, that

$$h_m(f, k; x) \equiv \left\{ \frac{1}{m} \sum_{\nu=m}^{2m-1} |s_\nu(x) - f(x)|^k \right\}^{\frac{1}{k}} = O\left(\frac{1}{m^{r+\alpha}} \right) \tag{2. 6}$$

and

$$h_m(\tilde{f}, k; x) \equiv \left\{\frac{1}{m} \sum_{\nu=m}^{2m-1} |\tilde{s}_\nu(x) - \tilde{f}(x)|^k\right\}^{\frac{1}{k}} = O\left(\frac{1}{m^{r+\alpha}}\right),$$

where $\tilde{s}_\nu(x) = s_\nu(\tilde{f}; x)$.

Suppose that $2^{m_1} \leqq n - \lambda_n + 1 < 2^{m_1+1}$ and $2^{m_2} < n \leqq 2^{m_2+1}$. Then, by (2. 6), we have

$$\frac{1}{\lambda_n} \sum_{\nu=n-\lambda_n}^{n-1} |s_\nu(x) - f(x)|^k \leqq \frac{1}{\lambda_n} \sum_{m=m_1}^{m_2} \sum_{\nu=2^m}^{2^{m+1}-1} |s_\nu(x) - f(x)|^k \leqq$$

$$\leqq \frac{O(1)}{\lambda_n} \sum_{m=m_1}^{m_2} 2^{m(1-k(r+\alpha))} \equiv \sum\nolimits_1 .$$

If $k(r+\alpha) < \alpha$, then, by $n = O(\lambda_n)$ and $2^{m_2} < n$,

$$\sum\nolimits_1 \leqq O(1) \frac{1}{\lambda_n} 2^{m_2(1-k(r+\alpha))} = O\left(\frac{1}{n^{k(r+\alpha)}}\right);$$

and if $k(r+1) = 1$, then

$$\sum\nolimits_1 \leqq O(1) \frac{1}{\lambda_n} (m_2 - m_1) = O\left(\frac{1}{n}\left(1 + \log \frac{n}{n - \lambda_n + 1}\right)\right).$$

From these estimates the statement (8) obviously follows.

The statement for the conjugate function can similarly be verified.

In order to prove the statements (9) and (10) let us define the following functions. If r is an even integer, let

$$f_1(x) = \sum_{n=1}^{\infty} \frac{\cos nx}{n^{1+r+\alpha}}$$

and

$$f_2(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{2^{n\alpha}} \sum_{l=2^{n-1}+1}^{2^n} \left(\frac{\sin(5 \cdot 2^n - l)x}{(5 \cdot 2^n - l)^r l} - \frac{\sin(5 \cdot 2^n + l)x}{(5 \cdot 2^n + l)^r l}\right);$$

and if r is an odd integer, let

$$f_1(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{2^{n\alpha}} \sum_{l=2^{n-1}+1}^{2n} \left(\frac{\cos(5 \cdot 2^n - l)x}{(5 \cdot 2^n - l)^r l} - \frac{\cos(5 \cdot 2^n + l)x}{(5 \cdot 2^n + l)^r l}\right)$$

and

$$f_2(x) = \sum_{n=1}^{\infty} \frac{\sin nx}{n^{1+r+\alpha}}.$$

It is proved that the r-th derivatives of these functions belong to Lip α (see [9], Satz IV). We shall now verify the inequalities (9) and (10) only for even r since the other case would be an analogous computation.

On account of $(r+\alpha)k=1$ and $n=O(\lambda_n)$, we have

$$T_n(f_1,\lambda,k;0)=\left\{\frac{1}{\lambda_n}\sum_{\nu=n-\lambda_n}^{n-1}\left(\sum_{m=\nu+1}^{\infty}\frac{1}{m^{r+1+\alpha}}\right)^k\right\}^{\frac{1}{k}}\geqq$$

$$\geqq C_1(k)\left\{\frac{1}{\lambda_n}\sum_{\nu=n-\lambda_n}^{n-1}\frac{1}{\nu}\right\}^{\frac{1}{k}}\geqq C_2(k)\left\{\frac{1}{n}\log\frac{n}{n-\lambda_n+1}\right\}^{\frac{1}{k}}\geqq$$

$$\geqq C_2(k)\left\{\frac{1}{n}\left(1+\log\frac{n}{n-\lambda_n+1}\right)\right\}^{\frac{1}{k}}\geqq C_4(k)n^{-r-\alpha}\left(1+\log\frac{n}{n-\lambda_n+1}\right)^{\frac{1}{k}},$$

which gives (9).

The proof of (10) is more complicated. We will first prove (10) in case that the sequence $\left\{\frac{n}{n-\lambda_n}\right\}$ is bounded. Let $n=12\cdot 2^m$ and let $\alpha_1=\max(n-\lambda_n, 22\cdot 2^{m-1})$, $\alpha_2=\max(\alpha_1, 23\cdot 2^{m-1})$ and $\alpha_3=\max\left(\alpha_2, n-\left[\frac{\lambda_n+1}{2}\right]\right)$. Then

$$T_n(\tilde f_2,\lambda,k;0)=\left\{\frac{1}{\lambda_n}\sum_{\nu=n-\lambda_n}^{n-1}|\tilde s_\nu(0)-\tilde f(0)|^k\right\}^{\frac{1}{k}}\geqq$$

$$\geqq\left\{\frac{1}{\lambda_n}\left(\sum_{\nu=\alpha_1}^{\alpha_2-1}+\sum_{\nu=\alpha_2}^{\alpha_3}\right)\left|\frac{1}{n^\alpha}\sum_{l=\nu-10\cdot 2^m+1}^{2^{m+1}}\frac{1}{n^r l}\right|^k\right\}^{\frac{1}{k}}.$$

It is clear by $n=O(\lambda_n)$, that

$$\sum_{\nu=\alpha_1}^{\alpha_2-1}\left|\frac{1}{n^\alpha}\sum_{l=\nu-10\cdot 2^m+1}^{2^{m+1}}\frac{1}{n^r l}\right|^k\geqq(\alpha_2-\alpha_1)\left|\frac{1}{n^\alpha}\sum_{l=\alpha_2-10\cdot 2^m}^{2^{m+1}}\frac{1}{n^r l}\right|^k\geqq$$

$$\geqq(\alpha_2-\alpha_1)\left|\frac{1}{n^{r+1+\alpha}}\min(2^{m-1},\lambda_n)\right|^k\geqq C_1(k)(\alpha_2-\alpha_1)\frac{1}{n^{(r+\alpha)k}}.$$

Similarly,

$$\sum_{\nu=\alpha_2}^{\alpha_3}\left|\frac{1}{n^\alpha}\sum_{l=\nu-10\cdot 2^m+1}^{2^{m+1}}\frac{1}{n^r l}\right|^k\geqq(\alpha_3-\alpha_2)\left|\frac{1}{n^\alpha}\sum_{l=\alpha_3-10\cdot 2^m}^{2^{m+1}}\frac{1}{n^r l}\right|^k\geqq$$

$$\geqq(\alpha_3-\alpha_2)\left|\frac{1}{n^{r+\alpha+1}}\min\left(2^{m-1},\frac{\lambda_n}{n}\right)\right|^k\geqq C_2(k)(\alpha_3-\alpha_2)\frac{1}{n^{(r+\alpha)k}}.$$

Hence we get

$$T_n(\tilde f_2,\lambda,k;0)\geqq C_3(k)\left\{(\alpha_3-\alpha_1)\frac{1}{\lambda_n}\frac{1}{n^{(r+\alpha)k}}\right\}^{\frac{1}{k}}\geqq C_4(k)\frac{1}{n^{r+\alpha}}$$

which proves (10) under the assumption that $\left\{\frac{n}{n-\lambda_n}\right\}$ is bounded.

If $\frac{n}{n-\lambda_n}$ tends to infinity, then the proof is simpler. Suppose that $4\cdot 2^m < n \leqq \leqq 4\cdot 2^{m+1}$ and $4\cdot 2^\mu \leqq n-\lambda_n+4 < 4\cdot 2^{\mu+1}$; and that $m > \mu+2$. Then we have

$$\begin{aligned} T_n(\tilde{f}_2, \lambda, k; 0) &\geqq \left\{\frac{1}{\lambda_n}\sum_{p=\mu+1}^{m-1}\sum_{\nu=4\cdot 2^p+1}^{4\cdot 2^{p+1}} |\tilde{s}_\nu(0)-\tilde{f}(0)|^k\right\}^{\frac{1}{k}} \geqq \\ &\geqq \left\{\frac{1}{\lambda_n}\sum_{p=\mu+1}^{m-1}\sum_{\nu=11\cdot 2^{p-1}}^{12\cdot 2^{p-1}} |\tilde{s}_\nu(0)-\tilde{f}(0)|^k\right\}^{\frac{1}{k}} \equiv \left\{\frac{1}{\lambda_n}\sum_{p=\mu+1}^{m-1} I_p\right\}^{\frac{1}{k}}. \end{aligned} \tag{2.7}$$

I_p can easily be estimated as follows

$$\begin{aligned} I_p &\geqq \sum_{\nu=11\cdot 2^{p-1}}^{23\cdot 2^{p-2}} \left(\frac{1}{2^{p\alpha}}\sum_{l=\nu-10\cdot 2^{p-1}+1}^{2p} \frac{1}{6^r\cdot 2^{pr}l}\right)^k \geqq \\ &\geqq \sum_{\nu=11\cdot 2^{p-1}}^{23\cdot 2^{p-2}} \left(\frac{1}{2^{p\alpha}}\sum_{l=23\cdot 2^{p-2}-10\cdot 2^{p-1}+1}^{2^p} \frac{1}{6^r 2^{pr}l}\right)^k \geqq \\ &\geqq C_1(r,k)2^{p-2}\frac{1}{2^{p(r+\alpha)k}} = C_2(r,k) \equiv C_2 > 0. \end{aligned}$$

From this and (2.7) we get

$$T_n(\tilde{f}_2, \lambda, k; 0) \geqq C_2^{\frac{1}{k}}\left(\frac{1}{\lambda_n}(m-\mu-2)\right)^{\frac{1}{k}} = C_2^{\frac{1}{k}}\left\{\frac{1}{n}\left(\frac{1}{\log 2}\log\frac{n}{n-\lambda_n+1}-6\right)\right\}^{\frac{1}{k}}.$$

Since $(r+\alpha)k=1$ and $\frac{n}{n-\lambda_n+1}\to\infty$, it is easy to see that, for n large enough,

$$T_n(\tilde{f}_2, \lambda, k; 0) \geqq \frac{c}{n^{r+\alpha}}\left(1+\log\frac{n}{n-\lambda_n+1}\right)^{\frac{1}{k}}$$

where $c>0$ is independent of n.

We have completed our proof.

Ad Theorem 7. To prove (12), by (11) and (1.4), it is sufficient to choose $p(>1)$ so close to 1 that $1-\frac{1}{p}\leqq\delta$, which is possible.

Ad Theorem 8. In order to prove (13), it is sufficient to show that, by suitable choice of p occurring in (1.4), the estimate

$$n^{1-\frac{1}{p}}\left\{\sum_{\nu=0}^{n}(A_{n-\nu}^{(\alpha-1)})^p\right\}^{\frac{1}{p}} = O(A_n^\alpha) \tag{2.8}$$

holds true. Let us choose p such that $(\alpha-1)p>-1$. Then we have

$$\sum_{\nu=0}^{n}(A_{n-\nu}^{(\alpha-1)})^p=\left(\sum_{\nu=0}^{\left[\frac{n}{2}\right]}+\sum_{\nu=\left[\frac{n}{2}\right]+1}^{n}\right)(A_{n-\nu}^{(\alpha-1)})^p\leqq$$

$$\leqq O(1)\sum_{\nu=0}^{\left[\frac{n}{2}\right]}n^{(\alpha-1)p}+O(1)\sum_{\nu=1}^{\left[\frac{n}{2}\right]}\nu^{(\alpha-1)p}=O(n^{1+(\alpha-1)p}).$$

From this we obtain

$$n^{1-\frac{1}{p}}\left\{\sum_{\nu=0}^{n}(A_{n-\nu}^{(\alpha-1)})^p\right\}^{\frac{1}{p}}=O(n^\alpha)$$

which is equivalent to (2. 8) as was required to be proved.

Ad Theorems 9 and 10. In [10] it is proved that if $\bar{\lambda}(x)$ is a positive, monotonic function with the property

$$\sum_{n=m}^{\infty}\frac{1}{n\bar{\lambda}(n)}\leqq K\frac{1}{\bar{\lambda}(m)}, \tag{2. 9}$$

then the conditions

$$\sum_{n=1}^{\infty}\frac{1}{\bar{\lambda}(n)}E_n^{(2)}(f)<\infty,$$

$$\int_0^1\frac{1}{t^2\bar{\lambda}\left(\frac{1}{t}\right)}\left(\int_0^{2\pi}[f(x+t)-f(x-t)]^2\,dx\right)^{\frac{1}{2}}dt<\infty$$

are equivalent.

The function $\lambda(x)$ occurring in Theorem 9 satisfies (2. 9), viz.

$$\sum_{n=m}^{\infty}\frac{1}{n\sqrt{n\lambda_n}}\leqq\frac{1}{\sqrt{\lambda_m}}\sum_{n=m}^{\infty}\frac{1}{n^{3/2}}\leqq K\frac{1}{\sqrt{m\lambda_m}};$$

(14) and (15) are therefore equivalent.

By virtue of the equivalence, it is sufficient to prove that condition (14) implies the $|V,\lambda|$-summability of (1).

We proved in [7] (Satz VIII) that any orthogonal series $\sum c_n\varphi_n(x)$ is $|V,\lambda|$-summable a.e. under the condition

$$\sum_{m=0}^{\infty}\left(\sum_{n=\mu_m+1}^{\mu_{m+1}}c_n^2\right)^{\frac{1}{2}}<\infty,$$

where $\mu_0=1$ and $\mu_m=\sum_{k=0}^{m-1}\lambda_{\mu_k}$. Therefore, it only remains to prove that the inequality

$$\sum_{m=0}^{\infty}C_m\equiv\sum_{m=0}^{\infty}\left\{\sum_{n=\mu_m+1}^{\mu_{m+1}}(a_n^2+b_n^2)\right\}^{\frac{1}{2}}<\infty \tag{2. 10}$$

follows from condition (14). Since $\sum_{k=n+1}^{\infty}(a_k^2+b_k^2)=(E_n^{(2)}(f))^2$, we obtain by a simple computation that

$$\sum_{n=2}^{\infty} C_n = \sum_{m=1}^{\infty}\sum_{n=2^m}^{2^{m+1}-1} C_n \leqq \sum_{m=1}^{\infty} 2^{\frac{m}{2}}\left\{\sum_{n=2^m}^{2^{m+1}-1} C_n^2\right\}^{\frac{1}{2}} \leqq \sum_{m=1}^{\infty} 2^{\frac{m}{2}} E_{\mu_{2^m}}^{(2)}. \tag{2.11}$$

If we can show that

$$\sum_{n=\mu_{2^{m-1}}+1}^{\mu_{2^m}} \frac{1}{\sqrt{n\lambda_n}} \geqq \frac{1}{4} 2^{\frac{m}{2}}, \tag{2.12}$$

then, by (2. 11) and (2. 12), we have

$$\sum_{n=2}^{\infty} C_n \leqq 4\sum_{m=1}^{\infty}\sum_{n=\mu_{2^{m-1}}+1}^{\mu_{2^m}} \frac{1}{\sqrt{n\lambda_n}} E_{\mu_{2^m}}^{(2)} \leqq 4\sum_{n=1}^{\infty}\frac{1}{\sqrt{n\lambda_n}} E_n^{(2)},$$

i.e. the condition (14) implies (2. 10) indeed as we stated.

It remains to prove (2. 12). Using the definition of μ_n, we get

$$\sum_{n=\mu_{2^{m-1}}+1}^{\mu_{2^m}} \frac{1}{\sqrt{n\lambda_n}} \geqq \sum_{k=2^{m-1}}^{2^m-1}\sum_{n=\mu_k+1}^{\mu_{k+1}} \frac{1}{\sqrt{n\lambda_n}} \geqq \sum_{k=2^{m-1}}^{2^m-1} \lambda_{\mu_k}\frac{1}{\sqrt{\mu_{k+1}\lambda_{\mu_{k+1}}}} \geqq$$

$$\geqq \sum_{k=2^{m-1}}^{2^m-1} \lambda_{\mu_k}\frac{1}{2\sqrt{\mu_k\lambda_{\mu_k}}} \geqq \frac{1}{2}\sum_{k=2^{m-1}}^{2^m-1}\left(\frac{\lambda_{\mu_k}}{\sum_{i=0}^{k-1}\lambda_{\mu_i}}\right)^{\frac{1}{2}} \geqq \frac{1}{2}\sum_{k=2^{m-1}}^{2^m-1}\frac{1}{\sqrt{k}} \geqq \frac{1}{4}2^{\frac{m}{2}},$$

and this completes the proof.

References

[1] G. ALEXITS—D. KRÁLIK, Über die Approximation im starken Sinne, *Acta Sci. Math.*, **26** (1965), 93—101.

[2] N. K. BARY, *A treatise on trigonometric series.* I (New York, 1964).

[3] S. N. BERNSTEIN, Sur l'ordre de la meilleure approximation des fonctions analytiques, *Mém. Acad. Belgique*, **4** (1912), 1—103.

[4] A. B. Ефимов, О прибцижении периодических функций суммами Валле-Пуссена, *Изв. Акад. Наук СССР*, **23** (1959), 737—770.

[5] G. H. HARDY—J. E. LITTLEWOOD, Sur la série de Fourier d'une fonction à carré sommable, *C. R. Acad. Sci. Paris*, **156** (1913), 1307—1309.

[6] H. LEBESGUE, Sur les integrales singulières, *Ann. de Toulouse*, **1** (1909), 25—117.

[7] L. LEINDLER, Über die de la Vallée Poussinsche Summierbarkeit allgemeiner Orthogonalreihen, *Acta. Math. Acad. Sci. Hung.*, **16** (1965), 375—387.

[8] L. LEINDLER, On the absolute summability factors of Fourier series, *Acta Sci. Math.*, **28** (1967), 323—336.

[9] L. LEINDLER, Über die Approximation im starken Sinne, *Acta Math. Acad. Sci. Hung.*, **16** (1965), 255—262.

[10] L. LEINDLER, Über verschiedene Konvergenzarten trigonometrischer Reihen, *Acta. Sci. Math.*, **25** (1964), 233—249.

[11] J. MARCINKIEWICZ, Sur l'interpolation, *Studia Math.*, **6** (1936), 1—17, 67—81.

[12] CH. DE LA VALLÉE POUSSIN, *Leçons sur l'approximation des fonctions d'une variable réelle* (Paris, 1919).

[13] A. ZYGMUND, On the convergence and summability of power series on the circle of convergence, *Proc. London Math. Soc.*, **47** (1941), 326—350.

[14] A. ZYGMUND, *Trigonometric series.* I (Cambridge, 1959).

(Received August 20, 1967)

On a theorem of L. Takács

By L. GEHÉR in Szeged

L. TAKÁCS [1] proved the following generalization of the classical ballot theorem in probability theory:

Let $\Phi(u)$ be a nondecreasing function on the interval $I=[0, t]$, for which $\Phi'(u)=0$ almost everywhere and $\Phi(0)=0$. Set $\Phi(t+u) = \Phi(t)+\Phi(u)$ for $0 \leqq u \leqq t$. Define the function δ on I as follows; $\delta(u)=1$ if $\Phi(v)-\Phi(u) \leqq v-u$ for every v such that $u \leqq v \leqq u+t$, and $\delta(u)=0$ otherwise. Then

$$\int_0^t \delta(u)\, du = t-\Phi(t) \qquad \text{whenever } \Phi(t) \leqq t.$$

In the following we shall give a new proof of this theorem or rather of a somewhat more general theorem, which we formulate in a measure theoretic form:

Theorem. *Let m and μ be non-negative complete regular measures on a circle C such that m is absolutely continuous and μ is singular with respect to the Lebesgue measure on C, and $m(C) \geqq \mu(C)$. Denote by $\Delta[m, \mu]$ the set of the points $x \in C$ for which*

$$m[x, y) \geqq \mu[x, y)$$

holds for all $y \in C$, where $[x, y)$ denotes the semi-closed arc from x to y in the positive direction. Then $\Delta[m, \mu]$ is m-measurable (as well as Lebesgue measurable) and we have

$$m(\Delta[m, \mu]) = m(C)-\mu(C). \tag{1}$$

In the particular case that m is the Lebesgue measure, our theorem is obviously equivalent to TAKÁCS's theorem.

The proof of the theorem will be done in several steps.

1. First suppose that μ is concentrated into finitely many points $x_1, x_2, \cdots, x_n$ and that $\alpha_1, \alpha_2, \cdots, \alpha_n$ are the corresponding measures. Then we prove the theorem by induction on n. The case $n=1$ is trivial. Indeed, if we associate with x_1 the minimal arc $(\xi_1, x_1]$ of measure α_1 (m being continuous with $m(C) \geqq \mu(C)$, this can be done), the complementary set of $(\xi_1, x_1]$ coincides with $\Delta[m, \mu]$. Suppose the theorem is

true for n and then we prove it for $n+1$. Let μ be concentrated in the points $x_1, x_2, \cdots, x_{n+1}$ with the corresponding measures $\alpha_1, \alpha_2, \cdots, \alpha_{n+1}$. We associate with every point x_i $(i=1, 2, \cdots, n+1)$ the minimal arc $(\xi_i, x_i]$ of m-measure α_i. If these arcs are mutually non-overlapping then the complementary set of the union of these arcs coincides with $\Delta[m, \mu]$, which proves our statement. In case that there exists an x_k contained in some $(\xi_i, x_i]$, consider the atomic measure μ' that we obtain from μ by shifting the measure α_k from the point x_k to the point x_i so that x_k will have measure 0 and x_i will have measure $\alpha_i+\alpha_k$. This measure μ' has only n atoms and we have $\Delta[m, \mu]=\Delta[m, \mu']$. Using our induction hypothesis, our theorem is proved.

2. Let μ be an arbitrary atomic measure, its atoms being at the points $x_1, x_2, \cdots$, with the measures $\alpha_1, \alpha_2, \cdots$, respectively. Denote by μ_n $(n=1, 2, \cdots)$ the atomic measure whose atoms are at $x_1, x_2, \cdots, x_n$ with the corresponding measures $\alpha_1, \alpha_2, \cdots, \alpha_n$. Our statement easily follows from the fact $\Delta[m, \mu]=\bigcap_{n=1}^{\infty} \Delta[m, \mu_n]$.

Let namely $x\in\Delta[m, \mu]$, then a fortiori we have $x\in\Delta[m, \mu_n]$ for every n, which gives that $\Delta[m, \mu]\subseteq\bigcap_{n=1}^{\infty}\Delta[m, \mu_n]$. Conversely, suppose that $x\in\bigcap_{n=1}^{\infty}\Delta[m, \mu_n]$. Then for any n and $y\in C$ we have $m[x, y)\geqq\mu_n[x, y)$, from which we get $m[x, y)\geqq \lim_{n\to\infty}\mu_n[x, y)=\mu[x, y)$, i.e. $x\in\Delta[m, \mu]$. Summarizing, we obtain, that $\Delta[m, \mu]=\bigcap_{n=1}^{\infty}\Delta[m, \mu_n]$, which was to be proved.

3. Let μ be an arbitrary singular measure and denote by Z a set of (Lebesgue) measure 0 such that μ vanishes on every measurable set contained in the complementary set of Z. Choose to Z countable many covering systems $\{I_i^{(n)}\}_{i=1}^{\infty}$ $(n=1, 2, \cdots)$ of disjoint open arcs such that

$$\bigcup_{i=1}^{\infty} I_i^{(n)}\supset\bigcup_{i=1}^{\infty} I_i^{(n+1)} \qquad \text{and} \qquad \sum_{i=1}^{\infty} m(I_i^{(n)}) < \frac{1}{2^n} \qquad (n = 1, 2, \ldots).$$

To the n-th covering system we associate the atomic measure ν_n whose atoms are at the starting points of the arcs $I_i^{(n)}$ $(i=1, 2, \cdots)$, each with measure $\mu(I_i^{(n)})$. Set

$$\Delta_n = \Delta[m, \nu_n]\cup\left(\bigcup_{i=1}^{\infty} I_i^{(n)}\right).$$

Then we have

$$\Delta_1\supset\Delta_2\supset\cdots \qquad \text{and} \qquad \Delta[m, \mu]\subset\bigcap_{n=1}^{\infty}\Delta_n.$$

In an analogous manner we construct a second system of atomic measures by associating with the n-th covering system an atomic measure ν_n' whose atoms are at the terminal points of the arcs $I_i^{(n)}$ $(i=1, 2, \cdots)$, each with measure $\mu(I_i^{(n)})$. Set

$$\Delta_n' = \Delta[m, \nu_n'] - \left(\bigcup_{i=1}^{\infty} I_i^{(n)} \right);$$

it can be easily seen that

$$\Delta_1' \subset \Delta_2' \subset \cdots \qquad \text{and} \qquad \Delta[m, \mu] \supset \bigcup_{n=1}^{\infty} \Delta_n'.$$

Summarizing, we obtain $\bigcup_{n=1}^{\infty} \Delta_n' \subset \Delta[m, \mu] \subset \bigcap_{n=1}^{\infty} \Delta_n$. Since

$$\lim_{n\to\infty} m(\Delta_n) = \lim_{n\to\infty} m(\Delta_n') = m(C) - \mu(C),$$

we get that $\Delta[m, \mu]$ is m-measurable and $m(\Delta[m, \mu]) = m(C) - \mu(C)$, i. e. (1).

Remark. The theorem remains true if m is only continuous, but μ is singular *with respect to m*. The proof is essentially the same.

Reference

[1] L. Takács, *Combinatorial Methods in the Theory of Stochastic Processes* (New York—London—Sydney, 1967).

(Received March 1, 1967)

Generators for groups of permutation polynomials over finite fields

By CHARLES WELLS in Cleveland (Ohio, U. S. A.)*)

1. Introduction. Let $GF(q)$ denote the finite field of order $q=p^n$. If Φ is a function from $GF(q)$ to $GF(q)$, a polynomial f over $GF(q)$ is said to represent Φ if $f(\xi)=\Phi(\xi)$ for all $\xi \in GF(q)$. If follows from the Lagrange interpolation formula that every such function Φ is represented by a unique polynomial f of degree $\leqq q-1$. (No such simple theorem is true over the ring of integers (mod p^n); see CARLITZ [5], NÖBAUER [12], RÉDEI and SZELE [13].)

A *permutation polynomial* is simply a polynomial which represent a permutation. The first systematic investigation of permutation polynomials was undertaken by DICKSON [8, 9]; the permutation polynomials over $GF(p)$ had previously been investigated by HERMITE [11]. Other references to early work done on special cases may be found in DICKSON [8].

DICKSON's work suggested much of the work done since with permutation polynomials. His longest and most detailed investigation culminated in his listing of all the permutation polynomials of degree $\leqq 6$ for all $GF(q)$. (We note here that CAVIOR [6] extended these results partially to octic binomial permutation polynomials.)

By means of this list DICKSON proved that the symmetric group on 7 letters was generated by the permutations x^5 and $\alpha x+\beta$ (α, $\beta \in GF(7)$, $\alpha \neq 0$). This suggested our Theorem 4. 1, first proved by CARLITZ [2]. By a modification of CARLITZ's method, FRYER [10] found generators for the alternating group on p letters (Theorem 4. 6).

The present paper contains a number of new theorems on generators of the symmetric group on q letters and its subgroups. These include a sharpening of CARLITZ's result (Theorem 4. 2) and the presentation of generators of three small subgroups (Theorem 4. 4). A more interesting result is the discovery of several sets of generators for the alternating group on q letters (Theorems 4. 7 and 4. 8).

*) Supported by National Science Foundation Grant GF—1891.

None of these sets of generators is a direct generalization of FRYER's Theorem; it may be that that result cannot be generalized in a satisfactory way.

Theorems on generators of S_{q+1} and A_{q+1} are found in Sections 6—8 by means of a device used by BURNSIDE [1; p. 185], CARLITZ [3], and others. An element ∞ is added to $GF(q)$, forming the extended domain $\overline{GF(q)}$. Rational functions $f(x)/g(x)$ (where f and g are polynomials over $GF(q)$) are well defined as mappings of $\overline{GF(q)}$ into $\overline{GF(q)}$, and all permutations of $\overline{GF(q)}$ are representable by rational functions, as CARLITZ showed [3, p. 326—327]. Theorems 7. 2, 7. 3, 8. 2, and 8. 3 exhibit generators of S_{q+1} and A_{q+1} in terms of rational functions.

2. Preliminaries. Throughout the following, q will be assumed to be fixed and greater than 2. Many of the theorems are false for $q=2$ (for example Lemma 4. 3).

$GF(q)$ always includes $GF(p)$. In this paper the elements of $GF(p)$ will be written as integers; it will be understood that k and $k+mp$ are the same element of $GF(p)$ for all m. It is in this sense that a formula like (3. 2) below should be understood.

If two polynomials represent the same function, then they differ by a polynomial multiple of x^q-x. The *reduced form* of a polynomial will here be taken to be the remainder obtained when the polynomial is divided by x^q-x. When two permutation polynomials are combined by the operation of composition, the result may be assumed to be in reduced form; in this sense the set of permutation polynomials of degree $\leqq q-1$ represents the symmetric group S_q. It is not hard to prove that in fact a permutation polynomial cannot have degree $q-1$.

It is convenient to write

$$\langle g(x)\rangle \langle f(x)\rangle = \langle h(x)\rangle \tag{2. 1}$$

when $f(g(x))\equiv h(x)$ (mod x^q-x). Then $\langle g(x)\rangle$ is the function represented by $g(x)$. However, except when it is convenient to write out formulas like (2. 1), we shall follow the usual practice of identifying the polynomial and the function.

3. We collect here some elementary facts about permutation polynomials. In the first place, it follows from the cancellation laws that αx and $x+\beta$ are permutation polynomials for any β and any $\alpha\neq 0$ in $GF(q)$. It follows from a theorem of DICKSON [9; p. 59] that x^b is a permutation polynomial for any integer b such that $(b, q-1)=1$. This may also be proved directly: let ϱ be a primitive root of $GF(q)$, set $\alpha=\varrho^r$, $\beta=\varrho^s$, and note that $\alpha^b=\beta^b$ if and only if $rb\equiv sb$ (mod $q-1$).

In particular, x^{q-2} is a permutation polynomial. It is, in fact, the function that takes every nonzero element into its inverse.

For later use we note the following rules of calculation:

$$(3.1)\qquad \langle x+\alpha\rangle\langle x+\beta\rangle=\langle x+\alpha+\beta\rangle \qquad (\alpha,\ \beta\in GF(q)).$$

$$(3.2)\qquad \langle x+\alpha\rangle^{s}=\langle x+s\alpha\rangle \qquad (\alpha\in GF(q),\ s \text{ any integer}).$$

$$(3.3)\qquad \langle \alpha x\rangle^{s}=\langle \alpha^{s}x\rangle \qquad (\alpha\in GF(q),\ \alpha\neq 0,\ s \text{ any integer}).$$

$$(3.4)\qquad \langle x^{q-2}\rangle^{2}=\langle x\rangle.$$

$$(3.5)\qquad \langle \alpha x\rangle\langle x+\beta\rangle=\langle \alpha x+\beta\rangle \qquad (\alpha,\ \beta\in GF(q),\ \alpha\neq 0).$$

$$(3.6)\qquad \langle x^{q-2}\rangle\langle \alpha x\rangle=\langle \alpha x^{q-2}\rangle \qquad (\alpha\in GF(q),\ \alpha\neq 0).$$

Since the composition of two permutations is a permutation, it follows that the functions on the right side of the above equations are all permutations.

4. Generators of $A_q^{'}$. In [2] CARLITZ proved:

Theorem 4. 1. *S_q is generated by*

$$(4.1)\qquad \alpha x+\beta,\ x^{q-2} \qquad (\alpha,\ \beta\in GF(q),\ \alpha\neq 0).$$

The proof consists in noting that the polynomial

$$(4.2)\qquad g_{\gamma}(x) = -\gamma^{2}[((x-\gamma)^{q-2}+\gamma^{-1})^{q-2}-\gamma]^{q-2}$$

represents the transposition (0γ).

Of course, several sets of generators of the abstract symmetric group are known; see, for example, COXETER and MOSER [7; pp. 63—66]. The transpositions form such a set. The value of the generators found in this section is that they are simple as polynomials; it is evident from (4. 2) that simplicity as polynomials and simplicity as permutations are not equivalent.

We may simplify Theorem 4. 1 as follows. Let ϱ be a primitive root of $GF(q)$.

Theorem 4. 2. *S_q is generated by*

$$(4.3)\qquad \varrho x,\ x+1, \ \textit{and}\ \ x^{q-2}.$$

Proof. Let $\alpha,\ \beta\in GF(q)$, $\alpha\beta\neq 0$. Let $\alpha=\varrho^{s}$, $\beta=\varrho^{t}$. Then the proof follows from (3. 3) and

$$(4.4)\qquad \langle \alpha x+\beta\rangle=\langle \varrho x\rangle^{s-t}\langle x+1\rangle\langle \varrho x\rangle^{t}.$$

By an elaboration of these methods we may find generators for the alternating group A_q. Since the polynomials given in (4. 1) and (4. 3) do not necessarily represent even permutations, we first prove

Lemma 4. 3. *For all α, $x+\alpha$ and $(x^{q-2}+\alpha)^{q-2}$ are even; αx is even if and only if α is a nonzero square; x^{q-2} is even if and only if* $q\equiv 3 \pmod 4$.

Proof. By (3. 1) $x+\alpha$ is composed of p^{n-1} cycles of length p. Thus, if p is odd, or if $q=2^n$ and $n>1$, then $x+\alpha$ is even. Since $\langle (x^{q-2}+\alpha)^{q-2}\rangle = \langle x^{1-2}\rangle \cdot \langle x+\alpha\rangle \langle x^{q-2}\rangle$, it is even regardless of what x^{q-2} is.

By (3. 3) αx is a power of ϱx (as permutations), which is a cycle of length $q-1$. The second clause follows from this and the fact that every element of $GF(2^n)$ is square.

As a permutation, x^{q-2} consists of disjoint transpositions containing all elements of $GF(q)$ except 0, 1, -1. It therefore contains $\frac{1}{2}(q-3)$ transpositions when q is odd and $\frac{1}{2}(q-2)$ when q is even (since then $1=-1$). This proves the lemma.

We now define the following sets:

$$L_q = \{\alpha x+\beta \mid \alpha,\ \beta\in GF(q),\ \alpha\neq 0\}$$

$$AL_q = \{\alpha^2 x+\beta \mid \alpha,\ \beta\in GF(q),\ \alpha\neq 0\}$$

$$Q_q = \{(x^{q-2}+\alpha)^{q-2} \mid \alpha\in GF(q)\}.$$

The following equations imply that L_q, AL_q, and Q_q are actually groups

$$\langle \alpha x+\beta\rangle\langle \gamma x+\delta\rangle = \langle \alpha\gamma x+\beta\gamma+\delta\rangle \tag{4. 5}$$

$$\langle (x^{q-2}+\alpha)^{q-2}\rangle\langle (x^{q-2}+\beta)^{q-2}\rangle = \langle (x^{q-2}+\alpha+\beta)^{q-2}\rangle. \tag{4. 6}$$

Evidently the order of L_q is $q(q-1)$, that of AL_q is $\frac{1}{2}q(q-1)$, q odd, and that of Q_q is q. Q_q is isomorphic to the additive group of $GF(q)$. We have

Theorem 4. 4. *L_q is generated by ϱx and $x+1$. AL_q is generated by $\varrho^2 x$ and $x+1$. The elements of Q_q may be obtained from $(x^{q-2}+1)^{q-2}$ and $\varrho^2 x$. Furthermore, $AL_q \subseteqq A_q$, $Q_q\subseteqq A_q$.*

Proof. The first two sentences follow from (4. 4) and the fact that every element in a finite field is the sum of two squares (see [12; p. 46]). The third sentence follows from the last mentioned fact, (3. 3), and

$$\langle (x^{q-2}+\alpha^2)^{q-2}\rangle = \langle \alpha^2 x\rangle\langle (x^{q-2}+1)^{q-2}\rangle\langle \alpha^{-2}x\rangle. \tag{4. 7}$$

The last sentence follows from Lemma 4. 3.

The groups L_q and AL_q were first considered by BURNSIDE [1; pp. 181—185].

We now prove a lemma on generators for the alternating group A_n on n letters $\{0, 1, \cdots, n-1\}$. Let $R=(0\ 1\ 2)$ and $S=(0\ 1\ 2 \cdots n-1)$.

Lemma 4. 5. *For odd n, A_n is generated by R and S.*

Proof. We have

$$(0\ 1\ 3) = S^{-1}R^{-1}SR$$

and

$$(0\ 1\ i+1) = (0\ 1\ i)S^{-i+1}RS^{i-1}(0\ 1\ i)^{-1}(0\ 1\ i-1),$$

for $i=3, \cdots, n-1$. Since the permutations $(0\ 1\ i)$ $(i=2, 3, \cdots, n-1)$ generate A_n, this proves the lemma.

Using this lemma, FRYER [10] proved

Theorem 4. 6. *Let p be an odd prime. Then A_p is generated by $x+1$ and mx^{p-2}, where m is any square if $p \equiv 3$ (mod 4) and any nonsquare if $p \equiv 1$ (mod 4). Otherwise $x+1$ and mx^{p-2} generate S_p.*

Now FRYER's proof depends on the fact that in $GF(p)$ the permutation $x+1$ is a single cycle containing all the elements of the field, and so is the S of Lemma 4. 5. But for general $GF(p^n)$, $x+1$ contains n cycles and FRYER's proof does not work.

However, we may find generators for the general case using (4. 2). For the elements $(0\ 1\ \alpha)$ $(\alpha \in GF(q))$ generate A_q, and

$$(0\ 1\ \alpha) = (0\ 1)(0\ \alpha) = \langle g_1(x)\rangle \langle g_\alpha(x)\rangle =$$

$$= \langle x-1\rangle Q\langle x+1\rangle Q\langle x-1\rangle Q\langle -x\rangle \langle x-\alpha\rangle Q\langle x-\alpha^{-1}\rangle Q\langle x-\alpha\rangle Q\langle -\alpha^2 x\rangle,$$

where Q denotes the permutation x^{q-2}.

Now for $q \equiv 0$ or 1 (mod 4) this has the form

$$E(OEO)E(OEEO)E(OEO)E \tag{4. 8}$$

where E stands for any even permutation and O for any odd one. Grouping in the manner shown we obtain the generators $\alpha^2 x+\beta$ $(\alpha, \beta \in GF(q), \alpha \neq 0)$ and $(x^{q-2}+\gamma)^{q-2}(\gamma \in GF(q))$. For $q \equiv 3$ (mod 4) we have

$$EEEEEEOEEEEEEO, \tag{4. 9}$$

but we may bring the two odd permutations together by noting that $-x$ commutes (as a permutation) with αx and with x^{q-2}, and that

$$\langle -x\rangle \langle x+1\rangle = \langle x-1\rangle \langle -x\rangle.$$

After this is done we may group them as in (4. 8) to obtain the same set of generators. We therefore have

Theorem 4. 7. *The alternating group A_q is generated by its subgroups AL_q and Q_q.*

Of course, (4. 9) implies the existence of a simpler set of generators, which is incorporated in the following theorem:

Theorem 4. 8. *The alternating group A_q is generated by $\varrho^2 x$, $x+1$, and any one of the elements in the following list:*

(i) $(x^{q-2}+1)^{q-2}$ (*all* q),

(ii) x^{q-2} ($q \equiv 3 \pmod 4$),

(iii) αx^{q-2} ($q \equiv 1 \pmod 4$, α *not square*; *or* $q \equiv 3 \pmod 4$, α *square*).

Proof. The theorem follows from Theorems 4. 4 and 4. 7, and from (4. 9), Lemma 4. 3, and the following formula:

$$\langle (x^{q-2}+1)^{q-2} \rangle = \langle \alpha x^{q-2} \rangle \langle x+\alpha \rangle \langle \alpha x^{q-2} \rangle. \tag{4. 10}$$

5. Another method of proof. The fact that AL_q and x^{q-2} generate A_q whenever αx^{q-2} is even may also be deduced independently by a method resembling the proof of FRYER's Theorem (Theorem 4. 6). It follows from Lemma 4. 5 by properly renumbering the elements of $GF(q)$ that the permutations

$$(0\ 1\ \varrho) \quad \text{and} \quad T = (0\ 1\ \varrho\ \varrho^2 \cdots \varrho^{q-2}) \tag{5. 1}$$

generate A_q. Let $s=1$ if $q \equiv 1 \pmod 4$ and $s=2$ if $q \equiv 3 \pmod 4$, and let U be the permutation $\varrho^s x^{q-2}$. A lengthy calculation shows that

$$(0\ 1\ \varrho) = T^{-1}[(T\,U\,T)^2\, U (T\,U\,T)^2]^4 T$$

so that T and U generate A_q.

Since $T = \langle \varrho x \rangle \langle g_1(x) \rangle$, we may deduce by a method like that in (4. 8) and (4. 9) that A_q is generated by

$$(\varrho x-1)^{q-2}, \quad (x^{q-2}-1)^{q-2}, \quad \varrho^2 x, \quad x+1, \quad \text{and} \quad \varrho x^{q-2} \tag{5. 2}$$

when $q \equiv 1 \pmod 4$ and

$$-\varrho x, \quad x-1, \quad x^{q-2} \quad \text{and} \quad \varrho^2 x^{q-2} \tag{5. 3}$$

when $q \equiv 3 \pmod 4$. But we may replace $(x^{q-2}-1)^{q-2}$ by $(x^{q-2}+1)^{q-2}$ in (5. 2) since the former is merely the p-lst power of the latter (as permutations), and we may eliminate $(\varrho x-1)^{q-2}$ by means of the equation

$$\langle (\varrho x-1)^{q-2} \rangle = \langle \varrho^2 x - p \rangle \langle \varrho x^{q-2} \rangle.$$

We may replace $-\varrho x$ by $\varrho^2 x$ in (5. 3) because the former is the $\frac{1}{4}(q-1)$st power of the latter. Finally we may substitute $\varrho^m x^{q-2}$ for $\varrho^s x^{q-2}$ for m the proper parity (that is, we must have $m \equiv s \pmod 2$) in both (5. 2) and (5. 3) by means of

$$\langle \varrho^s x^{q-2} \rangle = \langle \varrho^m x^{q-2} \rangle \langle \varrho^{s-m} x^{q-2} \rangle.$$

This shows that αx^{q-2} and AL_q generate A_q when αx^{q-2} is even.

6. The extended domain. Let $r(x)=f(x)/g(x)$ be a rational function over $GF(q)$, where f and g are relatively prime polynomials over $GF(q)$ and g is primary. If g has a root $\beta \in GF(q)$, then r does not represent a function from $GF(q)$ into $GF(q)$ since $r(\beta)$ is undefined. We may evade this difficulty by adding an element ∞ to $GF(q)$ obeying the following rules of calculation:

$$r(\beta) = \begin{cases} f(\beta)g(\beta)^{-1} & (g(\beta) \neq 0) \\ \infty & (g(\beta) = 0) \end{cases} \tag{6. 1}$$

and

$$r(\infty) = \begin{cases} \infty & (\deg g < \deg f) \\ 0 & (\deg g > \deg f) \\ \operatorname{sgn} f & (\deg g = \deg f). \end{cases} \tag{6. 2}$$

From (6. 1) and (6. 2) we may deduce the usual rules of calculation. For example for $\alpha, \beta, \gamma, \delta \in GF(q)$ we have

$$\gamma \cdot \infty = \gamma/0 = \infty \qquad (\gamma \neq 0) \tag{6. 3}$$

$$\frac{\alpha\infty+\beta}{\gamma\infty+\delta} = \frac{\alpha}{\gamma} \qquad (\gamma \neq 0) \tag{6. 4}$$

$$f(\infty) = \infty \qquad (f \in GF[q, x], \deg f \geqq 1). \tag{6. 5}$$

The structure obtained from $GF(q)$ by adding ∞ in this manner is called the *extended domain* and is denoted by $\overline{GF(q)}$.

CARLITZ [3; pp. 326—327] showed that every permutation of $\overline{GF(q)}$ is representable by a rational function. In fact, he shows that every permutation is representable in the form $g(t(x))$, where g is a polynomial over $GF(q)$, and t is a member of the *general linear fractional group* of functions of the form

$$t(x) = \frac{\alpha x+\beta}{\gamma x+\delta} \qquad (\alpha, \beta, \gamma, \delta \in GF(q),\ \alpha\delta-\beta\gamma \neq 0).$$

7. Generators for S_{q+1}. We may find generators for S_{q+1} by using the following lemma:

Lemma 7. 1. *Let Ψ and Φ be in S_n, with Φ a cycle containing $n-1$ elements and Ψ a transposition containing the element not in Φ. Then Φ and Ψ generate S_n.*

The proof follows from the formula

$$(1\,2\cdots n-1)^{-k}\,(0\;1)(1\,2\cdots n-1)^k=(0\;k+1)\;(1\leqq k\leqq n-2).$$

Now considered as permutations of $\overline{GF(p)}$, $1/x^{p-2}$ is $(0\;\infty)$ and $x+1$ is $(0\;1\;2\cdots p-1)$, so by the lemma they generate S_{p+1}.

For the general case $q=p^n$ we consider the permutation $\varrho x+1$, where ϱ is a primitive root of $GF(q)$. This permutation takes $\infty\to\infty$, and

$$0\to1\to\varrho+1\to\varrho^2+\varrho+1\to\varrho^3+\varrho^2+\varrho+1\to\cdots\to\varrho^{q-2}+\varrho^{q-3}+\cdots+\varrho+1=0.$$

Since ϱ is a primitive q-1st root of unity, it follows that the above series contains no zeros except for the first and last elements. Hence, $\varrho x+1$ is a Ψ as in the lemma, and we have

Theorem 7. 2. *S_{p+1} is generated by $x+1$ and $1/x^{p-2}$. For all $q=p^n$, S_{q+1} is generated by $\varrho x+1$ and $1/x^{q-2}$.*

From this we have immediately

Theorem 7. 3. *S_{q+1} is generated by $1/x$, x^{q-2}, ϱx, and $x+1$.*

Theorem 7. 3 is, by Theorem 4. 4, equivalent to a theorem proved by CARLITZ [3; p. 328] (his proof uses the canonical form mentioned at the end of Section 6).

8. To find generators for the alternating group A_{q+1}, we first prove a lemma analogous to Lemma 4. 3.

Lemma 8. 1. *The permutation $x+\alpha$ is even over $\overline{GF(q)}$ for all $\alpha\in GF(q)$; x^{q-2} is even if and only if $q\equiv3$ (mod 4); $1/x$ is odd if and only if $q\equiv3$ (mod 4).*

Proof. The first two clauses follow immediately from Lemma 4. 3, since $x+\alpha$ and x^{q-2} leave ∞ unchanged. The other follows from

$$\langle x^{q-2}\rangle=\left\langle\frac{1}{x}\right\rangle(0\infty). \tag{8. 1}$$

This proves the lemma.

Now A_{q+1} is generated by the elements $(0\infty\alpha)$ $(\alpha\in GF(q))$. But

$$(0\infty\alpha)=(\alpha\infty)(0\infty)$$

and for any $\alpha\in F(q)$, including $\alpha=0$,

$$(\alpha\infty)=\left\langle\frac{1}{(x-\alpha)^{q-2}}+\alpha\right\rangle. \tag{8. 2}$$

Since

$$\left\langle \frac{1}{x^{q-2}} \right\rangle = \left\langle \frac{1}{x} \right\rangle \langle x^{q-2} \rangle = \langle x^{q-2} \rangle \left\langle \frac{1}{x} \right\rangle$$

we may write $(0\infty\alpha)$ in two ways:

(8. 3) $$(0\infty\alpha) = \langle x+\alpha \rangle \left\langle \frac{1}{x} \right\rangle \langle x^{q-2} \rangle \langle x+\alpha \rangle \langle x^{q-2} \rangle \left\langle \frac{1}{x} \right\rangle$$

and

(8. 4) $$(0\infty\alpha) = \langle x+\alpha \rangle \langle x^{q-2} \rangle \left\langle \frac{1}{x} \right\rangle \langle x+\alpha \rangle \left\langle \frac{1}{x} \right\rangle \langle x^{q-2} \rangle.$$

Grouping the third, fourth and fifth factors together in each of (8. 3) and (8. 4), we find that when $q \equiv 0$ or 1 (mod 4), A_{q+1} is generated by $1/x$, SL_q, and Q_q, where SL_q is the group of permutations of the form $x+\alpha$ $(\alpha \in GF(q))$; and when $q \equiv 3$ (mod 4), A_{q+1} is generated by x^{q-2}, SL_q, and Q'_q, where Q'_q is the group of permutations $(x^{-1}+\beta)^{-1}(\beta \in GF(q))$.

Now it is easy to see that Q_q, Q'_q, and SL_q are all isomorphic to the additive group of $GF(q)$ in the obvious manner. When $q=p$, these groups are cyclic, and we have the following particularly simple theorem:

Theorem 8. 2. *A_{p+1} is generated by*

$$\frac{1}{x}, \quad x+1, \quad (x^{p-2}+1)^{p-2} \qquad (p \equiv 0, 1 \,(\mathrm{mod}\, 4))$$

or

$$x^{p-2}, \quad x+1, \quad (x^{-1}+1)^{-1} \qquad (p \equiv 3 (\mathrm{mod}\, 4)).$$

By Theorem 4. 4 the elements of Q_q may be obtained from $(x^{q-2}+1)^{q-2}$ and $\varrho^2 x$, and the elements of SL_q from $\varrho^2 x$ and $x+1$ (since $SL_q \subseteqq AL_q$). Similarly $\varrho^2 x$ and $(x^{-1}+1)^{-1}$ give the elements of Q'_q. Hence, we have

Theorem 8. 3. *A_{q+1} is generated by*

$$\frac{1}{x}, \quad x+1, \quad \varrho^2 x, \quad (x^{q-2}+1)^{q-2} \qquad (q \equiv 0, 1 \,(\mathrm{mod}\, 4))$$

or

$$x^{q-2}, \quad x+1, \quad \varrho^2 x, (x^{-1}+1)^{-1} \qquad (q \equiv 3 (\mathrm{mod}\, 4)).$$

References

[1] W. BURNSIDE, *Theory of Groups of Finite Order* (New York, 1955).
[2] L. CARLITZ, Permutations in a finite field, *Proc. Amer. Math. Soc.*, **4** (1953), 538.
[3] L. CARLITZ, A note on permutation functions in a finite field, *Duke Math. J.*, **29** (1962), 120—123.
[4] L. CARLITZ, Permutations in finite fields, *Acta Sci. Math.*, **24** (1953), 196—203.
[5] L. CARLITZ, Functions and polynomials (mod p^n), *Acta Arithmetica*, **9** (1964), 67—78.
[6] S. CAVIOR, A note on octic permutation polynomials, *Mathematics of Computation*, **17** (1963), 450—452.
[7] H. S. M. COXETER and W. MOSER, *Generators and Relations for Discrete Groups*, second edition (Berlin, 1965).
[8] L. E. DICKSON, The analytic representation of substitutions on a power of a prime number of letters with a discussion of the linear group, *Annals of Math.*, **11** (1896—97), 65—110.
[9] L. E. DICKSON, *Linear Groups with an Exposition of Galois Field Theory* (New York, 1958).
[10] K. FRYER, Note on permutations in a finite field, *Proc. Amer. Math. Soc.*, **6** (1955), 1—2.
[11] CH. HERMITE, Sur les fonctions de sept lettres, *C. R. Acad. Sci. Paris.*, **57** (1863), 750—757.
[12] W. NÖBAUER, Bemerkungen über die Darstellung von Abbildungen durch Polynome und Rationale Funktionen, *Monatshefte für Math.*, **68** (1964), 138—142.
[13] L. RÉDEI and T. SZELE, Algebraisch-zahlentheoretische Betrachtungen über Ringe. I, *Acta Math.* **79** (1947), 291—320.

(Received February 24, 1966)

Objektentheoretische Untersuchungen über die kovarianten Ableitungen in allgemeinen Linienelementräumen

Von A. MOÓR in Szeged*)

§ 1. Einleitung

Ein allgemeiner Linienelementraum $\mathfrak{M}_n$ ist eine Mannigfaltigkeit der Grundelemente (x^i, v^i), $(i=1, 2, \cdots, n)$, die dem Transformationsgesetz:

$$(1.1)\qquad \begin{cases} \hat{x}^i = \hat{x}^i(x^1, x^2, \cdots, x^n) \\ \hat{v}^i = \hat{v}^i(\bar{v}^1, \bar{v}^2, \cdots, \bar{v}^n), \qquad \bar{v}^j = \dfrac{\partial \hat{x}^j}{\partial x^r} v^r \end{cases}$$

genügen, wo die Funktionen $\hat{v}^i(\bar{v})$ homogen von erster Ordnung in den $\bar{v}^i$ sind, ferner

$$(1.1a)\qquad \operatorname{Det}\left(\frac{\partial \hat{x}^i}{\partial x^k}\right) \neq 0, \quad \operatorname{Det}\left(\frac{\partial \hat{v}^i}{\partial \bar{v}^k}\right) \neq 0$$

besteht. Die Bedingungen (1.1a) sichern die Existenz der inversen Transformationen. Im $\mathfrak{M}_n$ soll noch eine Übertragungstheorie der erweiterten kontra- und kovarianten Vektoren existieren [3]. Diese erweiterten Vektoren genügen bezüglich der Transformationen (1.1) dem Transformationsgesetz:

$$(1.2a)\qquad \hat{X}^i = \frac{\partial \hat{v}^i}{\partial v^r} X^r \quad \text{bzw.} \quad (1.2b)\qquad \hat{Y}_i = \frac{\partial v^s}{\partial \hat{v}^i} Y_s.$$

Die Übertragungstheorie haben wir in unserer Arbeit [3] entwickelt, deren Resultate wir im folgenden weitgehend benützen werden. Die wichtigsten Transformationsformeln stellen wir zusammen.

Nach den Bezeichungen

$$\hat{q}^i_j \overset{\text{def}}{=} \frac{\partial \hat{x}^i}{\partial x^j} \equiv \frac{\partial \bar{v}^i}{\partial v^j}, \qquad q^i_j \overset{\text{def}}{=} \frac{\partial x^i}{\partial \hat{x}^j},$$

$$\hat{p}^i_j \overset{\text{def}}{=} \frac{\partial \hat{v}^i}{\partial \bar{v}^j}, \quad p^i_j \overset{\text{def}}{=} \frac{\partial \bar{v}^i}{\partial \hat{v}^j},$$

*) Jetzt in Sopron (Ungarn).

folgt auf Grund von (1. 1), daß die $\hat{q}^i_j$ bzw. q^i_j allein von den x^i bzw. $\hat{x}^i$ abhängig sind, ferner die Relationen

$$(1.3) \qquad \hat{q}^i_r q^r_j = \hat{q}^r_j q^i_r = \delta^i_j,$$

$$(1.4) \qquad \hat{p}^i_r p^r_j = \hat{p}^r_j p^i_r = \delta^i_j$$

bestehen. Die Transformationsformeln (1. 2a) und (1. 2b) können somit in der Form:

$$(1.5a) \qquad \hat{X}^i = \hat{p}^i_s \hat{q}^s_r X^r, \qquad (1.5b) \quad \hat{Y}_i = q^s_i p^t_i Y_s$$

geschrieben werden (vgl. für die inversen Trasformationen von (1. 1) die Formeln (2. 4) von [3]). Die Transformationsformeln der Übertragungsparameter sind die folgenden (vgl. die Formeln (3. 8) und (4. 6) von [3]):

$$(1.6) \qquad \hat{M}_j{}^i{}_k = M_a{}^b{}_c p^t_j q^a_t \hat{p}^i_r \hat{q}^r_b p^s_k q^c_s - \hat{p}^i_{ts} p^t_j p^s_k,$$

$$(1.6^*) \qquad \hat{L}^{*i}_j{}_k = L^{*b}_a{}_c p^r_j q^a_r \hat{p}^i_s \hat{q}^s_b q^c_k + \hat{p}^i_{bc} p^b_j \hat{q}^c_t q^r_k L^{*t}_o{}_r + \hat{p}^i_{bc} \hat{q}^c_r p^b_j q^r_{tk} \bar{v}^t + \hat{p}^i_r \hat{q}^r_t q^t_{sk} p^s_j,$$

wo der Index „o" die Kontraktion mit v^i, ferner

$$\hat{p}^i_{ts} \overset{\text{def}}{=} \frac{\partial^2 \hat{v}^i}{\partial \bar{v}^t \partial \bar{v}^s}, \qquad q^t_{sk} \overset{\text{def}}{=} \frac{\partial^2 x^t}{\partial \hat{x}^s \partial \hat{x}^k}$$

bezeichnen. Bei der Herleitung dieser Formeln haben wir außer den zitierten Formeln von [3] die Relationen

$$v^i = q^i_s \bar{v}^s, \qquad \frac{\partial \hat{v}^i}{\partial v^j} = \hat{p}^i_r \hat{q}^r_j, \qquad \frac{\partial v^j}{\partial \hat{v}^k} = q^j_s p^s_k$$

ferner (1. 3) und (1. 4) beachtet.

Es gelten noch für $M_j{}^i{}_k$ die Relationen

$$M_o{}^i{}_k \equiv M_k{}^i{}_o \equiv 0.$$

Nach diesen Vorbereitungen können wir unsere Resultate kurz formulieren. Die kovariante Ableitung der Vektoren im Sinne der Theorie der geometrischen Objekte ist ein Tensor zweiter Stufe, der aus dem Vektor, aus seiner partiellen Ableitungen und aus einem Hilfsobjekt gebildet ist (vgl. [1], Kapitel IV). In unserer Arbeit [3] haben wir zwei kovariante Ableitungen im $\mathfrak{M}_n$-Raum bestimmt, undzwar:

$$\overset{*}{\nabla}_k X^i \overset{\text{def}}{=} \frac{\partial X^i}{\partial v^k} + M_j{}^i{}_k X^j, \qquad \nabla_k X^i \overset{\text{def}}{=} \frac{\partial X^i}{\partial x^k} - \frac{\partial X^i}{\partial v^j} L^{*j}_o{}_k + L^{*i}_j{}_k X^j,$$

bzw. für die erweiterten kovarianten Vektoren

$$\overset{*}{\nabla}_k Y_i \overset{\text{def}}{=} \frac{\partial Y_i}{\partial v^k} - M_i{}^j{}_k Y, \qquad \nabla_k Y_i \overset{\text{def}}{=} \frac{\partial Y_i}{\partial x^k} - \frac{\partial Y_i}{\partial v^j} L^{*j}_o{}_k - L^{*j}_i{}_k Y_j.$$

Diese fundamentalen kovarianten Ableitungen genügen den Transformationsformeln ([3], (4. 15)):

(1. 7a) $$\overset{*}{\nabla}_k \hat{X}^i = \hat{p}^i_r \hat{q}^r_s q^t_m p^m_k \overset{*}{\nabla}_t X^s,$$

(1. 7b) $$\nabla_k \hat{X}^i = \hat{p}^i_r \hat{q}^r_s q^t_k \nabla_t X^s$$

(1. 8a) $$\overset{*}{\nabla}_k \hat{Y}_i = q^r_t p^t_i q^m_s p^s_k \overset{*}{\nabla}_m Y_r,$$

(1. 8b) $$\nabla_k \hat{Y}_i = q^r_t p^t_i q^s_k \nabla_s Y_r.$$

Die erste kovariante Ableitung $\overset{*}{\nabla}_k$ ist vom Vektor selbst, von seinen partiellen Ableitungen nach v^k und von dem Hilfsobjekt $M_j{}^i{}_k$ abhängig. Im zweiten kovarianten Ableitung ∇_k kommt noch die partielle Ableitung nach x^k_π vor und statt der Übertragungsparameter $M_j{}^i{}_k$ ist $L^{*i}_{j\ k}$ vorhanden. Die allgemeinsten ersten und zweiten kovarianten Ableitungen, die wir mit ${}_{(1)}\nabla_k$ bzw. ${}_{(2)}\nabla_k$ bezeichnen werden, sind auch von diesen Größen abhängig, aber nach der Theorie der geometrischen Objekte haben sie eine allgemeinere Form, da z. B. Linearität nicht gefordert wird. Die allgemeinen ersten und zweiten kovarianten Ableitungen verallgemeinern somit die fundamentalen kovarianten Ableitungen $\overset{*}{\nabla}_k$ und ∇_k. Wir wollen in diesem Aufsatz eben diese allgemeinere Form bestimmen (vgl. unsere Sätze 1—4.).

Ein $\mathfrak{M}_n$-Raum heißt ein *gewöhnlicher Linienelementraum* (vgl. Aufsatz [2]) falls

(1.9) $$\hat{v}^i \equiv \bar{v}^i = \frac{\partial \hat{x}^i}{\partial x^r} v^r$$

gilt. In diesem Falle ist $\hat{p}^i_j = p^i_j = \delta^i_j$. Da im allgemeinen $\mathfrak{M}_n$-Raum die p^i_j und q^i_j voneinander unabhängig sind, bekommt man für die Bestimmung der allgemeinen kovarianten Ableitungen stärkere Bedingungen als im gewöhnlichen — d.h. durch (1. 9) charakterisierten — Linienelementraum. Das kommt besonders in der Formel der zweiten kovarianten Ableitung ${}_{(2)}\nabla_k X^i$ zum Ausdruck (vgl. [3], § 4).

§ 2. Die allgemeine erste kovariante Ableitung

Vor allem geben wir die Definition der allgemeinen ersten kovarianten Ableitung.

Definition 1. Die allgemeine erste kovariante Ableitung eines kontra- bzw. kovarianten erweiterten Vektors $\vec{Z}$ ist ein erweiterter Tensor zweiter Stufe: ${}_{(1)}\nabla\vec{Z}$, der aus $\partial_{v^i}\vec{Z}, \vec{Z}$ und aus einem Hilfsobjekt M mit dem Transformationsgesetz (1. 6) gebildet ist, und dessen kovariante Stufenzahl um eins größer als die von $\vec{Z}$ ist (vgl. [3], § 2).

Im folgenden werden wir immer annehmen, daß die vorhandenen Funktionen in ihren Veränderlichen stetig sind. Unter dieser Annahme gilt der folgende

Satz 1. *Wenn* $M^i{}_{[jk]}=0$, *so ist die allgemeine erste kovariante Ableitung eines erweiterten kontravarianten Vektors* X^i *vom Transformationsgesetz* (1. 2a) *eine Funktion von* $\overset{*}{\nabla}_k X^i$.

Beweis. Bezeichnen wir die allgemeine erste kovariante Ableitung des erweiterten Vektors X^i mit ${}_{(1)}\nabla_k X^i$, so gilt:

$$ {}_{(1)}\nabla_k \hat{X}^i = \hat{p}^i_r \hat{q}^r_s q^m_t p^t_k \, {}_{(1)}\nabla_m X^s; \tag{2. 1}$$

da die allgemeine erste kovariante Ableitung von X^i ein gemischter erweiterter Tensor ist. Nach der Definition 1. ist ${}_{(1)}\nabla_k X^i$ eine Funktion von $\partial_{v^l} X^j$, X^j und $M_j{}^i{}_k$, d.h.:

$$ {}_{(1)}\nabla_k X^i = F^i_k\left(\frac{\partial X^j}{\partial v^l}, X^j, M_j{}^l{}_m\right). $$

Auf Grund von (2. 1) bekommt man für die Funktion F^i_k das Funktionalgleichungssystem:

$$ \begin{aligned} &F^i_k\left(\hat{p}^j_a \hat{q}^a_r p^b_l q^s_b \frac{\partial X^r}{\partial v^s} + \hat{p}^j_{ab} p^b_l \hat{q}^a_c X^c, \quad \hat{p}^j_a \hat{q}^a_r X^r, \quad M_a{}^b{}_c p^t_j q^a_t \hat{p}^l_r \hat{q}^r_b p^s_m q^c_s - \hat{p}^l_{ab} p^a_j p^b_m\right) = \\ &= \hat{p}^i_r \hat{q}^r_s q^m_t p^t_k F^s_m\left(\frac{\partial X^j}{\partial v^l}, \ X^j, \ M_l{}^j{}_h\right), \quad \text{wo} \quad \hat{p}^j_{ab} = \frac{\partial^2 \hat{v}^j}{\partial \bar{v}^a \partial \bar{v}^b}. \end{aligned} \tag{2. 2}$$

Bestimmen wir jetzt die Koordinatentransformation (1. 1) in der Form:

$$ \hat{p}^i_j = \varrho \delta^i_j, \quad \hat{q}^i_j = \delta^i_j, \quad \hat{p}_a{}^l{}_b = \varrho M_a{}^l{}_b, \quad \varrho = konst. \tag{2. 3}$$

so wird nach (1. 3) und (1. 4)

$$ p^i_j = \frac{1}{\varrho}\delta^i_j, \quad q^i_j = \delta^i_j $$

und aus (2. 2) bekommt man:

$$ F^i_k\left(\frac{\partial X^j}{\partial v^l}, \ X^j, \ M_l{}^j{}_h\right) = F^i_k(\overset{*}{\nabla}_l X^j, \ \varrho X^j, \ 0). $$

F^i_k ist also von $M_l{}^j{}_k$ nur durch $\overset{*}{\nabla}_l X^j$ abhängig; wenn jetzt ϱ nach Null strebt, so folgt auf Grund der Stetigkeit, daß F^i_k nur von $\overset{*}{\nabla}_l X^j$ abhängig ist, w.z.b.w.

Jetzt gehen wir zur Bestimmung der allgemeinen ersten kovarianten Ableitung der erweiterten kovarianten Vektoren über. Es gilt der folgende

Satz 2. *Wenn* $M_{[j}{}^i{}_{k]}=0$, *so ist die allgemeine erste kovariante Ableitung eines erweiterten kovarianten Vektors* Y_i *der dem Transformationsgesetz* (1. 2b) *genügt, von* Y_i *und* $\overset{*}{\nabla}_k Y_i$ *abhängig.*

Beweis. ${}_{(1)}\nabla_k Y_i$ sei die allgemeine erste kovariante Ableitung von Y_i; die Transformationsformel dieser Größe ist nach unserer Definition 1:

$$(2.4)\qquad {}_{(1)}\nabla_k \hat{Y}_i = q_i^s p_i^t q_m^r p_k^m {}_{(1)}\nabla_r Y_s.$$

${}_{(1)}\nabla_k Y_i$ hat ferner die Form:

$${}_{(1)}\nabla_k Y_i = F_{ik}\left(\frac{\partial Y_j}{\partial v^l}, Y_j, M_j{}^l{}_h\right).$$

Auf Grund von (2.4) bekommt man für die Funktionen F_{ik} die Formeln:

$$(2.5)\qquad \begin{aligned} &F_{ik}\left(q_i^s p_j^t q_m^r p_l^m \frac{\partial Y_s}{\partial v^r} + q_i^s p_{jl}^t Y_s,\ q_i^s p_j^t Y_s,\ M_a{}^b{}_c p_j^t q_b^a \hat{p}_r^l \hat{q}_b^r p_m^s q_s^c - p_{ab}^l p_j^a p_m^b\right) = \\ &= q_i^s p_i^t q_h^r p_k^h F_{sr}\left(\frac{\partial Y_j}{\partial v^m}, Y_j, M_j{}^l{}_m\right), \quad \text{wo} \quad p_{jl}^t = \frac{\partial^2 \bar{v}^t}{\partial \hat{v}^j \partial \hat{v}^l}. \end{aligned}$$

Nach (1.4) ist aber

$$p_j^r \hat{p}_{rk}^i = -\hat{p}_r^i p_{jt}^r \hat{p}_k^t,$$

wie man das nach einer partiellen Ableitung beider Seiten von (1.4) nach $\bar{v}^k$ unmittelbar verifizieren kann, und somit wird aus (2.5) nach der Substitution (2.3) wenn noch in (2.3) $\varrho = 1$ gesetzt wird:

$$F_{ik}\left(\frac{\partial Y_j}{\partial v^l}, Y_j, M_j{}^l{}_m\right) = F_{ik}(\overset{*}{\nabla}_l Y_j, Y_j, 0)$$

und das beweist den Satz 2.

An zwei Beispielen, die unten angegeben sind, sieht man, daß außer $\overset{*}{\nabla}_k Y^i$ bzw. $\overset{*}{\nabla}_k Y_i$ auch andere Formen von ${}_{(1)}\nabla_k X^i$ bzw. ${}_{(1)}\nabla_k Y_i$ möglich sind:

$$\text{I.}\qquad {}_{(1)}\nabla_k X^i = \overset{*}{\nabla}_s X^i \overset{*}{\nabla}_k X^s,$$

$$\text{II.}\qquad {}_{(1)}\nabla_k Y_i = \overset{*}{\nabla}_k Y_i - Y_i Y_k.$$

Diese genügen der Transformationsformel (2.1) bzw. (2.4), da $\overset{*}{\nabla}_k$ eine tensorielle Operation ist. (Vgl. Satz 3 von [3].)

In den gewöhnlichen Linienelementräumen ist immer $\hat{p}_{jk}^i = 0$ und $\hat{p}_j^i = \delta_j^i$, somit ist ∂_{v^k} eine tensorielle Operation, die der Operation $\overset{*}{\nabla}_k$ entspricht. Auch $M_j{}^i{}_k$ ist in diesen Räumen ein Tensor, und somit bestimmt $M_j{}^i{}_k$ in den gewöhnlichen Linienelementräumen kein Hilfsobjekt.

Zum Schluß dieses Paragraphen wollen wir noch bemerken, daß zwar $\partial_{v[j} Y_{i]}$ selbst als eine erste kovariante Ableitung betrachtet werden kann, doch wegen der vorausgesetzten Symmetrie von $M_j{}^i{}_k$ in j, k ist

$$\partial_{v[j} Y_{i]} = \overset{*}{\nabla}_{[j} Y_{i]}$$

und somit gibt $\partial_{v[j} Y_{i]}$ keine neue Möglichkeiten für ${}_{(1)}\nabla_k Y_i$.

§ 3. Die allgemeine zweite kovariante Ableitung

In diesem Paragraphen werden wir annehmen, daß in dem Raum die folgende Forderung erfüllt ist:

Forderung 1. Es soll ein Koordinatensystem für die Elemente (x^i, v^i) existieren, in dem $L_j^{*i}{}_k$ in j, k symmetrisch ist. [1])

Mit der Forderung 1 ist unser Raum eine unmittelbare Verallgemeinerung der Finslerräume, da in diesen die entsprechenden Übertragungsparemeter $\Gamma_j^{*i}{}_k$ in j, k immer symmetrisch sind und für die gewöhnlichen Linienelementtransformationen, d.h. $\hat{v}^i = \bar{v}^i$ ist diese Relation invariant. In den $\mathfrak{M}_n$-Räumen gilt aber die Forderung 1 nur in einem System (x^i, v^i).

Definition 2. Die allgemeine zweite kovariante Ableitung ${}_{(2)}\nabla_k$ eines erweiterten kontra- bzw. kovarianten Vektors X^i bzw. Y_i ist ein Pseudotensor mit dem Transformationsgesetz

(3.1a) $\quad {}_{(2)}\nabla_k \hat{X}^i = \hat{p}^i_r \hat{q}^r_s q^t_{k\,(2)}\nabla_t X^s$ $\quad$ bzw. $\quad$ (3.1b) $\quad {}_{(2)}\nabla_k \hat{Y}_i = q^s_i p^t_i q^r_{k\,(2)}\nabla_r Y_s$

und ${}_{(2)}\nabla_k X^i$ bzw. ${}_{(2)}\nabla_k Y_i$ ist eine Funktion des Vektors selbst, die partiellen Ableitungen des Vektors und eines Hilfsobjekts $L_i^{*j}{}_k$ mit dem Transformationsgesetz (1.6*).

Jetzt gehen wir zur Bestimmung der Form der zweiten kovarianten Ableitung über. Es gilt der folgende

Satz 3. *Die zweite kovariante Ableitung eines erweiterten kontravarianten Vektors X^i ist bis auf einen skalaren Faktor mit $\nabla_k X^i$ identisch.*

Beweis. Die zweite kovariante Ableitung von X^i ist nach der Definition 2 von der Form

$$ {}_{(2)}\nabla_k X^i = F^i_k \left(X^j, \frac{\partial X^j}{\partial x^l}, \frac{\partial X^j}{\partial v^l}, L_l^{*j}{}_m \right), $$

wo F^i_k noch der Gleichung (3.1a) genügt. Setzen wir in (3.1a) $\hat{p}^i_r = \delta^i_r$, so bekommen wir den Fall des Linienelementraumes zurück, und nach Satz 3 von [2] (vgl. [2], Seite 49) folgt nach unserer Schreibweise

$$ {}_{(2)}\nabla_k X^i_k = F^i_k \left(\nabla_l X^j, \frac{\partial X^j}{\partial v^l} \right), $$

d.h. die zweite kovariante Ableitung ist eine Funktion von $\nabla_l X^j$ und $\partial_{v^l} X^j$. Im $\mathfrak{M}_n$-Raum muß aber (3.1a) erfüllt sein, und das gibt für F^i_k das Funktionalgleichungssystem

$$ (3.2) \qquad F^i_k \left(\hat{p}^j_a \hat{q}^a_b q^c_l \nabla_c X^b, \hat{p}^j_a \hat{q}^a_r p^b_l q^s_b \frac{\partial X^r}{\partial v^s} + \hat{p}^j_{ab} p^b_l \hat{q}^a_c X^c \right) = \hat{p}^i_r \hat{q}^r_s q^t_k F^s_t \left(\nabla_l X^j, \frac{\partial X^j}{\partial v^l} \right). $$

---

[1]) Diese Voraussetzung wird im Satz 3 [2] ausgenützt.

Wählen wir jetzt $\hat{p}^j_{ab}$ so, daß

$$\hat{p}^j_{ab} X^a = -\frac{\partial X^j}{\partial v^b}, \quad \text{ferner} \quad \hat{p}^j_a = \hat{q}^j_a = p^j_a = q^j_a = \delta^j_a,$$

so folgt aus (3. 2):

$$F^i_k\left(\nabla_l X^j, \frac{\partial X^j}{\partial v^l}\right) = F^i_k(\nabla_l X^j, 0).$$

${}_{(2)}\nabla_k X^i$ ist also nur von $\nabla_k X^i$ abhängig:

(3. 3) $$ {}_{(2)}\nabla_k X^i = f^i_k(\nabla_l X^j), \quad f^i_k(\nabla_l X^j) \overset{\text{def}}{=} F^i_k(\nabla_l X^j, 0).$$

Das Funktionalgleichungssystem (3. 2) geht somit in

(3. 4) $$f^i_k(\hat{p}^j_a \hat{q}^a_b q^c_l \nabla_c X^b) = \hat{p}^i_r \hat{q}^r_s q^t_k f^s_t(\nabla_l X^j)$$

über. Wählen wir jetzt:

$$\hat{q}^r_s = \delta^r_s, \quad q^r_s = \delta^r_s,$$

so folgt aus (3. 4)

(3. 5) $$f^i_k(\hat{p}^j_b \nabla_l X^b) = \hat{p}^i_r f^r_k(\nabla_l X^j).$$

Wenn wir jetzt $\nabla_l X^j$ fix halten, während $\hat{p}^j_m$ veränderlich ist, z.B. $\nabla_l X^j = \delta^j_l$, so folgt aus (3. 5):

(3. 6) $$f^i_k(\hat{p}^j_l) = \hat{p}^i_r a^r_k, \quad a^r_k \overset{\text{def}}{=} f^r_k(\delta^j_l) = konst.$$

f^i_k ist also eine homogen-lineare Funktion. Aus (3. 4) wird nach (3. 6), wo die Veränderlichen jetzt die $\hat{p}^i_r$ sind:

$$a^r_k \hat{p}^i_s \hat{q}^s_t q^m_r \nabla_m X^t = \hat{p}^i_r \hat{q}^r_s q^t_k a^m_t \nabla_m X^s.$$

Für $\hat{p}^i_h = \delta^i_h$ und nach einer Kontraktion mit q^j_i wird nach den Formeln (1. 3):

$$q^m_r(a^r_k \nabla_m X^j - \delta^r_k a^s_m \nabla_s X^j) \equiv 0.$$

Das gilt aber dann nur dann für jedes zulässige q^m_r [det $q^m_r \neq 0$], falls

$$a^r_k \nabla_m X^j = \delta^r_k a^s_m \nabla_s X^j$$

besteht. Eine Verjüngung bezüglich r und k gibt unmittelbar

(3. 7) $$\nabla_m X^j = \text{Skalar } a^s_m \nabla_s X^j.$$

Im allgemeinen ist aber

$$\text{Det}(\nabla_s X^j) \neq 0,$$

somit existiert die Lösung des Gleichungssystems

$$(\nabla_k X^j)\Phi^l_j = \delta^l_k.$$

Aus (3. 7) folgt dann nach einer Überschiebung mit Φ^l_j, daß a^s_m bis auf einen Skalaren Faktor eben δ^s_m ist. Es folgt aber daraus auf Grund von (3. 6) und (3. 3):

$$ {}_{(2)}\nabla_k X^i = \text{Skalar } \nabla_k X^i, $$

w.z.b.w.

Bemerkung. Der Skalar in (3. 7) braucht nicht eine Konstante zu sein, da bei der Herleitung von (3. 6) im allgemeinen kann $\nabla_l X^i_{|}$ einen festen, aber von (x^i, v^i) abhängigen Wert haben.

Die Form der zweiten kovarianten Ableitung der erweiterten kovarianten Vektoren mit dem Transformationsgesetz (1. 2b) ist durch den folgenden Satz festgelegt:

Satz 4. *Die zweite kovariante Ableitung eines erweiterten kovarianten Vektors Y_i ist eine Funktion von Y_i selbst, ferner ist sie von $\nabla_k Y_i$ und $\partial_{v[j} Y_{i]}$ abhängig.*

Beweis. Nehmen wir in den Grundtransformationen (1. 1) $\hat{v}^i = \bar{v}^i$, so reduziert sich unser Raum auf einen gewöhnlichen Linienelementraum. Nach Satz 4 von [2] ist dann die zweite kovariante Ableitung von Y_i, $\nabla_k Y_i$ und $\partial_{\dot{v}^k} Y_i$ abhängig. — Im Satz 4 von [2] ist noch die explizite Abhängigkeit von v^i ausgedrückt; wir haben aber in unseren jetzigen Definitionen 1 und 2 die explizite Abhängigkeit von v^i nicht betont, da v^i schießlich Grundelement ist. — Beachten wir noch, daß

$$ \partial_v Y_i \equiv \partial_{v(k} Y_{i)} + \partial_{v[k} Y_{i]}, $$

so können wir annehmen, daß die zweite kovariante Ableitung von Y_i die Form

$$ {}_{(2)}\nabla_k Y_i = F_{ik}(Y_j, \nabla_l Y_j, \partial_{v(j} Y_{l)}, \partial_{v[j} Y_{l]}) \tag{3. 8} $$

hat.

In der Formel (3. 8) sind Y_j, $\nabla_l Y_j$ und $\partial_{v[j} Y_{l]}$ der Reihe nach erweiterter Vektor, Pseudotensor und erweiterter rein kovarianter Tensor, wie das aus (1. 5b) und (1. 6*) leicht bewiesen werden kann. Hingegen hat $\partial_{v(k} Y_{l)}$ die Transformationsformel:

$$ \partial_{v(l} \hat{Y}_{j)} = q^s_t p^t_j q^r_m p^m_l \partial_{v(r} Y_{s)} + q^s_t p^t_{jl} Y_s. \tag{3. 9} $$

Setzen wir nun ${}_{(2)}\nabla_k Y_i$ aus (3. 8) in (3. 1b) ein, so bekommt man für F_{ik} ein Funktionalgleichungssystem. Wählt man jetzt:

$$ q^s_t = p^s_t = \delta^s_t, \quad p^s_{jl} Y_s = -\partial_{v(j} Y_{l)}, $$

so erhält man:

$$ F_{ik}(Y_j, \nabla_l Y_j, \partial_{v(j} Y_{l)}, \partial_{v[j} Y_{l]}) = F_{ik}(Y_j, \nabla_l Y_j, 0, \partial_{v[j} Y_{l]}). $$

Nach (3. 8) drückt das eben den Satz 4 aus.

Weitere Elimination der im Satz 4 angegebenen Größen aus der Formel der allgemeinen zweiten kovarianten Ableitung der erweiterten kovarianten Vektoren ist nicht möglich, wie das aus dem folgenden Beispiel folgt: Es sei $H^{rt}_{\approx}$ die Lösung des Gleichungssystems

$$\partial_{v[s} Y_{r]} H^{rt}_{i} = \delta^{t}_{s}.$$

Offenbar ist H^{rt} ein allgemeiner schiefsymmetrischer Tensor zweiter Stufe mit dem Transformationsformel:

$$\hat{H}^{rt} = \hat{p}^{r}_{a} \hat{q}^{a}_{s} \hat{p}^{t}_{b} \hat{q}^{b}_{m} H^{sm}.$$

Es ist nun

$${}_{(2)}\nabla_k Y_i = \nabla_k Y_t H^{ts} Y_s Y_i$$

eine entsprechende Formel für ${}_{(2)}\nabla_k Y_i$, die der Gleichung (3. 1b) genügt.

§ 4. Schlußbemerkungen

Unsere Sätze 1—4 drücken aus, daß die erste und zweite kovariante Ableitungen aus gewissen fundamentalen Größen gebildet sind. Diese fundamentalen Größen sind die fundamentalen kovarianten Ableitungen $\overset{*}{\nabla}_k$ und ∇_k ferner $\partial_{v[j} Y_{i]}$. Selbstverständlich sind die allgemeinen kovarianten Ableitungen ${}_{(1)}\nabla_k$ und ${}_{(2)}\nabla_k$ nicht beliebige Funktionen der genannten Größen, sondern sie müssen den Relationen (2. 1), (2. 4), (3. 1a) und (3. 1b) genügen.

Wenn wir in den Grundtransformationen (1. 1) für $\hat{v}^i(\bar{v})$ nur in $\bar{v}^j$ lineare Funktionen zulassen, dann ist nach (1. 2) $M_j{}^i{}_k$ ein erweiterter Tensor und nach (1. 2) ist ∂_{v^k} eine tensorielle Operation. Statt $\overset{*}{\nabla}_k$ kann jetzt ∂_{v^k} als fundamentale erste kovariante Ableitung genommen werden und falls $M_j{}^i{}_k \neq 0$ ist, können sehr verschiedenartige Funktionen für ${}_{(1)}\nabla_k X^i$ bestimmt werden. Z.B.:

$${}_{(1)}\nabla_k X^i = \frac{\partial X^i}{\partial v^r} M_s{}^r{}_k X^s.$$

Ist $M_j{}^i{}_k \equiv 0$, so übernimmt $\partial_{v^k} X^i$ die Rolle von $\overset{*}{\nabla}_k X^i$.

Die kovariante Ableitung ${}_{(1)}\nabla_k$ ist sehr ähnlich zur allgemeinen kovarianten Ableitung in den Punkträumen, wie das durch die Vergleichung der Sätze 1 und 2 mit den Sätzen 3 und 4 der Arbeit [4] unmittelbar verifiziert werden kann. Die zweite kovariante Ableitung zeigt hingegen wesentliche Unterschiede im Vergleich zu der kovarianten Ableitung der gewöhnlichen Linienelementräume. Das kommt besonders im Satz 3 zum Ausdruck, wenn wir diesen Satz mit dem Satz 3 von [2] vergleichen. Besonders in diesem Satz zeigt sich die einschränkende Wirkung der allgemeinen Transformationen (1. 1).

Zuletzt wollen wir noch darauf hinweisen, daß die Symmetrie von $M_j{}^i{}_k$ in j, k nicht eine wesentliche Bedingung ist. Hätten wir diese Bedingung nicht gestellt, so hätten wir etwas allgemeinere Type für die erste kovariante Ableitung bekommen.

Literatur

[1] J. Aczél und S. Gołąb, *Funktionalgleichungen der Theorie der geometrischen Objekte* (Warszawa, 1960).
[2] A. Moór, Untersuchungen über die kovariante Ableitung in Linienelementräumen, *Publ. Math. Debrecen*, 7 (1960), 41—53.
[3] A. Moór, Übertragungstheorie bezüglich der allgemeinen Linienelementtransformationen, *Publ. Math. Debrecen*, **13** (1966), 263—287.
[4] A. Moór, Über die kovariante Ableitung der Vektoren, *Acta Sci. Math.*, **19** (1958), 237—246.

(Eingegangen am 5. September 1967)

On mixing sequences of σ-algebras

By J. MOGYORÓDI in Budapest

1. Let $\{\Omega, \mathscr{A}, P\}$ be a probability space, i.e. let $\mathscr{A}$ be a σ-algebra of some subsets of the basic set Ω and for $A \in \mathscr{A}$ let $P(A)$ denote the probability measure of A. Random variables are defined as functions on Ω, which are measurable with respect to the σ-algebra $\mathscr{A}$. The elements of $\mathscr{A}$ will be called events. $\bar{A}$ denotes the event consisting in the non-occurrence of the event A. The elements of Ω will be denoted by ω. If A is an event then χ_A denotes the indicator of A, i.e. $\chi_A = 1$, if $\omega \in A$ and $\chi_A = 0$, if $\omega \notin A$. Let $\mathscr{A}' \subset \mathscr{A}$ be another σ-algebra of some subsets of Ω. We denote by $M(\xi|\mathscr{A}')$ the conditional expectation of the random variable ξ, i.e. such a random variable, which is measurable with respect to $\mathscr{A}'$ and for which

$$\int_A M(\xi|\mathscr{A}')\,dP = \int_A \xi\,dP$$

holds, if $A \in \mathscr{A}'$. If, especially, $\xi = \chi_A$ then $M(\chi_A|\mathscr{A}')$ $(A \in \mathscr{A})$ will be called the conditional probability of the event A with respect to $\mathscr{A}'$ and will be denoted by $P(A|\mathscr{A}')$. The simplest properties of the conditional expectation will be used in this paper.

A. RÉNYI ([1]) calls the sequence $\{B_n\}$ of random events *mixing with density d* if for any fixed random event E

$$\lim_{n\to+\infty} P(B_n E) = dP(E) \tag{1}$$

holds, where d is a fixed number, $0<d<1$. It follows that $\lim_{n\to+\infty} P(B_n) = d$. For mixing sequences $\{B_n\}$ with density d the following limit relation also holds: for every fixed random event E

$$\lim_{n\to+\infty} P(E|B_n) = P(E). \tag{1'}$$

A sequence $\{B_n\}$, satisfying the relation (1′), is *mixing* in the sense of the definition of A. RÉNYI if the limit $\lim_{n\to+\infty} P(B_n) = d$ exists and $0<d<1$.

It is obvious that, if the sequence $\{B_n\}$ is mixing with density d, then the sequence $\{\bar{B}_n\}$ is also mixing with density $1-d$. Thus we have for every fixed E

$$\lim_{n\to+\infty} P(E|\bar{B}_n) = P(E). \tag{1''}$$

The facts expressed by (1′) and (1″) can be unified in the following manner: let $\mathcal{F}_n$ denote the class $\{\Omega, B_n, \bar{B}_n, O\}$ of sets, where O is the empty set and Ω is the basic set. It is easily seen that $\mathcal{F}_n$ is a σ-algebra. For every fixed event E we have with probability 1

$$P(E|\mathcal{F}_n) = P(E|B_n)\chi_{B_n} + P(E|\bar{B}_n)(1-\chi_{B_n}),$$

where $P(E|\mathcal{F}_n)$ denotes the conditional probability of E with respect to the σ-algebra $\mathcal{F}_n$, χ_{B_n} and $1-\chi_{B_n}$ are the indicators of the events B_n and $\bar{B}_n$, respectively. From the fact that $\{B_n\}$ is a mixing sequence with density d, it follows that we have with probability 1

$$\lim_{n\to+\infty} P(E|\mathcal{F}_n) = P(E) \tag{2}$$

and so this limit relation holds in probability measure, too.

However, relation (2) implies the mixing property (1) of the sequences $\{B_n\}$ and $\{\bar{B}_n\}$ only if $\lim_{n\to+\infty} P(B_n) = d$ exists and $0 < d < 1$.

Relation (2) suggests a more general formulation of the notion of mixing sequences of events.

Let $\{\mathcal{G}_n\}$ $(n=1, 2, \cdots; \mathcal{G}_n \subset \mathcal{A})$ be a sequence of σ-algebras.

Definition. The sequence $\{\mathcal{G}_n\}$ of σ-algebras is called mixing if for every fixed event E the sequence

$$P(E|\mathcal{G}_n) \qquad (n=1, 2, \cdots) \tag{3}$$

converges in probability to $P(E)$, as $n\to+\infty$.

Examples, showing that the class of the mixing sequences of σ-algebras is not empty, are given e.g. in papers [1], [2], where concrete mixing sequences of events in the sense of the definition of A. Rényi have been studied.

Rosenblatt ([3]) introduced the notion of the strongly mixing and Kolmogorov ([6]) used the notion of the completely regular sequences of σ-algebras. Both notions require more than that of the regular sequences of σ-algebras. A sequence $\mathcal{G}_n \supset \mathcal{G}_{n+1}$ $(n=1, 2, \cdots)$ of σ-algebras is called regular if the σ-algebra $\bigcap_{n=1}^{\infty} \mathcal{G}_n$ is a trivial one. (The trivial σ-algebra is that σ-algebra which contains only sets with measure 0 or 1.) In this note we shall show that the properties of mixing sequences of σ-algebras are close to those of regular sequences.

2. First we shall show the following

Theorem 1. *Let $E\in\mathcal{A}$ be an arbitrary fixed event and let $\{\mathcal{G}_n\}$ be a mixing sequence of σ-algebras. Then*

$$\lim_{n\to\infty} \sup_{B\in\mathcal{G}_n} |P(EB) - P(E)P(B)| = 0. \tag{4}$$

Proof. It can be easily seen that

$$\sup_{B\in\mathcal{G}_n}|P(EB)-P(E)P(B)| = \sup_{B\in\mathcal{G}_n}\left|\int_B (P(E|\mathcal{G}_n)-P(E))\,dP\right| \leqq$$

$$\leqq \sup_{B\in\mathcal{G}_n}\int_B |P(E|\mathcal{G}_n)-P(E)|\,dP \leqq \int_\Omega |P(E|\mathcal{G}_n)-P(E)|\,dP.$$

We have $0\leqq P(E|\mathcal{G}_n)\leqq 1$ with probability 1. Since the sequence $\{\mathcal{G}_n\}$ is mixing we obtain (4) from (3) by LEBESGUE's convergence theorem.

Remark. If, for $n=1, 2, \cdots, \mathcal{G}_n\supset\mathcal{G}_{n+1}$, then relation (4) is the necessary and sufficient condition for $\{\mathcal{G}_n\}$ to be a regular sequence of σ-algebras. (See e.g. [5].)

It is known that the set-theoretical intersection of σ-algebras is also σ-algebra. By the union of σ-algebras we mean that σ-algebra which is generated by the set-theoretical union of these σ-algebras. It is obvious that the union of trivial σ-algebras is their set-theoretical union. We mean by the "limes inferior" of a sequence of σ-algebras the σ-algebra

$$\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\mathcal{G}_k$$

and we denote it by $\liminf_{n\to\infty}\mathcal{G}_n$.

Theorem 2. *Let $\{\mathcal{G}_n\}$ be a mixing sequence of σ-algebras. Then $\liminf_{n\to+\infty}\mathcal{G}_n$ is a trivial σ-algebra.*

Proof. It is enough to show that for $n=1, 2, \cdots$ the σ-algebras $\bigcap_{k=n}^{\infty}\mathcal{G}_k$ are trivial. Applying assertion (4) of Theorem 1 to the event $E\in\bigcap_{k=n}^{\infty}\mathcal{G}_k$ (and writing E instead of B in (4)) we obtain $P(E)=(P(E))^2$. From this our assertion follows.

The following theorem asserts that regular sequences of σ-algebras are mixing.

Theorem 3. *Let $\{\mathcal{G}_n\}$ be a monotonically decreasing sequence of σ-algebras, i.e. for $n=1, 2, \cdots$ let $\mathcal{G}_n\supset\mathcal{G}_{n+1}$. In order that the sequence $\{\mathcal{G}_n\}$ be mixing, it is necessary and sufficient that the sequence $\{\mathcal{G}_n\}$ be regular.*

Proof. If the sequence $\{\mathcal{G}_n\}$ is mixing then, by Theorem 2, $\liminf_{n\to\infty}\mathcal{G}_n=\bigcap_{n=1}^{\infty}\mathcal{G}_n$ is a trivial σ-algebra.

Conversely, if for $n=1, 2, \cdots$ we have $\mathcal{G}_n\supset\mathcal{G}_{n+1}$, then for every fixed event E the sequence

$$P(E|\mathcal{G}_n)\qquad (n=1, 2, \cdots)$$

is a martingale which converges with probability 1 (and, consequently, in probability measure, too) to the random variable

$$P(E|\bigcap_{n=1}^{\infty}\mathcal{G}_n).$$

(See e.g. Doob [4], Section VII, Theorem 4. 3.) Since in our case $\bigcap_{n=1}^{\infty} \mathcal{G}_n$ is a trivial σ-algebra, the last conditional probability is equal to $P(E)$ with probability 1. This proves our assertion.

3. Let $\{\mathcal{G}_n\}$ be a sequence of σ-algebras. Further consider all the random variables defined on Ω and measurable with respect to the σ-algebra $\mathcal{A}$ which are square integrable. If $\mathcal{H}$ denotes the set of these random variables, then for $\xi(\omega) \in \mathcal{H}$ $\eta(\omega) \in \mathcal{H}$ we define the inner product of ξ and η by

$$\int_{\Omega} \xi \cdot \eta \, dP,$$

and we denote it by (ξ, η). The norm of a random variable ξ is defined as $(\xi, \xi)^{\frac{1}{2}}$ and it is denoted by $\|\xi\|$. Then $\mathcal{H}$ is a Hilbert space with the inner product (ξ, η) Let $E \in \mathcal{G}_n$ $(n=1, 2, \cdots)$ and consider the linear combinations of the random variables $\chi_E - P(E)$, further consider also the limits in the mean of these linear combinations. The set of these random variables will be denoted by $\mathcal{H}_1$. Clearly, $\mathcal{H}_1$ is a subspace of the Hilbert space $\mathcal{H}$. Let us also consider that subspace $\mathcal{H}_2$ of $\mathcal{H}$ which is orthogonal to $\mathcal{H}_1$.

We prove now the following assertion which is well known. We prove it only for the sake of completeness.

Lemma *If $\xi \in \mathcal{H}_2$ is a random variable, then we have for every n*

$$P(M(\xi|\mathcal{G}_n) = M(\xi)) = 1.$$

Proof. It is enough to prove this assertion in case that $M(\xi)=0$. Let A be the event $\{\omega: M(\xi|\mathcal{G}_n) > \varepsilon\}$, where ε is an arbitrary fixed positive number. Then, since $M(\xi|\mathcal{G}_n)$ is $\mathcal{G}_n$-measurable, we have $A \in \mathcal{G}_n$. Further

$$\int_A M(\xi|\mathcal{G}_n) \, dP = \int_{\Omega} M(\xi\chi_A|\mathcal{G}_n) \, dP = \int_{\Omega} \xi\chi_A \, dP.$$

Since, by our supposition, ξ and $\chi_A - P(A)$ are orthogonal, we obtain that

$$\int_A M(\xi|\mathcal{G}_n) \, dP = P(A) M(\xi) = 0.$$

On the other hand

$$0 = \int_A M(\xi|\mathcal{G}_n) \, dP \geqq \varepsilon P(A) \geqq 0.$$

This results that $P(A)=0$. Let further B be the event $\{\omega: M(\xi|\mathcal{G}_n) < -\varepsilon\}$ $(\varepsilon > 0)$. Then $B \in \mathcal{G}_n$ and we obtain in such a way as above that

$$\int_B M(\xi|\mathcal{G}_n) \, dP = 0.$$

On the other hand

$$0 = \int_B M(\xi|\mathscr{G}_n)\,dP = -\varepsilon P(B) \leqq 0.$$

So we have $P(B)=0$. Since

$$P(|M(\xi|\mathscr{G}_n)| > \varepsilon) = P(A)+P(B) = 0,$$

and $\varepsilon>0$ was chosen arbitrarily, we proved that

$$M(\xi|\mathscr{G}_n) = 0, \qquad (n = 1, 2, \cdots)$$

with probability 1.

We shall use this Lemma in the following assertion which facilitates to decide whether a sequence $\{\mathscr{G}_n\}$ of σ-algebras is mixing or not.

Theorem 4. *Let $\{\mathscr{G}_n\}$ be a sequence of σ-algebras. A necessary and sufficient condition for $\{\mathscr{G}_n\}$ to be mixing is that for every fixed $E\in\mathscr{G}_k$ $(k=1, 2, \cdots)$ the sequence*

$$P(E|\mathscr{G}_n) \qquad (n = 1, 2, \cdots)$$

of random variables converge in probability to $P(E)$.

Proof. The necessity part of the assertion is obvious. The sufficiency part of the proof can be performed as follows. Let $\varepsilon>0$ be an arbitrary fixed number and let E be an arbitrary event. Then

$$\varepsilon P(|P(E|\mathscr{G}_n)-P(E)| > \varepsilon) \leqq M(|M((\chi_E-P(E))|\mathscr{G}_n)|). \tag{5}$$

Let us decompose $\chi_E-P(E)$ in the form

$$\xi_1+\xi_2$$

where $\xi_1\in\mathscr{H}_1$ and $\xi_2\in\mathscr{H}_2$; $\mathscr{H}_1$ and $\mathscr{H}_2$ being defined as above. Since $M(\chi_E-P(E))=0$, further $M(\xi_1)=0$, one has $M(\xi_2)=0$. So by our Lemma we have for every n

$$M(\xi_2|\mathscr{G}_n) = 0$$

with probability 1. On the other hand

$$M((\chi_E-P(E))|\mathscr{G}_n) = M(\xi_1|\mathscr{G}_n)+M(\xi_2|\mathscr{G}_n) = M(\xi_1|\mathscr{G}_n)$$

with probability 1. So we have

$$M(|M((\chi_E-P(E))|\mathscr{G}_n)|) = M(|M(\xi_1|\mathscr{G}_n)|). \tag{6}$$

ξ_1, being element of $\mathscr{H}_1$, can be approximated in the mean by finite linear combinations of the elements $\chi_A-P(A)$, $(A\in\mathscr{G}_k,\ k=1, 2, \cdots)$. Denote the sequence, approximating ξ_1 in the mean, by $\eta_1, \eta_2, \cdots$ $(\eta_j \in \mathcal{H}_1)$. For every fixed k the sequence

$$M(\eta_k|\mathcal{G}_n) \qquad (n = 1, 2, \cdots)$$

converges, obviously, in probability to 0. Let $\delta > 0$ be arbitrary and let us put k such that $\|\xi_1 - \eta_k\| < \delta$ be satisfied. Then fix k. It is easily seen that η_k is bounded and so is $M(\eta_k|\mathcal{G}_n)$ with probability 1. Now we have

$$(7) \qquad M(|M(\xi_1|\mathcal{G}_n)|) \leqq M(|M((\xi_1 - \eta_k)|\mathcal{G}_n)|) + M(|M(\eta_k|\mathcal{G}_n)|).$$

The second member on the right-hand side of (7) by LEBESGUE'S theorem converges to 0, while the first is smaller than

$$M(|M((\xi_1 - \eta_k)|\mathcal{G}_n)|) \leqq M(M(|\xi_1 - \eta_k||\mathcal{G}_n)) \leqq \|\xi_1 - \eta_k\|.$$

Conferring (7), (6) and (5) we obtain that

$$(8) \qquad \limsup_{n\to\infty} P(|P(E|\mathcal{G}_n) - P(E)| > \varepsilon) \leqq \frac{\delta}{\varepsilon}.$$

Since $\varepsilon > 0$ and $\delta > 0$ vary independently each of other, (8) means the assertion of the theorem.

Theorem 4 gives similar conditions for $\{\mathcal{G}_n\}$ to be mixing as the conditions of the theorem of A. RÉNYI ([1]) for the sequence of events $\{B_n\}$ to be mixing with density $d\,(0 < d < 1)$.

Theorem 5. *Let $\{\mathcal{G}_n\}$ be a mixing sequence of σ-algebras and let z be a random variable having finite mean-value. Then the sequence*

$$M(z \mid \mathcal{G}_n) \qquad (n = 1, 2, \cdots)$$

of random variables converges in probability to $M(z)$.

Proof. The assertion of the theorem is true if z is of the form:

$$\sum_{k=1}^{j} c_k \chi_{E_k},$$

where c_k $(k = 1, 2, \cdots, j)$ is a real number, $E_k \in \mathcal{A}$ $(k = 1, 2, \cdots, j)$ are events such that $E_k \cap E_l = \emptyset$, and $\bigcup_{k=2}^{j} E_k = \Omega$, further χ_{E_k} denotes the indicator of E_k and j is finite positive integer. Since $M(z)$ is finite, the random variable z can be approximated in L^1 norm by the random variables of the mentioned form as close as we please. Let z^* be such a random variable for which

$$M(|z^* - z|) < \varepsilon$$

holds. Then

$$\int_\Omega |M(z|\mathcal{G}_n)-M(z)|\,dP \leqq \int_\Omega |M(z|\mathcal{G}_n)-M(z^*|\mathcal{G}_n)|\,dP+$$

$$+\int_\Omega |M(z^*|\mathcal{G}_n)-M(z^*)|\,dP+\int_\Omega |z^*-z|\,dP \leqq$$

$$\leqq \int_\Omega M(|z-z^*||\mathcal{G}_n)\,dP+\int_\Omega |M(z^*|\mathcal{G}_n)-M(z^*)|dP+M(|z-z^*|).$$

The first and the third terms on the right hand side of this inequality are smaller than ε and the second converges to zero. This proves the theorem.

4. It is interesting to investigate the analogon of Theorem 4 in case of the almost everywhere convergence. Theorem 6 makes this for martingales.

Theorem 6. *Let $\{\mathcal{G}_n\}$ be a sequence of σ-algebras and suppose that for every event E the conditional probabilities $P(E|\mathcal{G}_n)$ $(n=1, 2, \cdots)$ form a martingale. If for every fixed $E\in\mathcal{G}_k$ $(k=1, 2, \cdots)$ we have*

$$P\big(\lim_{n\to+\infty} P(E|\mathcal{G}_n) = P(E)\big) = 1,$$

then the same holds for every event E.

Proof. We have for arbitrary fixed E

$$0\leqq P(E|\mathcal{G}_n)\leqq 1.$$

Thus by the convergence theorem of the martingales (cf. Doob [4], Section VII, Theorem 4. 1.) the limit

$$\lim_{n\to+\infty} P(E|\mathcal{G}_n) = \xi_E(\omega),$$

exists with probability 1, where $\xi_E(\omega)$ is a random variable. We have further $M(\xi_E(\omega))=P(E)$. So it remains to prove that $P(\xi_E(\omega)=P(E))=1$. Let us consider for this purpose $M(|\xi_E(\omega)-P(E)|)$. We have

$$(9)\qquad M(|\xi_E(\omega)-P(E)|)\leqq M(|\xi_E(\omega)-P(E|\mathcal{G}_n)|)+M(|M((\chi_E-P(E))|\mathcal{G}_n)|).$$

By Lebesgue's theorem, the first term on the right hand side converges to zero as $n\to\infty$. For dealing with the second member, let us decompose the random variable $\chi_E-P(E)$ into the form

$$\xi_1+\xi_2,$$

where $\xi_1\in\mathcal{H}_1$ and $\xi_2\in\mathcal{H}_2$; $\mathcal{H}_1$ and $\mathcal{H}_2$ being defined as above. Since $\xi_1\in\mathcal{H}_1$, one has $M(\xi_1)=0$ and so $M(\xi_2)=0$. By our Lemma we have for every n

$$M(\xi_2|\mathcal{G}_n) = 0.$$

with probability 1. On the other hand

$$M((\chi_E - P(E))|\mathcal{G}_n) = M(\xi_1|\mathcal{G}_n) + M(\xi_2|\mathcal{G}_n) = M(\xi_1|\mathcal{G}_n)$$

with probability 1. So

$$M(|M((\chi_E - P(E))|\mathcal{G}_n)|) = M(|M(\xi_1|\mathcal{G}_n)|). \tag{10}$$

ξ_1, being element of $\mathcal{H}_1$, can be approximated in the mean by finite linear combinations of the elements $\chi_A - P(A)$ $(A \in \mathcal{G}_k,\ k = 1, 2, \cdots)$. Denote this sequence, approximating ξ_1 in the mean, by $\eta_1, \eta_2, \cdots$. For every fixed k we have with probability 1

$$\lim_{n \to +\infty} M(\eta_k|\mathcal{G}_n) = 0.$$

Let $\varepsilon > 0$ be arbitrary and let us put k such that $\|\xi_1 - \eta_k\| < \varepsilon$ be satisfied. Then fix k. It is obvious that η_k is bounded and so is $M(\eta_k|\mathcal{G}_n)$. Now we have

$$M(|M(\xi_1|\mathcal{G}_n)|) \leqq M(|M((\xi_1 - \eta_k|\mathcal{G}_n)|) + M(|M(\eta_k|\mathcal{G}_n)|). \tag{11}$$

The second member on the right hand side of (11) converges to 0, while the first is smaller than $\|\xi_1 - \eta_k\|$. Conferring (11), (10), and (9) we see that

$$\begin{aligned} M(|\xi_E(\omega) - P(E)|) &\leqq \liminf_{n\to\infty} M(|\xi_E(\omega) - P(E|\mathcal{G}_n)|) + \\ &+ \liminf_{n\to\infty} M(|M((\xi_1 - \eta_k)|\mathcal{G}_n)|) + \liminf_{n\to\infty} M(|M(\eta_k|\mathcal{G}_n)|) \leqq \varepsilon. \end{aligned} \tag{12}$$

Since $\varepsilon > 0$ was chosen arbitrarily, the inequality (12) means our assertion.

5. By the aid of the mixing sequences of σ-algebras sequences of random events, which are mixing with density $d\,(0 < d < 1)$ can be constructed as follows:

Theorem 7. *Let $\{\mathcal{G}_n\}$ be a mixing sequence of σ-algebras and $\{B_n\}$ a sequence of random events, for which $B_n \in \mathcal{G}_n$, further $\lim_{n\to+\infty} P(B_n) = d$ exists. Then $\{B_n\}$ is a mixing sequence of events with density d.*

Proof. Let E be an arbitrary event. By our supposition the condition of Theorem 1 is satisfied. So we have

$$|P(EB_n) - dP(E)| \leqq \sup_{B \in \mathcal{G}_n} |P(EB) - P(E)P(B)| + P(E)|P(B_n) - d|.$$

Letting $n \to \infty$, the limit of the right hand side will be 0. This proves the theorem.

6. Consider now some consequences of the above results. We say that a sequence $\{\zeta_n\}$ $(n = 1, 2, \cdots)$ of random variables is mixing if the sequence of the corresponding σ-algebras $\mathcal{G}_n$ $(n = 1, 2, \cdots)$ generated by the random variable ζ_n is mixing.

Theorem 8. *If the sequence ζ_n $(n = 1, 2, \cdots)$ of random variables is mixing and η is an arbitrary random variable the ζ_n is asymptotically independent of η. If,* *in addition, ζ_n converges in probability to a random variable ζ, then ζ is constant with probability* 1.

Proof. The first assertion follows immediately from Theorem 1. In fact, x and y being arbitrary real numbers, by (4) we obtain if $n\to\infty$

$$|P(\zeta_n < x, \eta < y) - P(\zeta_n < x)P(\eta < y)| \leqq$$
$$\leqq \sup_{B\in\mathcal{G}_n} |P(B, \eta < y) - P(B)P(\eta < y)| \to 0.$$

From this it follows especially that if ζ_n converges in probability to ζ, then for every real x

$$\lim_{n\to+\infty} P(\zeta_n < x, \zeta < x) = (P(\zeta < x))^2.$$

On the other hand, if $\varepsilon > 0$ is an arbitrary number

$$P(\zeta_n < x, \zeta < x) = P(\zeta_n < x, \zeta < x, |\zeta_n - \zeta| < \varepsilon) + P(\zeta_n < x, \zeta < x, |\zeta_n - \zeta| \geqq \varepsilon).$$

The second member on the right hand side converges to 0, while the first satisfies the inequality

$$P(\zeta_n < x - \varepsilon, |\zeta_n - \zeta| < \varepsilon) \leqq P(\zeta_n < x, \zeta < x\, |\zeta_n - \zeta| < \varepsilon) \leqq P(\zeta_n < x).$$

If x and $x-\varepsilon$ are continuity points of the distribution function of ζ, then the right hand side converges to $P(\zeta < x)$ and the lim inf of the left hand side of the inequality is greater than $P(\zeta < x - \varepsilon)$. Since $\varepsilon > 0$ was chosen arbitrarily, we see that

$$\lim_{n\to\infty} P(\zeta_n < x, \zeta < x) = P(\zeta < x).$$

So we have

$$P(\zeta < x) = (P(\zeta < x))^2,$$

which means that for every real x

$$P(\zeta < x) = 0, \quad \text{or} \quad 1.$$

This proves our assertion.

Theorem 9. *Let $\{\mathcal{G}_n\}$ be a mixing sequence of σ-algebras in the probability space $\{\Omega, \mathcal{A}, P\}$. If Q is another probability measure, defined on $\mathcal{A}$, and it is absolutely continuous with respect to P, then for every event E the sequence*

$$Q(E|\mathcal{G}_n) \qquad (n = 1, 2, \cdots)$$

of conditional Q-probabilities converges in P (and, consequently, in Q)-probability to $Q(E)$, as $n\to\infty$.

Proof. $Q(E|\mathcal{G}_n)$, being conditional Q-probability, is a random variable, which is measurable with respect to $\mathcal{G}_n$ and for every $A \in \mathcal{G}_n$ we have

$$Q(EA) = \int_A Q(E|\mathcal{G}_n)\, dQ = \int_A Q(E|\mathcal{G}_n)\lambda(\omega)\, dP,$$

where $\lambda(\omega)$ is the Radon—Nikodým derivative of Q with respect to P. Now we have

$$Q(EA)=\int_A Q(E|\mathcal{G}_n)\lambda\,dP=\int_A M\big(Q(E|\mathcal{G}_n)\lambda|\mathcal{G}_n\big)\,dP=\int_A Q(E|\mathcal{G}_n)M(\lambda|\mathcal{G}_n)\,dP.$$

On the other hand for every $A\in\mathcal{G}_n$

$$Q(EA)=\int_A \chi_E\lambda\,dP=\int_A M(\chi_E\lambda|\mathcal{G}_n)\,dP.$$

Since the conditional expectation is uniquely determined mod P, we have with probability 1

$$Q(E|\mathcal{G}_n)M(\lambda|\mathcal{G}_n)=M(\chi_E\lambda|\mathcal{G}_n).$$

By Theorem 5 the random variables

$$M(\lambda|\mathcal{G}_n)\quad\text{and}\quad M(\chi_E\lambda|\mathcal{G}_n)$$

converge in probability to 1 and to

$$M(\chi_E\lambda)=\int_E \lambda\,dP=Q(E),$$

respectively. From this and from the preceding equality our theorem follows.

Corollary. *If $A_n\in\mathcal{G}_n$, $\lim_{n\to+\infty}P(A_n)=d$, then under the conditions of Theorem 9 we have for every event E*

$$\lim_{n\to\infty}Q(A_nE)=dQ(E),$$

i.e., if a sequence $\{A_n\}$ is mixing with density d in the probability space $\{\Omega,\mathcal{A},P\}$, then it is mixing with the same density in $\{\Omega,\mathcal{A},Q\}$ provided that Q is absolutely continuous probability measure with respect to P.

Proof. By Theorem 9, $Q(E|\mathcal{G}_n)$ converges in probability to $Q(E)$ and so

$$Q(A_nE)=\int_{A_n} Q(E|\mathcal{G}_n)\,dQ\to Q(E)\lim_{n\to\infty}P(A_n)=dQ(E),$$

as $n\to\infty$, because $\lim_{n\to\infty}Q(A_n)=\lim_{n\to\infty}\int_{A_n}M(\lambda|\mathcal{G}_n)dP=\lim_{n\to\infty}P(A_n)=d.$

As another consequence of Theorem 9 we prove now

Theorem 10. *Let $\{\mathcal{G}_n\}$ be a mixing sequence of σ-algebras in the probability space $\{\Omega,\mathcal{A},P\}$ and $\{\zeta_n\}$ a sequence of random variables such that ζ_n is $\mathcal{G}_n$-measurable $(n=1,2,\cdots)$. Let further Q be a probability measure which is absolutely continuous with respect to P. If*

$$\lim_{n\to+\infty}P(\zeta_n<x)=F(x),$$

where $F(x)$ is a distribution function and the limit relation holds for every fixed x which is a continuity point of $F(x)$, then we have at every continuity point of $F(x)$

$$\lim_{n\to+\infty} Q(\zeta_n < x) = F(x).$$

Remark. Theorem 10 is a generalization of Theorem 3.1 of [5] and of the corresponding theorem of [1], where a similar assertion has been proved for regular sequences of σ-algebras.

Proof. Let x be an arbitrary fixed continuity point of $F(x)$. Then the event $A_n=\{\omega:\zeta_n(\omega)<x\}$ belongs to $\mathcal{G}_n$. So by the Corollary to Theorem 9 (putting Ω instead of E) we obtain our assertion.

References

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

[2] A. Rényi—P. Révész, On mixing sequences of random variables, *Acta Math. Acad. Sci. Hung.*, **9** (1958), 338—343.

[3] M. Rosenblatt, A central limit theorem and a strong mixing condition, *Proc. Nat. Acad. Sci. USA*, **42** (1956), 43.

[4] J. L. Doob, *Stochastic processes* (New York, 1953).

[5] T. Nemetz—Gy. Varga, Határeloszlástételek kiterjesztéséről, *Magyar Tudományos Akadémia III. Osztály Közleményei*, **14** (1964), 416—421.

[6] В. А. Волконский—Ю. А. Розанов Некоторые предельные теоремы для случайных функций, *Теория вероятностей и её применения*, **4** (1959), 186—207.

[7] И. А. Ибрагимов, Некоторые предельные теоремы для стационарных процессов, *Теория вероятностей и её применения*, **7** (1962), 361—392.

(Received September 16, 1967)

On the sum $\sum dd(f(n))$

By I. KÁTAI in Budapest

1.

Let $f(n)$ denote an irreducible polynomial with integer coefficients. We assume that $f(n)>0$ for $n\geqq 1$. Suppose further that $f(n)\not\equiv cn$. Let $d(n)$ denote the number of divisors of n, and $dd(n)$ the number of divisors of $d(n)$. The letters $p, q, p_1, p_2, \cdots$, $q_1, q_2, \cdots$ stand for prime numbers. For the sake of brevity we write $x_1=\log x$, $x_2=\log x_1, \cdots$. We shall prove the following results.

Theorem 1. *If the degree of $f(n)$ is $\leqq 3$, then*

$$\sum_{n\leqq x} dd(f(n)) = cxx_2+O\left(x\sqrt{x_2}\right), \tag{1.1}$$

where c is a positive constant.

Theorem 2. *If the degree of $f(n)$ is $\leqq 2$, then*

$$\sum_{p\leqq x} dd(f(p)) = c' \operatorname{li} x\cdot x_2+O\left(\operatorname{li} x\cdot\sqrt{x_2 x_3}\right), \tag{1.2}$$

where c' is a positive constant.

Remarks. It seems probable that the relations (1.1)—(1.2) hold without any restriction on the degree of $f(n)$. For the proof of (1.1) in the case $r=3$ we use a result of C. HOOLEY concerning the power-free values of polynomials [1]. (This question previously was investigated by P. ERDŐS in [2].) For the proof of (1.2) we use some well-known theorems on the distribution of prime numbers in arithmetical progressions.

2. Notation

The function $U(n)$ is the number of distinct prime factors of n. (a, b) is the highest common factor of a and b. $\varrho(n)$ denotes the number of (incongruent) roots of the congruence $f(\nu)\equiv 0 \pmod n$, and $\lambda(n)$ the number of those roots for which $(\nu, n)=1$. The letter m denotes square-free numbers.

We shall say that K is a "square-full" number if it contains every prime-divisors at least on the second power. Let $\mathfrak{U}$ denote the set of the square-full numbers. It is evident, that every integer n can be represented in the form $n=Km$, where $K\in\mathfrak{U}$, $(m, K)=1$. This representation is unique. We say that K is the square-full part and m is the square-free part of n. Let $\mathscr{B}_K$ denote the set of n's, square-full part of which is K.

Let $\mu(n)$ denote the Möbius-function.

For $K\in\mathfrak{U}$ we introduce the notation:

(2.1) $$k=d(K),\quad k=2^\alpha k_1 (k_1 \text{ is odd}),\quad k_2=d(k),\quad k_3=d(k_1);$$

(2.2) $$a(K)=k_2-U(K)k_3.$$

Thus for $f(n)\in B_K$ we have

(2.3) $$ddf(n)=k_3U(f(n))+a(K).$$

Let $B_K(x)$ (resp. $\bar{B}_K(x)$) the number of n's (resp. p's) in the interval $[1, x]$ for which $f(n)$ (resp. $f(p)$) belongs to $\mathscr{B}_K$. Let $C_l(x, \eta)$ (resp. $\bar{C}_l(x, \eta)$) the number of n's (resp. p's) in the interval $[1, x]$ for which $f(n)\equiv 0 \pmod l$ but $f(n)\not\equiv 0 \pmod{q^2}$ (resp. $f(p)\equiv 0 \pmod l$ but $f(p)\not\equiv 0 \pmod{q^2}$), when $1\leqq q\leqq \xi$ and $q\nmid l$. Let $C_l(x)=C_l(x,\infty)$, $\bar{C}_l(x)=\bar{C}_l(x,\infty)$.

The following relations obviously hold:

(2.4) $$B_K(x)=\sum_{\nu|K}\mu(\nu)C_{K\nu}(x),$$

(2.5) $$\bar{B}_K(x)=\sum_{\nu|K}\mu(\nu)\bar{C}_{K\nu}(x).$$

$\varepsilon_1, \varepsilon_2, \varepsilon_3$ denote sufficiently small positive constants. We use the symbol $\ll$ in VINOGRADOV's sense.

3. Lemmas

Lemma 1. [3] *The following relations hold:* a) $\varrho(ab)=\varrho(a)\varrho(b)$, *if* $(a, b)=1$; b) $\varrho(p^\alpha)\ll\alpha$; c) $\varrho(p^\alpha)=\varrho(p)$, *if* $p\nmid D$ *(D denotes the discriminant of $f(n)$). Further* $\varrho(p^\alpha)=\lambda(p^\alpha)$, *when p is sufficiently large.*

We shall use the following result of P. TURÁN.

Lemma 2. [4]
$$\sum_{n\leqq x}(U(f(n))-x_2)^2\ll xx_2.$$

Combining the method of TURÁN with the Rodossky—Tatuzawa theorems, we can prove the following

Lemma 3.

$$\sum_{p \leqq x} (U(f(p)) - x_2)^2 \ll \frac{x}{x_1} x_2 \log x_2 . \tag{3. 1}$$

Using additionally the result of BOMBIERI in the theory of large sieve [6], we could prove that the left hand side of (3. 1) has the order $x x_1^{-1} x_2$.

Lemma 4. ([8])

$$\sum_{n \leqq x} [d(f(n))]^\alpha \ll x \cdot x_1^{c(\alpha)} \quad \textit{if} \quad \alpha \geqq 1.$$

$c(\alpha)$ is a suitable constant which depends only on α and f.

Corollary.

$$\sum_{\substack{n \leqq x \\ U(f(n)) > \beta x_2}} d(f(n)) \ll \frac{x}{x_1^2}, \text{ if } \beta \text{ is large enough.}$$

Let $N(x, y)$ denote the number of those n's in $1 \leqq n \leqq x$, for which $p^2 | f(n)$ with some $p > y$.

C. HOOLEY proved

Lemma 5. ([1])

$$N(x, x_1) \ll x \cdot x_1^{-A/x_3} \qquad (A > 0, \text{ suitable constant}).$$

Lemma 6. *Let $b_n \ll n_4^\varepsilon$ be a sequence of positive numbers. Then*

$$\sum_{K > y} \frac{b_K}{K} \ll y^{-\frac{1}{2}+\varepsilon} \quad \textit{for} \quad y \to \infty .$$

The proof is simple and so can be omitted.

Applying the sieve method, we can prove the following

Lemma 7.

$$C_h(x, x) = x \frac{\varrho(h)}{h} \prod_{p \nmid h} \left(1 - \frac{\varrho(p^2)}{p^2}\right) + O(x x_1^{-1})$$

uniformly for $1 \leqq h \leqq x_1^2$.

Lemma 8. *Let $f(n)$ be an irreducible polynomial of degree 2. Then for fixed h the number of the solutions of $f(n) = hs^2$ ($1 \leqq n \leqq x$, n, s integers) is at most $O(x_1)$ uniformly in h.*

For the proof see [7], Lemma 2.

Lemma 9.

$$\bar{C}_h(x, x^{1/2}) = \operatorname{li} x \cdot \frac{\lambda(h)}{h} \prod_{p \nmid h} \left(1 - \frac{\lambda(p^2)}{p^2}\right) + O(x x_1^{-2}),$$

uniformly in $1 \leqq h \leqq x_1$.

The proof goes with the standard application of the sieve method using in addition the prime number theorems in the form:

$$\pi(x, k, l) = \frac{\operatorname{li} x}{\varphi(k)} (1 + O(x_1^{-2})), \tag{3.2}$$

uniformly for $1 \leqq k \leqq x_1^3$, $(k, l) = 1$ (see [5], and the Brun—Titchmarsh inequality stating that

$$\pi(x, k, l) < C_\delta \frac{\operatorname{li} x}{\varphi(k)}, \quad \text{for} \quad k < x^{1-\delta} \qquad (\delta > 0) \text{ ([5])}. \tag{3.3}$$

4. The proof of Theorem 1

$$\sum\nolimits_K = \sum_{\substack{n \leqq x \\ f(n) \in B_K}} ddf(n); \quad \sum\nolimits_{K,A} = \sum_{\substack{n \leqq x \\ f(n) \in B_K}} U(f(n)).$$

Using (2. 3) we have

$$\sum\nolimits_K = k_3 \sum\nolimits_{K,A} + a(K) B_K(x).$$

Let $\xi = x_1^\delta$, and let δ be a sufficiently small positive constant. First we prove that

$$\sum_{K > \xi} \sum\nolimits_K \ll x. \tag{4.1}$$

Applying the Corollary to Lemma 4, it is enough to prove that

$$\sum_{K > \xi} (x_2 k_3 + k_2) B_K(x) \ll x.$$

Since $B_K(x) \ll \frac{x \varrho(K)}{K} + \varrho(K)$, by Lemma 6 we obtain

$$\sum_{\xi \leqq K \leqq x} (k_3 x_2 + k_2) B_K(x) \ll x x_2 \sum_{\xi \leqq K \leqq x} \frac{k_3 \varrho(K)}{K} + x \sum_{K \geqq \xi} \frac{k_2 \varrho(K)}{K} \ll x x_2 \xi^{-1/3} \ll x.$$

Let now $K > x$. $K = p_1^{\alpha_1} \cdots p_1^{\alpha_r}$, $p_1 < p_2 < \cdots < p_j \leqq x^{1/4} < p_{j+1} < \cdots < p_r$. Let $K = K_1 K_2$, $K_1 = p_1^{\alpha_1} \cdots p_j^{\alpha_j}$.

Let

$$\sum_{K > x} (x_2 k_3 + k_2) B_K(x) = \sum\nolimits_a + \sum\nolimits_b + \sum\nolimits_c$$

where in the sums $\sum_a$, $\sum_b$, $\sum_c$ we sum over those K for which: a) $K_1 \leqq \xi$; b) $\xi < K_1 \leqq x$; c) $K_1 > x$ holds, respectively.

Since for $K_1 \leqq \xi$ the inequality

$$(k_3 \leqq) k_2 \leqq dd(K) \ll dd(K_1) \ll [d(K_1)]^{\varepsilon} \ll \exp\left(2\varepsilon \frac{\log \xi}{\log\log \xi}\right) \ll \exp\left(\varepsilon_1 \frac{x_2}{x_3}\right)$$

holds, by Lemma 5 we have

$$\sum\nolimits_a \ll x_2 \exp\left(\varepsilon_1 \frac{x_2}{x_3}\right) N(x, x^{1/4}) \ll x x_2 \exp\left(-\frac{A}{2}\frac{x_2}{x_3}\right) \ll x.$$

For $\sum_b$ we have

$$\sum\nolimits_b \ll \sum_{\xi \leqq K_1 \leqq x} (k_3 x_2 + k_2) C_{K_1}(x) \ll x x_2 \sum_{\xi < K_1 < x} \frac{d(K_1)\varrho(K_1)}{K} \ll x x_2 \xi^{-1/3} \ll x.$$

For the estimation of $\sum_c$ let K_3 denote the maximal square-full divisor of K_1 in the interval $x^{1/4} \leqq K_3 \leqq x$. ($K_3$ exists since the greatest prime factor of K_1 is $\leqq x^{1/4}$.) Consequently, we have

$$\sum\nolimits_c \ll x^{1+\varepsilon} \sum_{x^{1/4} < K_3 \leqq x} \frac{\varrho(K)}{K} \ll x.$$

So (4. 1) holds.

Since

$$\sum_{K \leqq \xi} a(K) B_K(x) \ll x \sum_{K \leqq \xi} \frac{K^{\varepsilon}\varrho(K)}{K} \ll x,$$

for the proof of (1. 1) it is enough to prove that

$$\sum_{K \leqq \xi} k_3 \sum\nolimits_{K,A} = c x x_2 + O\left(x\sqrt{x_2}\right).$$

By the Cauchy—Schwarz inequality we have

$$T = \sum_{K \leqq \xi} k_3 \{\sum\nolimits_{K,A} - x_2 B_K(x)\} \ll \sum_{K \leqq \xi} \sum_{f(n) \in \mathcal{B}_K} k_3 |U(f(n)) - x_2| \ll$$

$$\ll \Big(\sum_{K \leqq \xi} k_3^{1/2} B_K(x)\Big)^{1/2} \Big(\sum_{n \leqq x} |U(f(n)) - x_2|^2\Big)^{1/2} = \sum\nolimits_1^{1/2} \cdot \sum\nolimits_2^{1/2}.$$

Since

$$\sum\nolimits_1 \ll x \sum_{K \leqq \xi} k_3^{1/2} \frac{\varrho(K)}{K} \ll x$$

and by Lemma 2 $\sum_2 \ll x x_2$, we have $T \ll x x_2^{1/2}$.

Now we prove that

$$\sum_{K \leqq \xi} k_3 B_K(x) = cx + O\left(x \exp\left(-\frac{A}{2}\frac{x_2}{x_3}\right)\right), \tag{4. 2}$$

hence Theorem 1 follows.

Applying (2. 4) we have

$$\sum_{K \leqq \xi} k_3 B_K(x) = \sum_{K \leqq \xi} k_3 \sum_{\nu | K} \mu(\nu) C_{K\nu}(x) = \sum_{K \leqq \xi} k_3 \sum_{\nu | K} \mu(\nu) C_{K\nu}(x, x) +$$

$$+ O\Big(\sum_{K \leqq \xi} k_3 \sum_{\nu | K} |\mu(\delta)| |C_{K\nu}(x, x) - C_{K\nu}(x)|\Big) = \sum\nolimits_3 + O\big(\sum\nolimits_4\big).$$

Since in the sum $\sum_4$ the relations $d(K) \ll \exp\left(3\delta \frac{x_2}{x_3}\right)$, $k_3 \leqq d^{\varepsilon}(K)$ hold, by Lemma 5

$$\sum\nolimits_4 \ll \exp(4\delta x_2/x_3) N(x, x) \ll x x_1^{-A/2x_3},$$

if δ is small enough.

Further by Lemma 7

$$\sum\nolimits_3 = cx + O(x x_1^{-1} \xi^2) = cx + O(x x_1^{-A/2x_3}),$$

where

$$c = \sum_K \frac{k_3}{K} \left\{ \sum_{\nu | K} \mu(\nu) \frac{\varrho(K\nu)}{\nu} \right\} \prod_{p \nmid K} \left(1 - \frac{\varrho(p^2)}{p^2}\right).$$

5. The proof of Theorem 2

Let

$$S_K = \sum_{\substack{p \leqq x \\ f(p) \in B_K}} ddf(p); \qquad S_{K,A} = \sum_{\substack{p \leqq x \\ f(p) \in B_K}} U(f(p)).$$

By (2. 3)

$$S_K = k_3 \, S_{K,A} + a(K) \bar{B}_K(x).$$

Using the Corollary to Lemma 4, we have

$$\sum_{K > \xi} S_K \ll \sum_{K > \xi} (k_3 x_2 + k_2) \bar{B}_K(x) + O(x/x_1^2) = \sum + O\left(\frac{x}{x_1^2}\right).$$

Let

$$\sum = \sum\nolimits_1 + \sum\nolimits_2 + \sum\nolimits_3 + \sum\nolimits_4,$$

where in $\sum_1$: $\xi \leqq K \leqq x^{3/4}$, in $\sum_2$: $x^{3/4} < K \leqq x$, in $\sum_3$: $x \leqq K \leqq x^{7/4}$, and in $\sum_4$: $K \geqq x^{7/4}$.

For $K \leqq x^{3/4}$ we have by (3. 3) that

$$\bar{B}_K(x) \ll \frac{\lambda(K)}{\varphi(K)} \operatorname{li} x.$$

Consequently

$$\sum\nolimits_1 \ll \operatorname{li} x \sum_{K \geqq \xi} \frac{k_3 x_2 + k_2}{\varphi(K)} \ll x_2 \operatorname{li} x \cdot \xi^{-1/3} \ll \operatorname{li} x.$$

For $x^{3/4} < K \leqq x$ we use the trivial estimation

$$\bar{B}_K(x) \leqq B_K(x) \ll x \frac{\varrho(K)}{K},$$

$$\sum\nolimits_2 \ll x^{1+\varepsilon} \sum_{K \geqq x^{3/4}} \frac{\varrho(K)}{K} \ll \operatorname{li} x.$$

Since for $K \geqq x$

$$B_K(x) \ll \varrho(K) \ll x^{\varepsilon},$$

and the number of the square-full number in the interval $[1, x^{7/4}]$ is majorized by $x^{7/8+\varepsilon}$, so

$$\sum\nolimits_3 \ll \mathrm{li}\, x.$$

Finally, let $K \geqq x^{7/4}$. Let L^2 denote the greatest square divisor of K. Since K is a square-full number, so $L^2 \geqq K^{2/3}$ $(\geqq x^{7/6})$.

It is obvious, that

$$\sum\nolimits_4 \ll x^{\varepsilon} \sum_{K \geqq x^{7/4}} \sum_{\substack{f(n) \equiv 0 (\mathrm{mod}\, K) \\ n \leqq x}} 1 \ll x^{\varepsilon} \sum_{L^2 \geqq x^{7/6}} \sum_{\substack{f(n) = hL^2 \\ n \leqq x}} 1.$$

Since the degree of $f(n)$ is 2, so $h \ll x^{5/6}$. Changing the order of summation and applying Lemma 8, we have

$$\sum\nolimits_4 \ll x^{\varepsilon} \sum_{h \leqq c x^{5/6}} \sum_{\substack{f(n) = hL^2 \\ n \leqq x}} 1 \ll \mathrm{li}\, x.$$

Consequently

$$\sum_K S_K = \sum_{K \leqq \xi} S_K + O(\mathrm{li}\, x).$$

Taking into account that

$$\sum_{K \leqq \xi} |a(K)|\, \bar{B}_K(x) \ll \mathrm{li}\, x \sum_{K \leqq \xi} \frac{|a(K)|}{\varphi(K)} \ll \mathrm{li}\, x,$$

we have

$$\sum_K S_K = \sum_{K \leqq \xi} k_3 S_{K,A} + O(\mathrm{li}\, x).$$

By Lemma 3 we obtain that

$$\Big| \sum_{K \leqq \xi} k_3 S_{K,A} - x_2 \sum_{K \leqq \xi} k_3 \bar{B}_K(x) \Big| \ll \Big(\sum_{K \leqq \xi} k_3^2 \bar{B}_k(x) \Big)^{1/2} \Big(\sum_{p \leqq x} \big(U(f(p)) - x_2\big)^2 \Big)^{1/2} \ll$$

$$\ll (\mathrm{li}\, x)^{1/2} (\mathrm{li}\, x \cdot x_2 \cdot x_3)^{1/2} \ll \mathrm{li}\, x \cdot \sqrt{x_2 x_3}.$$

Consequently for the proof of Theorem 2 it is enough to prove that

$$\sum_{K \leqq \xi} k_3 \bar{B}_K(x) = d\, \mathrm{li}\, x \cdot x_2 + O\big(\mathrm{li}\, x \cdot \sqrt{x_2 x_3}\big). \tag{5.1}$$

The proof of (5. 1) is very similar to that of (4. 2) and so it can be omitted.

References

[1] C. Hooley, On the power free values of polynomials, *Mathematika*, **14** (1967), 21—26.
[2] P. Erdős, Arithmetical properties of polynomials, *J. London Math. Soc.*, **28** (1953), 416—425.
[3] P. Erdős, On the sum $\Sigma df(k)$, *J. London Math. Soc.*, **27** (1952), 7—15.
[4] P. Turán, Über einige Verallgemeinerungen eines Satzes von Hardy und Ramanujan, *J. London Math. Soc.*, **11** (1936), 125—133.
[5] K. Prachar, *Primzahlverteilung* (Berlin, 1957).
[6] E. Bombieri, On the large sieve, *Mathematika*, **12** (1965), 201—225.
[7] F. V. Atkinson and Lord Cherwell, On arithmetical functions, *Quarterly J. Math.* (Oxford), **20** (1949), 65—79.
[8] Van der Corput, Une inegalité au nombre des diviseurs, *Nederl. Wetensch. Proc.*, **42** (1939), 547—553.

(Received May 2, 1967, revised December 30, 1967)

On a classification of primes

By I. KÁTAI in Budapest

1. Let l, D be coprime natural numbers. The letters $p, p_1, \cdots, q, q_1, \cdots$ denote prime numbers. Let $\mathcal{A}_{D,l}$ denote the set of those p for which $q \nmid p+1$, if $q \equiv l \pmod{D}$. Let $N(x, D, l)$ be the number of the elements of $\mathcal{A}_{D,l}$ which are smaller than x. It seems to be interesting to know whether $N(x, D, l) \to \infty$ for $x \to \infty$ or not. Using the variance-method due to YU. V. LINNIK [1], or the method of C. HOOLEY [2] combined with BOMBIERI's large sieve theorem (see [3]), we deduce the inequality

$$N(x, 4, 3) \gg \frac{x}{(\log x)^4}. \tag{1.1}$$

Sharpening the method of HOOLEY we are also able to prove that

$$N(x, D, l) \gg \frac{x}{(\log x)^4}, \tag{1.2}$$

provided that there exists some Dirichlet character $\chi \pmod{D}$ such that $\chi(l) = -1$. For the remaining cases we are unable to prove that $N(x, D, l) \to \infty$ for $x \to \infty$.

Furthermore, by SELBERG's sieve method we obtain

$$N(x, D, l) \ll x/(\log x)^{1-1/\varphi(D)}. \tag{1.3}$$

It seems probable that this is the exact order of $N(x, D, l)$.

We shall give a detailed proof of the inequalities (1.1)—(1.2) in another paper. Here we investigate only the special case $l = D-1$, D prime, and one of its applications.

Let $\varphi(n)$ denote the Euler function, and $\sigma(n)$ the sum of the positive divisors of n. Let $\varphi(n) = \varphi_1(n)$, $\sigma(n) = \sigma_1(n)$, $\varphi_k(n) = \varphi(\varphi_{k-1}(n))$, $\sigma_k(n) = \sigma(\sigma_{k-1}(n))$ for all $k \geqq 2$.

Let D be a fixed odd prime. We say, that the prime number q belongs to the rth class, if $\varphi_r(q) \equiv 0 \pmod{D}$ but $\varphi_k(q) \not\equiv 0 \pmod{D}$, whenever $k < r$. Let $f(D, r, x)$ denote the number of the primes in the rth class smaller than x. Using the prime-number theorem for arithmetical progressions and the eratosthenian sieve, ERDŐS [4] proved that

$$f(D,1,x)=\bigl(1+o(1)\bigr)\frac{x}{(D-1)\log x};\qquad f(D,2,x)=\bigl(1+o(1)\bigr)\frac{D-2}{D-1}\cdot\frac{x}{\log x}.$$

But he has left open the problem whether $f(D, 3, x)\to\infty$ as $x\to\infty$.

We formulate now analogous questions for $\sigma(n)$ instead of $\varphi(n)$.

We say, that the prime number q belongs to the rth class, if $\sigma_r(q)\equiv 0 \pmod D$ but $\sigma_k(q)\not\equiv 0 \pmod D$ whenever $k<r$. Let $g(D, r, x)$ denote the number of the primes in the rth class smaller than x. Using the same method as ERDŐS, it is easy to see that

$$g(D,1,x)=\bigl(1+o(1)\bigr)\frac{1}{D-1}\,\frac{x}{\log x};\qquad g(D,2,x)=\bigl(1+o(1)\bigr)\frac{D-2}{D-1}\,\frac{x}{\log x}.$$

In this paper we shall prove, that $g(D, 3, x)\gg x(\log x)^{-4}$ if $x\to\infty$. The method cannot be applied to the lower estimation of $f(D, 3, x)$.

Theorem. *We have* $$g(D,3,x)\gg\frac{x}{(\log x)^4}.$$

Remark. Sharpening the method we are able to improve this inequality (see [5]).

2. For the proof we need some lemmas.

Lemma 1. (E. BOMBIERI [3])

$$\sum_{D\le Y}\ \max_{\substack{l \,(\mathrm{mod}\, D)\\ (l,D)=1}}\ \max_{z\le x}\left|\pi(z,D,l)-\frac{\operatorname{li} z}{\varphi(D)}\right|\ll\frac{x}{(\log x)^A},$$

where

$$Y=x^{1/2}(\log x)^{-B},\quad B\ge 4A+40,$$

A being an arbitrary constant.

Let $\chi(n)$ denote a character $\bmod D$ such that $\chi(-1)=-1$. Let further

$$r(n)=\sum_{d\mid n}\chi(d)=\prod_{p^\alpha\| n}\{1+\chi(p)+\cdots+\chi(p^\alpha)\}.$$

Let

$$K(x)=\sum_{\substack{q\le x\\ q\not\equiv -1 \,(\mathrm{mod}\, D)}} r(q+1)\,|\mu(q+1)|.$$

Using the method of C. HOOLEY [2] combined with the theorem of BOMBIERI (Lemma 1), we can prove the following

Lemma 2. $K(x)=A_D\operatorname{li}x+O\bigl(\operatorname{li}x\cdot(\log\log x)^{-\alpha}\bigr)$, *where* $\alpha>0$, $A_D\ne 0$ *are suitable constants.*

We shall give a detailed proof of this assertion in another paper.

Lemma 3. *Let $N(k, x)$ denote the number of the couples of primes satisfying the conditions $p+1=kq$, $p \leqq x$. Then*

$$N(k,x) \ll \frac{x}{\varphi(k)\log^2 \frac{x}{k}}.$$

For the proof see PRACHAR [6] p. 51, Theorem 4. 6.

Let

$$M(x,y) = \sum_{n \leqq x}{}' |r(n)|,$$

where the dash means that we sum over those n all prime divisors of which are smaller than y.

Lemma 4. *We have*

$$M(x,y) < x \exp\left(-\frac{\log_3 y}{\log y}\log x + c\log_2 y + O\left(\frac{\log_2 y}{\log_3 y}\right)\right),$$

when $1<y(x)<x$; $y(x)\to\infty$ *as* $x\to\infty$; $c=\sqrt{2}\left(1-\frac{1}{D-1}\right)^{1/2}$.

The proof is similar to the proof of RANKIN's theorem (PRACHAR [6], p. 158) and so we omit it.

Let $f(n)$ be a totally additive arithmetical function defined as follows:

$$f(p) = \begin{cases} 1 & \text{when} \quad y<p<x^{1/3} \quad \text{and} \quad p\equiv -1 \pmod D, \\ 0 & \text{otherwise.}\end{cases}$$

Using BOMBIERI's theorem, we obtain:

Lemma 5. *We have*

$$\sum_{q\leqq x}\{f(q+1)-A_{x,y}\}^2 \ll \operatorname{li} x \cdot A_{x,y},$$

where

$$A_{x,y} = \sum_{\substack{y<p<x^{1/3} \\ p\equiv -1 \pmod D}} \frac{1}{p}.$$

Corollary. *If* $\frac{\log x}{\log y}\to\infty$, *then the number of those q for which $f(q+1)=0$ is at most $o(\operatorname{li} x)$.*

Lemma 6. $\sum_{q\leqq x} |r^2(q+1)| \ll x\log^2 x.$

The proof is simple and can be omitted.

3. Proof of the Theorem

The letters p, $q, Q, p_1, p_2, \cdots, q_1, q_2, \cdots, Q_1, Q_2, \cdots,$ denote prime numbers.

Let $\mathfrak{A}_r$ denote the set of those q which belong to the rth class. It is evident that those q in the sum $K(x)$, for which

$$r(q+1)\,\mu(q+1) \neq 0,$$

are not belonging to the classes $\mathfrak{A}_1, \mathfrak{A}_2$. Indeed, if $r(q+1)\,\mu(q+1) \neq 0$, $q \not\equiv -1 \pmod D$ then

(3.1) $$q+1 = q_1 q_2 \cdots q_r, \ (q_1 < \cdots < q_r) \quad \text{and} \quad \chi(q_i) \neq -1,$$

i.e. $q_i + 1 \not\equiv 0 \pmod D$.

If for a q, represented in the form (3.1), there exists a Q, such that $Q \equiv -1 \pmod D$ and $\sigma(q+1) = (q_1+1)\cdots(q_r+1) \equiv 0 \pmod Q$, but $\sigma(q+1) \not\equiv 0 \pmod{Q^2}$ then $q \in \mathfrak{A}_3$.

Let

(3.2) $$z_0 = (\log x)^5, \ z_1 = z_0^{\log_2 x}, \ z_2 = x^{1/\log_2 x}.$$

Let S_1 denote the set of those q which are represented in the form (3.1), and for which there exists a prime number Q, $Q > z_0$, $Q \equiv -1 \pmod D$ such that

$$\sigma(q+1) \equiv 0 \pmod{Q^2}.$$

Let S_2 denote the set of those q for which

$$\sigma(q+1) \not\equiv 0 \pmod Q,$$

if $Q > z_0$ and $Q \equiv -1 \pmod D$.

Let

(3.3) $$S_i(x) = \sum_{\substack{q \leq x \\ q \in S_i \\ D \nmid q+1}} |r(q+1)|\,|\mu(q+1)| \qquad (i = 1, 2),$$

and let

$$A_3(x) = \sum_{\substack{q \leq x \\ q \in \mathfrak{A}_3}} |r(q+1)|\,|\mu(q+1)|.$$

Obviously

(3.4) $$A_3(x) \geqq |K(x)| - |S_1(x)| - |S_2(x)|.$$

Lemma 7. *We have*

(3.5) $$S_1(x) = o(\mathrm{li}\, x),$$

(3.6) $$S_2(x) = o(\mathrm{li}\, x).$$

Proof. Since $|r(m)| \leqq d(m)$, where $d(m)$ denotes the number of divisors of m, so we have $S_1(x) \ll \sum_1 + \sum_2$ with

$$\sum_1 = \sum_{\substack{z_0 < Q \leqq x \\ Q \equiv -1 (\mathrm{mod}\ D)}} \sum_{\substack{q_1 \equiv q_2 \equiv -1 (\mathrm{mod}\ Q) \\ q_1, q_2 \leqq x}} d(q_1 q_2 m), \quad \sum_2 = \sum_{\substack{z_0 < Q \leqq x \\ Q \equiv -1 (\mathrm{mod}\ D)}} \sum_{\substack{q \equiv -1 (\mathrm{mod}\ Q^2) \\ q \leqq x}} d(qm).$$

We obtain evidently, that

$$\sum_1 \ll x \log x \sum_{z_0 < Q \leqq x} \sum_{\substack{q_1 \equiv q_2 \equiv -1 (Q) \\ q_1, q_2 \leqq x}} \frac{1}{q_1 q_2} \ll \frac{x (\log x)^3}{z_0} \ll \frac{x}{\log^2 x}.$$

Similarly, we have

$$\sum_2 \ll \sum_{z_0 < Q \leqq x} \frac{x \log^3 x}{Q^2} \ll \frac{x}{\log^2 x}$$

and so (3. 5) is proved.

In order to prove (3. 6), let

$$S_2(x) = S_3(x) + S_4(x),$$

where in $S_3(x)$ we sum over those $q+1$ the greatest prime divisor of which is smaller than z_2, and in $S_4(x)$ over the others.

Using Lemma 4, we easily deduce that

$$S_3(x) \leqq M(x, z_2) \ll \frac{x}{\log^2 x}.$$

We consider now $S_4(x)$. For the q occurring in the sum $S_4(x)$ let $q+1 = A(q) B(q)$, where

$$A(q) = \prod_{\substack{p | q+1 \\ p \leqq z_1}} p, \quad B(q) = \prod_{\substack{p | q+1 \\ p > z_1}} p.$$

Let p^* denote the maximal prime divisor of $q+1$, and write

$$B^*(q) \cdot p^* = B(q), \quad A(q) B^*(q) = k.$$

Since, for a fixed k, by Lemma 3 it follows that

$$\sum_{A(q) B^*(q) = k} r(q+1) \ll |r(k)| N(k, x) \ll \frac{|r(k)|}{\varphi(k)} \frac{x}{\log^2 \frac{x}{k}},$$

so we have

$$S_4(x) \ll \sum_{k \leqq \frac{x}{z_2}} |r(k)| N(k, x) \ll \frac{x (\log_2 x)^2}{\log^2 x} \sum_{k \leqq \frac{x}{z_2}} \frac{|r(k)|}{\varphi(k)} \ll$$

$$\ll \frac{x (\log_2 x)^2}{\log^2 x} \prod_{p \leqq z_1} \left\{1 + \frac{|r(p)|}{p-1}\right\} \cdot \prod_{\substack{z_1 < p \leqq x \\ p \in \mathcal{T}}} \left\{1 + \frac{|r(p)|}{p-1}\right\}.$$

Here $\mathcal{T}$ denotes the set of those p for which $p+1\not\equiv 0 \pmod{Q}$, if $Q\equiv -1 \pmod{D}$ and $Q>z_0$.

Obviously

$$\prod_{p<z_1}\left\{1+\frac{|r(p)|}{p-1}\right\}\ll(\log z_1)^2\ll(\log_2 x)^4.$$

Furthermore, applying Lemma 5 and the Corollary to $y=z_0$, $u\geqq z_1$, we have

$$\log\prod_{\substack{z_1<p\leqq x\\ p\in\mathcal{T}}}\left\{1+\frac{|r(p)|}{p-1}\right\}<2\sum_{2^\nu\leqq\frac{x}{z_1}}\;\sum_{\substack{2^\nu z_1<p\leqq 2^{\nu+1}z_1\\ p\in\mathcal{T}}}\frac{1}{p}<\varepsilon\log_2 x,$$

whence it follows

$$\prod_{\substack{z_1<p<x\\ p\in\mathcal{T}}}\left\{1+\frac{|r(p)|}{p-1}\right\}\ll(\log x)^{\varepsilon}.$$

So (3. 6) holds.

Taking into account the inequality (3. 4), from Lemma 2 and Lemma 7 it follows that

$$A_3(x)\gg \operatorname{li} x.$$

Using the Cauchy—Schwartz inequality and Lemma 6, we obtain

$$\frac{x}{\log x}\ll A_3(x)\ll\Bigg\{\sum_{\substack{q\leqq x\\ q\in\mathfrak{A}_3}}1\Bigg\}^{1/2}\Bigg\{\sum_{q\leqq x}|r^2(q+1)|\Bigg\}^{1/2}\ll g(D,3,x)^{1/2}\cdot x^{1/2}\log x.$$

Hence the assertion of the Theorem evidently follows.

Refereces

[1] Ю. В. Линник, Дисперсионный метод в бинарных аддитивных задачах, Изд. Ленинградского Университета, 1961.

[2] C. Hooley, On the representation of numbers as the sum of two squares and a prime, *Acta Math.*, **97** (1957), 189—210.

[3] E. Bombieri, On the large sieve, *Mathematika*, **12** (1965), 201—225.

[4] P. Erdős, Remarks on number theory. II, *Mat. Lapok*, **12** (1961), 161—169 (in Hungarian).

[5] I. Kátai, Statistical theorems for the number of prime factors of integers, *unpublished.*

[6] K. Prachar, *Primzahlverteilung* (Berlin—Heidelberg—Göttingen, 1957).

(Received September 26, 1967)

Bibliographie

N. I. Achieser—I. M. Glasmann, Theorie der linearen Operatoren im Hilbert-Raum (Mathematische Lehrbücher und Monographien, I. Abteilung. Mathematische Lehrbücher, Bd. 4), vierte, unveränderte Auflage, XIV+369 Seiten, Berlin, Akademie-Verlag, 1965.

This book, the German translation of the Russian original (1950), contains a complete presentation of the classical theory of self-adjoint operators in Hilbert space, and a summary of progress (prior to 1950) achieved in this field by Soviet mathematicians.

There are seven chapters and two appendices.

The first four chapters show no essential methodical difference from the corresponding parts of other well-known books in functional analysis. They give a survey of standard definitions and theorems in the theory of Hilbert space (geometry of Hilbert space, linear functionals and operators, special operators: projections, isometric and unitary operators, the notion of the spectrum and resolvent).

Chapter V deals with the elementary form of the spectral theory of completely continuous operators. The spectral theorem, stated in terms of reduction to diagonal form, is proved for a normal completely continuous operator by elementary means.

Chapter VI contains the spectral analysis of self-adjoint and unitary operators (the spectral theorem for normal operators is not mentioned). The chapter begins with a discussion of the moment problems, including BOCHNER's theorem on positive definite functions. The spectral theorem for unitary operators is proved by one of these moment theorems. STONE's theorem on integral representation for groups of unitary operators is proved by BOCHNER's theorem. For the spectral theorem of self-adjoint operators two proofs are given; the first of them uses the integral representation for the resolvent operators of a self-adjoint operator, while the second one is based on the spectral theorem for unitary operators and on Cayley transform. Other topics treated are: operators with simple spectrum; spectral types; functions of a self-adjoint operator; spectral resolution of the Schrödinger operators.

Chapter VII contains the theories of J. VON NEUMANN and M. G. KREIN on the extension of symmetric operators.

Appendix 1 deals with self-adjoint extensions of symmetric operators in a larger Hilbert space. The main result is that any symmetric operator in a Hilbert space H can be extended to a self-adjoint operator defined on a Hilbert space K, which is usually larger than H. This theorem is used to define a generalized resolvent of a symmetric operator, and an integral representation for this resolvent is obtained with respect to a generalized resolution of the identity.

Appendix 2 contains a detailed description of the foregoing theory as applied to ordinary linear differential operators of the n-th order.

L. Gehér (Szeged)

Н. И. Ахиезер — И. М. Глазман, Теория линейных операторов в гильбертовом пространстве [N. I. Achieser — I. M. Glasmann, Theorie der linearen Operatoren im Hilbertraum], Zweite, umgearbeitete und ergänzte Ausgabe, 544 Seiten, Verlag „Nauka", Moskau, 1966.

Das in seiner ersten Ausgabe 1950 erschienene und seitdem auch in fremdsprachigen Übersetzungen verbreitete, ausgezeichnete Lehrbuch der Verf. wurde in dieser neuen Ausgabe wesentlich umgearbeitet und ergänzt.

Die wesentlichsten Veränderungen sind die folgenden: Das Kapitel über vollstetige Operatoren betrachtet jetzt auch die Theorie von F. RIESZ über diese Operatoren, den Satz über die Existenz nichttrivialer invarianter Unterärume für vollstetige Operatoren, sowie Ausführugen über nukleare Operatoren. Ein ganz neues Kapitel wurde den Störungsproblemen selbstadjungierter Operatoren gewidmet, insbesondere betrachtet man hier die Sätze von M. ROSENBLUM und T. KATO über die Invarianz (bis auf unitäre Äquivalenz) des absolut stetigen Teiles eines selbstadjungierten Operators bei Störung durch einen nuklearen Operator. Ein neuer umfangreicher Anhang wird den Integraloperatoren gewidmet (Theorien von CARLEMAN und VON NEUMANN). Auch in den anderen Kapiteln findet man manches neue Material, entsprechend der Weiterentwicklung der Theorie seit der ersten Ausgabe. Natürlich wollten (und konnten) die Verf. nicht *alles* einarbeiten, was man heute über Hilbertraumoperatoren weiß, dazu wäre ein einziges Lehrbuch viel zu wenig.

Doch haben die Verf. erreicht, ihr Buch noch lehrreicher zu gestalten. Gewiß wird es vielen Erfolg bei den Lesern haben.

B. Sz.-Nagy (Szeged)

Ralph P. Boas, Jr., Integrability theorems for trigonometric transforms (Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 38), VIII+65 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1967.

This monograph offers a survey of the present stage of a special but important collection of problems concerning trigonometric series about which a sizeable literature has grown up in recent years.

The typical problems of this subject are raised as follows. Suppose that a periodic function $f(x)$ is associated with a trigonometric cosine series $\frac{a_0}{2}+\sum_{n=1}^{\infty} a_n \cos nx$ (e. g. $f(x)$ is integrable in some sense and a_n are its Fourier coefficients). We ask two questions: a) if $\varphi(x)$ is a given function, and $f(x)$ belongs to a specified class of functions, what hypotheses on $\{a_n\}$ are equivalent to the fact that $f(x)\varphi(x)$ belongs to a certain class of functions (e. g. $f(x)\varphi(x) \in L^p(0, \pi)$); b) if $\{\alpha_n\}$ is a given sequence of numbers, and $\{a_n\}$ belongs to a given class of sequences, what conditions on $f(x)$ are necessary and sufficient in order that $\{\alpha_n a_n\}$ should belong to a certain class of sequences (e. g. $\{\alpha_n a_n\} \in l^p$). Similar questions cas be asked about sine series as well.

The author has not aimed at encyclopedic completeness; he presents only the more general and more characteristic theorems from the large number of results. Good arrangement helps the reader to get a picture on the material of the book.

This monograph was written at the suggestion of BÉLA SZ.-NAGY whose paper, "Séries et intégrales de Fourier des fonctions monotones non bornées", published in these *Acta*, **13** (1949), inspired the greater part of the material that is presented in this book. Correspondingly, one considers in particular the cases $\varphi(x) = x^{-\gamma}$ and $\alpha_n = n^{-\gamma}$.

§ 3 and § 4 treat the case $0<\gamma<1, p=1$, provided that the functions $f(x)$ and the coefficients a_n are positive or decreasing, respectively. In § 5 we can find a more detailed analysis of the cases $\gamma=0$ and $\gamma=1$, § 6 is devoted to L^p problems, $1<p<\infty$, and § 7 deals with the case $p=\infty$. In § 8 a number of various generalizations is considered, e.g. the author uses a more general class of multipliers instead of $x^{-\gamma}$ and $n^{-\gamma}$, or, e.g., he shows that conditional convergence of the series $\Sigma a_n \alpha_n$ or of the integral $f(x)\varphi(x)dx$ implies conditional convergence of the other. § 9 touches the question of trigonometric integrals, selecting a few theorems particularly relevant to the theme of the monograph.

The theorems in the book — apart from § 9 — are proved in a detailed, concise form; thus the reader receives an overall picture of the methods of proof. To show the sharpness of some theorems, the author presents striking examples. The author mentions a number of unsolved problems which will certainly inspire further research. Some unpublished recent results, communicated to the author, are also included. A comprehensive bibliography gives a survey of the former literature on the subject.

K. Tandori—F. Móricz (Szeged)

N. Bourbaki, Eléments de mathématique, Fascicule XV, **Espaces vectoriels topologiques,** Chapitres 1 et 2: **Espaces vectoriels topologiques sur un corps valué, Ensembles convexes et espaces localement convexes,** deuxième édition revue et corrigée, VI+178 pages, Paris, Hermann, 1966.

The first edition appeared in 1953.

This book contains the first two chapters of an exposition of the theory of linear topological spaces. One difference between the first and second editions is that a paragraph on weak duality, which was contained in chapter 4 of the first edition, was included now into chapter 2. A more essential difference is that a more detailed development is given for the notions of the inductive limit for locally compact convex topologies and of the extremal generators of convex cones.

The book is a careful, extremely well-organized account of the foundations of the subject, covering essentially all the elementary topological and geometric prerequisites for the duality theory. The exposition is very well documented with examples and problems.

The first chapter treats linear topological spaces over a valued field and gives, in the usual way, the notion of topology, continuity and equicontinuity, and uniform structure. Product spaces, subspaces, and complementary subspaces are discussed briefly. The open mapping theorem and the closed graph theorem are proved for metrisable complete spaces over a valued field.

Chapter 2 is devoted to convex sets and to the elementary theory of locally convex spaces. The scalar field is supposed to be the reals, except in the last section where the scalars are supposed to be the complex numbers. The chapter begins with the study of topologies defined by semi-norms. After geometric preliminaries, the connection between cones and partial orders is studied. Two sections deal with the different forms (analytic and geometric) of the Hahn—Banach theorem; they are proved under the most general conditions (the classical form of the Hahn—Banach theorem is given as an example). In connection with the geometric form of the Hahn—Banach theorem, separation theorems are proved. Theorems concerning inductive limits of sequences of locally convex spaces, extremal generators of convex cones and weak topologies are also given. KREIN—MILMAN's theorem on the existence of extreme points for convex sets is proved, too. Finally a number of results are extended to complex spaces. In an appendix, the Markov—Kakutani fixed point theorem for a commuting family of affine maps on convex compact subsets of a linear topological space is proved, and, as an application, the existence of Haar measure on a compact Abelian group is derived.

L. Gehér (Szeged)

Hans Grauert und Ingo Lieb, Differential- und Integralrechnung. I (Heidelberger Taschen bücher, Band 26) X+200 Seiten, Berlin—Heidelberg—New York, Springer Verlag, 1967.

Dies Buch ist der erste Teil einer dreibändigen Darstellung der Differential- und Integralrechnung; in den folgenden Bänden werden Funktionen mehrerer Veränderlichen, gewöhnliche Differentialgleichungen und Integrationstheorie behandelt.

Die einzelnen Kapiteln sind der Reihe nach: Die reellen Zahlen, Mengen und Folgen, Unendliche Reihen, Funktionen, Differentiation, Spezielle Funktionen und Taylorscher Satz, Integration.

In dem ersten, auch vom didaktischen Standpunkt aus gut geschriebenen Kapitel werden die Axiome des reellen Zahlkörpers mit ihren einfachsten Folgerungen, die Grundbegriffe der Zahlenmengen und die logischen Gesetze besprochen.

Der weitere Aufbau der Theorie zeigt gewisse Besonderheiten. Die Konvergenz einer Reihe wird z. B. damit definiert, daß der limes superior und der limes inferior der Folge gleich sind. Statt Stetigkeit wird zuerst die Halbstetigkeit von Funktionen definiert; die Differenzierbarkeit wird als totale Differenzierbarkeit eingeführt. Die speziellen Funktionen (Exponentialfunktion, Logarithmus, trigonometrische Funktionen) werden durch Potenzreihen eingeführt. Das Buch gibt einen elementaren Aufbau des Lebesgueschen Integralbegriffes mit Hilfe von Treppenfunktionen; bei diesem Aufbau spielen die halbstetigen Funktionen eine wichtige Rolle. Diese Definition des Integrals kann man direkt auf Funktionen mit Werten in lokalkonvexen Vektorräumen übertragen. Die tieferen Sätze der Lebesgueschen Theorie werden in diesem Band noch nicht behandelt; die wichtigsten Sätze über das Riemansche Integral sind aber bewiesen. Fragen der Interpolationstheorie und der numerischen Integration werden ausführlich diskutiert.

Alle Beweise sind bis in die Einzelheiten hinein ausgeführt. Im Text gibt es aber nur wenige Beispiele, und wegen ihrer Besonderheiten ist die Betrachtungsweise nicht als einfach zu bezeichnen. Das interessant aufgebaute Buch ist also für eine erste Einführung weniger geeignet.

K. Tandori (Szeged)

Paul R. Halmos, A Hilbert space problem book, XVII+365 pages, D. Van Nostrand Company, Inc., Princeton—Toronto—London, 1967.

Until about twenty years ago it seemed that the theory of Hilbert space operators is essentially confined to *normal* (in particular selfadjoint and unitary) operators or to selfadjoint operator algebras. For the study of *non-normal* operators the inner product structure of Hilbert space seemed to be of little effect, so that, for these operators, one did not hope to obtain much more information in the Hilbert space than in the general Banach space situation.

This view had to be revised later in many respects. More and more special results were obtained which hold on Hilbert space operators without holding for operators in general Banach spaces. One of the first and most striking of these results was that of J. von Neumann (1950) asserting that, for any (linear bounded) operator T in Hilbert space, and for any polynomial p, the norm of the opeartor $p(T)$ is bounded from above by the maximum of the function $|p(\lambda)|$ on the disc $|\lambda| \leqq \|T\|$. One has discovered interesting properties of a particular non-normal operator in Hilbert space, the "unilateral shift", simple or multiple; one has solved for it the invariant subspace problem by essential use of the theory of Hardy class analytic functions in the unit disc (Beurling, Lax, Halmos). New ways were devised to relate non-normal operators to normal ones, in particular the *dilation* of the operator (Naĭmark, Halmos, Sz.-Nagy, etc.). New classes of operators in Hilbert space were studied, which share some properties with the normal operators (*subnormal, hyponormal* operators, etc.). The *numerical range* $W(T)=\{(Th, h):\|h\|=1\}$ of operators was extensively studied. The problem of determining the commutators of Hilbert space operators, raised by the commutation relations of quantum mechanics, was solved. An operator valued analytic function $\Theta_T(\lambda)$ has been attached to operators T, in particular to contractions $T(\|T\| \leqq 1)$; by this "characteristic function" (rather than by the classical resolvent function) it was possible to obtain functional models for these operators, study their invariant subspaces, etc.

This sample taken from the variety of directions in which research work has been done recently concerning Hilbert space operators might suffice to show that we have to do already with a many-faced new theory.

The present "problem book" of Professor Halmos gives an excellent introduction to a large part of this new theory, to which he and the group of his collaborators have contributed many basic results. The first part of the book ("Problems") lists 199 problems, each introduced in a very instructive manner by an outline of the pertinent chapter of the theory. There follows a part ("Hints") where some short indications are given how to try to find a solution, pointing to what the author regards as the heart of the matter. The last and largest part of the book are the "Solutions", with many interesting further remarks and comments added.

This structure of the book appeals to the reader who is willing to learn in the active way thus achieving more lasting result and ability to further work on the subject. But also the reader whose concern is only to get a short, but authentic orientation on these topics, will find this structure of the book of advantage: reading the "Problems" and occasionally having a glimpse into the "Solutions", he is given a masterly presentation of a large part of recent development in the theory of Hilbert space operators.

The "Problems" — and accordingly the two other parts — are grouped into 20 chapters. Here are the titles (with some samples from the contents): 1. *Vectors and Spaces.* (If $\{e_n\}$ is an orthonormal basis for a Hilbert space H and if $\{f_n\}$ is an orthonormal set in H such that $\Sigma\|e_n - f_n\|^2 < \infty$, then $\{f_n\}$ is also a basis for H.) — 2. *Weak topology.* (The closed unit ball in Hilbert space is compact with respect to weak topology. The weak topology of an infinite dimensional Hilbert space is not metrizable.) — 3. *Analytic functions.* (Analytic Hilbert spaces; Hardy classes; kernel functions, etc.) — 4. *Infinite matrices.* (Schur test; Hilbert matrix.) — 5. *Boundedness and invertibility.* (Very simple proofs are given for the uniform boundedness and the closed graph theorems.) — 6. *Multiplication operators.* (Diagonal operators; multipliers of functional Hilbert spaces.) — 7. *Operator matrices.* (If C and D commute, and if D is invertible, then $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ is invertible iff $AD-BC$ is invertible.) — 8. *Properties of spectra.* (The boundary of the spectrum is included in the approximate point spectrum.) — 9. *Examples of spectra.* (Point spectrum, approximate point spectrum, and compression spectrum of diagonal operators, of multiplication operators, and of shifts.) — 10. *Spectral radius.* (In particular for "weighted shifts".) — 11. *Norm topology.* (The spectral radius $r(A)$ is a discontinuous function of A in the norm tolpogy.) — 12. *Strong and weak topologies* (for operators. Multiplication is discontinuous in the strong and weak topologies of operators. Is a bounded increasing sequence of Hermitian operators necessarily strongly convergent? uniformly convergent?) — 13. *Partial isometries.* (Existence, for every contraction A in H, of a dilation to a partial isometry $M(A)$ in $H \oplus H$, such that if A, A' are unitarily equivalent then so are $M(A)$, $M(A')$; if A and A' are invertible, then the converse also holds.) — 14. *Unilateral shift.* (Does it have a square root? Commutants of shifts. Every contraction A such that $A^n \to O$ strongly, is unitarily equivalent to a part of the adjoint of a unilateral shift. Invariant subspaces of a simple unilateral shift. If U is a unilateral shift of countable multiplicity then U has a cyclic vector.) — 15. *Compact operators.* (If C is compact and if $Ch=h$ implies $h=0$, then $I-C$ is invertible. Volterra operators are quasi-nilpotent.) — 16. *Subnormal operators.* (Is a contraction similar to a unitary operator necessarily unitary?) — 17. *Numerical range.* (Convexity. Questions concerning closure. The inequality $w(A^n) \leqq w(A)^n$ for the numerical radius.) — 18. *Unitary dilations.* (Every contraction has a unitary power dilation.) — 19. *Commutators of operators* (i. e. operators of the form $PQ-QP$ where P, Q are operators. If an operator A on a separable Hilbert space is not a scalar, then $A \oplus A \oplus \dots$ is a commutator.) — 20. *Toeplitz operators* (i. e. operators on the Hardy space H^2, of the form $T_\varphi f = P(\varphi f)$ where φ is a bounded measurable function and P is the orthogonal projection from L^2 onto H^2. The only compact T_φ is O. The spectrum of T_φ if φ is real-valued.)

Thus the topics dealt with range from fairly text-book material to the boundary of what is known. As it should be, the choice of the problems reflects to some extent the interests of the author ("every problem in the book puzzled me once") and this lends an intriguing personal flavor to the book.

In any respect, this new book by Professor HALMOS is a very valuable and useful contribution to the literature on functional analysis, and certainly will be welcomed by a large circle of readers.

Béla Sz.-Nagy (Szeged)

Béla Kerékjártó, Les fondements de la géométrie, Tome deux, **Géométrie projective,** 528 pages, 168 figures, Budapest, Akadémiai Kiadó, et Paris, Gauthier-Villars, 1966.

Cet ouvrage est une traduction de l'édition parue en langue hongroise en 1944 (A geometria alapjairól, Második kötet, Projektív geometria). Le texte était revisé par le rédacteur de cette édition, G. HAJÓS et par l'auteur de ce compte rendu; de plus, il était complété par quelques remarques du rédacteur. Ici nous ne voulons pas nous occuper du sommaire du livre (cf. *Acta Sci. Math.*, **11** (1946/48), 254), mais nous devons souligner que, jusqu'aujourd'hui, cet ouvrage donne le développement le plus complet de la géométrie projective. Ainsi, outre des théorèmes classiques, on y trouve un large exposé de la théorie des groupes de transformations projectives, l'introduction de la measure projective dans les géométries non-euclidiennes et l'examen critique des divers systèmes d'axiomes de la géométrie projective. C'est pourquoi le livre de KERÉKJÁRTÓ sera très profitable pour tous ceux qui veulent étudier ce discipline mathématique d'une façon approfondie.

G. Szász (Nyíregyháza)

E. Hewitt and K. Stromberg, Real and Abstract Analysis, VIII+476 pages, Berlin—Heidelberg—New York, Springer-Verlag, 1965.

Measure and integration theory has been restricted to R^n for a long time. However, in the frame of the development of mathematics as a whole this theory also arrived at higher levels of abstraction. DANIELL's work concerning integration on an arbitrary abstract space is already classical. The abstract forms of measure and integration theory have proved to be very useful also for several other branches of modern mathematics. Thus, it is no wonder that many monographs and textbooks have been devoted to the subject. The book under review is one of the most sympathetic books written in this field.

The authors' standpoint is that modern analysis draws on at least five disciplines. First, in dealing with measure theory, and even in studying the structure of real numbers, one must often use powerful machinery from the abstract theory of sets. Second, algebraic ideas and techniques prove to be sometimes essential in studying problems in analysis. Third, set-theoretic topology is an important tool in constructing and studying measures. Fourth, the method of functional analysis can often be used with much success to obtain fundamental results in analysis with relatively little effort. Fifth, analysis really is *analysis*. Handling inequalities, computing with actual functions, and obtaining actual numbers, is indispensable for the training of every mathematician. The spirit in which the book has been written was certainly suggested by this conception. In fact, though measures and integrals constitute the primary objects of study, all five of these subjects find a place in the book, each of them developed just to the extent which assures the book to be self-contained as much as possible. Another virtue of the book is that the authors always point out that the abstract forms of measure and integration theory came into being as a very generalization of that on R^n. Thus they do not let the reader get lost in the jungle of abstract concepts and acquire only a formal knowledge of the subject.

The book consists of six chapters.

Chapter 1: Set theory and algebra. It begins with elementary facts on sets and functions. Then there are sections discussing the axiom of choice and its equivalents, cardinals, ordinals, transfinite induction.

Chapter 2: Topology and continuous functions. It begins with a self-contained treatment of those parts of set-theoretic topology that have proved important for analysis. There follows a study of continuous functions, and functions closely related to the continuous ones. In this study, the methods of Banach algebras evidently play a role, thus Banach algebras are also defined and illustrated. A detailed and careful treatment is given of various versions of the Stone-Weierstrass theorem.

Chapter 3: The Lebesgue integral. After a rapid discussion of the Riemann and Riemann—Stieltjes integral, the authors proceed as follows. They start with a positive linear functional I defined on continuous functions with compact support on a locally compact Hausdorff space X. They extend I to an "upper integral" $\bar{\bar{I}}$ and define outer measure ι by means of $\bar{\bar{I}}$ à la BOURBAKI. Measurable sets and functions, and the measures are then defined à la CARATHÉODORY. Now if (X, A, μ) is an arbitrary abstract measure space and f is an arbitrary nonnegative extended real-valued function on X, then the "integral" of f is defined as

$$L_\mu(f) = \sup\left\{\sum_{k=1}^{n} \inf\{f(x): x \in A_k\}\, \mu(A_k)\right\}$$

where the supremum is taken over all possible measurable dissections $A_1, \ldots, A_n$ of X. In particular. if X is locally compact and the measure space (X, A, ι) derives from the functional I then $L_\iota(f) = I(f)$ for every nonnegative measurable function f on X (a generalized form of the Riesz representation theorem). In this way, a connection between the Daniell and Carathéodory approaches to integration theory is established.

Chapter 4: Function spaces and Banach spaces. One begins with L^p spaces. This leads to a study of abstract Banach spaces, centered around the three basic principles: the principle of uniform boundedness, the interior mapping principle, and the Hahn—Banach theorem. The conjugate of L^p is computed by means of uniform convexity and CLARKSON's inequalities. Then again abstract spaces, in particular Hilbert spaces are studied; the orthogonal expansions discussed in an abstract form lead to non-trivial facts about Fourier series.

Chapter 5: Differentiation. First a brief but reasonably complete treatment of the usual pointwise theory is persente d. The main result proved re is LEBESGUE's theore m that a function of finite variation has a finite derivative almost everywhere. The main tool here is VITALI's covering theorem. Then the conditions under which the classical equality $f(b) - f(a) = \int_a^b f'(x)\,dx$ is valid are explored. This leads to interesting measure-theoretic ideas which have little to do with differentiation but have applications in diverse fields. The main result here is the Lebesgue-Nikodým theorem which is thoroughly examined and several applications of it to problems in absctract analysis, such as integration by substitution, computation of the conjugate of L^1 and L^∞, the Riesz representation theorem, are given.

Chapter 6: Integration on product spaces. First comes a discussion of the Fubini theorem with applications to convolution, Fourier transform and to the proof of the Hardy—Littlewood maximal theorem on the line. The final topic of this chapter is that of infinite products of measure spaces with probabilistic interpretations including the zero-one law and the strong law of large numbers.

A number of excellent exercises is added which cover a very wide range of difficulty; the more difficult ones are provided with hints.

The authors did a very nice job in writing this book. Every analyst, student or researcher, will find in it much of interest and value.

I. Kovács (Szeged)

O. Haupt—H. Künneth, Geometrische Ordnungen (Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, Bd. 133), VIII+429 Seiten, Berlin—Heidelbeg—New York, Springer-Verlag, 1967.

Das Buch ist Problemen der geometrischen Ordnungen gewidmet, d.h. Problemen der folgenden Art: In einem kompakten metrischen Raum G ist ein System $\mathfrak{k}$ von gewissen (voll-) kompakten Teilmengen $K \subset G$, — die sog. Ordnungscharakteristiken — gegeben. Ein Bogen, eine Kurve, allgemein ein Kontinuum $C \subset G$ heißt vom Punktordnungswert m bezüglich $\mathfrak{k}$, wenn C mit jedem $K \in \mathfrak{k}$ maximal m Punkte gemeinsam hat. Man fragt nun nach der „Gestalt" von C bezüglich $\mathfrak{k}$, also nach den ordnungshomogenen Teilkontinuen von C, d.h. nach Teilkontinuen, deren sämtliche Teilkontinuen die gleiche Ordnung haben, sowie nach den ordnungssingulären Punkten von C, d.h. nach Punkten, die nur solche Umgebungen auf C haben, deren Ordnung verschieden von dem der ordnungshomogenen Teilkontinuen ist, und evtl. nach infinitesimalgeometrischen Eigenschaften von C. Daneben ergeben sich auch Fragen hinsichtlich der Komponentenordnungswerte von C, worunter die Anzahl der Zusammenhangskomponenten von $C \cap K$ zu verstehen ist. Ist weiterhin m die kleinste unter den ganzen Zahlen $m' \geqq 0$, für welche die Menge der $K \in \mathfrak{k}$, die mit C mehr als m' Punkte gemeinsam haben, nirgends dicht in $\mathfrak{k}$ ist, bezüglich der in $\mathfrak{k}$ in üblicher Weise erklärten Metrik, so wird m als der schwache Punktordnungswert von C bezüglich $\mathfrak{k}$ bezeichnet.

Im I. Teil ist der Grundbereich G zumeist eine abgeschlossene Kreisscheibe, evtl. ein topologisches Bild derselben. Die Ordnungscharakteristiken K sind Bögen und Kurven, welche den folgenden Axiomen genügen:

I. Ist K ein Bogen, so hat K mit der Begrenzung G_g von G genau seine beiden Endpunkte gemeinsam. Ist K eine Kurve, so hat K mit G_g höchtens einen Punkt gemeinsam.

II. Es existiert eine natürliche Zahl $k \geqq 1$, die sog. Grundzahl, sodaß durch irgendwelche k verschiedene Punkte x_i $(i=1, \ldots, k)$ einer $K(\in \mathfrak{k})$ und alle zu solchen x_i hinreichend benachbarten x_i' (wobei $x_i'=x_i$ sein kann) ein $K' \in \mathfrak{k}$ eindeutig bestimmt ist, der sich mit den x_i' stetig ändert.

Dazu kommt in einigen Betrachtungen folgende zusätzliche Voraussetzung hinzu:

Es sei $B \subset G$ ein Bogen und x_i $(i=1, \ldots, k)$ irgend k verschiedene Punkte von B. Ferner gebe es $K_n \in \mathfrak{k}$ $(n=1, 2, \ldots)$ und $x_{ni} \in K_n$ mit $x_i = \lim x_{ni}$ $(i=1, \ldots, k)$. Dann existiert $K \in \mathfrak{k}$ mit $x_i \in K$ $(i=1, \ldots, k)$.

Es werden zunächst im Abschnitt 1 einige Folgerungen aus den obigen Axiomen abgeleitet und dann die allgemeinen Eigenschaften der Kontinuen von höchstens endlichen Komponentenordnungswerten zusammengestellt, sowie für den Fall der Grundzahl $k=1$, alle ordnungshomogenen Kontinuen bestimmt.

Nach diesen allgemeinen Erörterungen werden im Abschnitt 2 die Änderungen untersucht, welche von den evtl. gemeinsamen Punkten aller Ordnungscharakteristiken verschiedener Punkte einer Ordnungschrakarteristik, die sie mit einem Bogen B gemeinsam hat, erleidet, wenn man k—1 Punkte unter ihnen festhält.

Im Abschnitt 3 werden die früher für Systeme von Ordnungscharakteristiken in der euklidischen Ebene für eine beliebige Grundzahl $k \geqq 2$ dargelegte Methode für den Fall der sog. topologisch projektiven Ebene, d.h. für den Fall der reellen projektiven Ebene (oder eines topologischen Bildes von ihr) als Grundgebiet und für $k=2$, angewendet. Nach Verallgemeinerung gewisser Sätze über konvexe Mengen, wird die Gestalt der Bögen und Kurven vom Punktordnungswert 3 bestimmt, insbesondere bei Bogen die Maximalzahl und bei Kurven die genaue Anzahl der singulären Punkten, und es wird bewiesen, daß sich diese Bögen und Kurven als Vereinigungen einer beschränkten Anzahl von Bogen, die den schwachen Punktordnungswert zwei besitzen ($\mathfrak{k}$-konvexe Bogen) darstellen laßen. Damit ist die Juelsche Theorie der Kurven 3. Ordnung verallgemeinert worden. Dann kommen Bemerkungen über Kurven höherer Ordnung, sowie der Satz über die Existenz ordnungsgomogener Bögen. Zum Schluß wird gezeigt, daß auch die Theorie von J. v. Sz. Nagy über Kurven vom Maximalklassenindex, d.h. üver Kurven, bei denen das Minimum der Anzahl der Tangenten von den Punkten der Ebene an die Kurve um zwei kleiner ist als das Maximum der Anzahl derselben, sowie auch Sätze von Möbius, A. Kneser und Kivikoski gültig bleiben.

Im Abschnitt 4 werden Systeme von Ordnungscharakteristiken mit beliebiger Grundzahl $k \geqq 2$ betrachtet und zunächst die untere und obere Schranken für die Anzahl der singulären Punkte von Bögen und Kurven untersucht. Dann wird die bekannte Eigenschaft der Viertelellipse, daß von je zweien der Schmiegkreise derselben immer einer den anderen umfaßt, auf ein System $\mathfrak{k}$ von Kurven, deren jede durch k ihrer Punkte bestimmt ist, wobei $k \geqq 3$ ungerade ist, und auf einen Bogen, der bezüglich $\mathfrak{k}$ den Punktordnungswert k besitzt, verallgemeinert. (Für gerades $k \geqq 4$ ergibt sich analog eine Verallgemeinerung des Verhaltens der Konvexbogen bzw. ihrer Stützgeraden.) Dann wird eine Kennzeichnung der Kurven von der zyklischen Ordnung vier (die also den Punktordnungswert 4 besitzen bezüglich des Systems der in G liegenden Kreise und Kreisbogen) und ihre Verallgemeinerung angegeben. Endlich wird der Satz von BÖHMER, nach welchem der durch fünf beliebige Punkte eines Ovals bestimmte Kegelschnitt eine Ellipse ist, falls der Schmiegkegelschnitt in jedem Punkte des Ovals eine Ellipse ist, und die damit zusammenhängenden Sätze von MOHRMANN, CARLEMANN und MUKHOPADHYAYA verallgemeinert. Es ergibt sich noch ein 2-Scheitelsatz für (ebene) Jordankurven, sowie die topologische Verallgemeinerung des Kneserschen 4-Scheitelsatzes.

Der II. Teil, der die Abschnitte 5—7 umfaßt, ist Problemen in n-dimensionalen und noch allgemeineren Räumen gewidmet. Im Abschnitt 5 werden die Kontinua höchstens endlichen Ordnungswertes bezüglich der Hyperebenen im n-dimensionalen projektiven Raum P_n behandelt, und dann werden zunächst Parameterbogen und ihre k-dimensionale Schmieg-, insbesondere Tangential- (halb-) ebenen in P_n untersucht. Unter Benutzung dieser Entwicklungen werden die Bogen in P_n vom Punktordnungswert n und n—1 näher untersucht, und die Monotonie der Halbtangenten eines Bogens vom Punktordnungswert n mit stetiger Tangente im E_n bewiesen. Nunmehr werden in projektiven P_n $(n \geqq 2)$ einige Eigenschaften der schwach ordnungsminimalen Kontinua, d. h. der Kontinuen $C \subset P_n$, die nicht in einer Hyperebene von P_n liegen, mit schwachem Punktordnungswert n bezüglich des Systems der Hyperebenen von P_n, wie die Besehränktheit der Anzahl ihrer Verzweigungspunkte, sowie ihre Darstellbarkeit als Vereinigung von endlich vielen einfachen Bögen u. s. w., bewiesen. Es werden im euklidischen Raum $E_n (n \geqq 2)$ die Kontinua behandelt, die für keine Hyprebene mehr als $n-1$ zu ihnen parallele Tangenten besitzen, sowie die Bögen vom schwachen Punktordnungswert m. Endlich wird eine Klassifikation der regulären und singulären Punkte eines offenen Bogens in $P_n (n \geqq 2)$ gegeben.

Im Abschnitt 6 werden die t-dimensionalen Kompakta von endlichem Punktordnungswert in E_n behandelt, wobei über einen Satz von G. NÖBELING berichtet wird, der eine Verallgemeinerung des Satzes auf (voll-) kompakte metrische Räume ist, nach welchem alle Kontinua der Ebene, die von höchstens endlichem Punktordnungswert bezüglich eines Systems $\mathfrak{k}$ von Ordnungscharakteristiken mit der Grundzahl $k=1$ sind, sich als abgeschlossene Hülle der Vereinigung von abzählbar vielen Bogen mit Punktordnungswert 1 darstellen lassen, und es werden Bemerkungen zu diesem Satz betreffs der Darstellbarkeit durch Lipschitz-Flächenstücke (d.h. durch t-Flächenstücke von beschränkter Dehnung) hinzugefügt.

Im Abschnitt 7 werden ordnungsgeometrische Probleme in kompakten metrischen Räumen behandelt. Es wird gezeigt, daß einige grundlegende Sätze, die im Abschnitt 1 lediglich für Systeme $\mathfrak{k}$ von Ordnungschrakteristiken in der Ebene bewiesen wurden, unter ziemlich allgemeinen Voraussetzungen gelten.

Wie es sich auch aus dieser kurzen Inhaltsübersicht herausstellt, ist das vorliegende Buch in erster Linie eine weitgehende Verallgemeinerung von früher unter engeren Voraussetzungen bewiesenen klassischen Sätzen, die die toplogische Natur derselben deutlich hervortreten läßt. Um aber einen Überlick über neuere Untersuchungen zu geben, werden im letzten Teil des Buches neben verschiedenen Ergänzungen auch Berichte über Arbeiten u. a. von D. DERRY, F. FABRICIUS-BJERRE, A. MARCHAUD und P. SCHERK gebracht.

Ein reichhaltiges Literaturverzeichnis ergänzt dieses schön ausgestattete, wertvolle Buch.

J. Strommer (Budapest)

Andor Kertész, Vorlesungen über Artinsche Ringe, 282 Seiten, Budapest, Akadémiai Kiadó, 1968.

Im Studium von beliebigen Gruppen spielen die sogenannten Endlichkeitsbedingungen eine wichtige Rolle, denn die allgemeine Gruppentheorie ist historisch aus der Theorie der endlichen Gruppen ausgewachsen. Eine Endlichkeitsbedingung für eine Gruppe ist eine solche Bedingung, die für jede endliche Gruppe, nicht aber für alle unendlichen Gruppen erfüllt ist. In der Ringtheorie spielen die Algebren endlichen Ranges über einem Körper eine gewissermaßen ähnliche Rolle, wie die endlichen Gruppen in der Gruppentheorie. Weiterhin sind in der Ringtheorie die Minimalbedingungen bzw. Maximalbedingungen für Ideale bzw. Rechtsideale die wichtigsten Endlichkeitsbedingungen, und die Algebren endlichen Ranges über einem Körper genügen diesen beiden Kettenbedingungen. Auf die Bedeutung der Kettenbedingungen in Ringen machte schon EMMY NOETHER aufmerksam. Ein Ring erfüllt bekanntlich dann die Minimalbedingung (bzw. Maximalbedingung) für Rechtsideale, wenn jede echt absteigende (bzw. aufsteigende) Kette von Rechtsidealen des Ringes endlich ist.

Die Artinschen Ringe sind die Ringe mit Minimalbedingung für Rechtsideale. Von EMIL ARTIN stammt nämlich die wichtige Entdeckung, daß die Struktursätze der Algebren endlichen Ranges über einem Körper sich auf die größere Klasse der Ringe mit Minimalbedingung für Rechtsiadeale übertragen lassen.

Abgesehen von einem Kapitel des Buches „The Theory of Rings" von N. JACOBSON (1943), ist die erste monographische Darstellung über Artinsche Ringe das Buch „Rings with minimum condition" von E. ARTIN, C. NESBITT und R. THRALL (1944). Die seitdem verflossenen Jahre haben eine Fülle weiterer Ergebnisse auf diesem Gebiet erbracht, die eine neue monographische Bearbeitung notwendig machten. Obwohl einige diesbezügliche wichtige Resultate in einem Kapitel des Buches „Structure of Rings" von N. JACOBSON (1956) bzw. in einem Kapitel des Buches „Éléments de mathématique, Algèbre" von N. BOURBAKI (1958) betrachtet sind, streben aber diese Kapitel nach keiner vollständigen Darstellung des Stoffes.

Das vorliegende, neue Buch über Artinsche Ringe entstand aus einem gemeinsamen Plane von T. SZELE und A. KERTÉSZ, an dessen Verwirklichung T. SZELE aber wegen seines frühzeitigen Tod (1955) leider nicht mitwirken konnte. Prof. A. KERTÉSZ widmet sein Buch dem Andenken von T. SZELE, und bringt seine Dankbarkeit an seinen Lehrer und Mitarbeiter auch hiermit zum Ausdruck.

Das Buch enthält 10 Kapitel und einen Anhang über diejenigen Aspekten der Theorie der Abelschen Gruppen, die für die Untersuchung der Artinschen Ringe nötig sind. Die Kapitel sind die folgenden:

I. Mengen, Relationen;
II. Der Ringbegriff;
III. Ringkonstruktionen;
IV. Moduln und Algebren;
V. Das Radikal;
VI. Allgemeines über Artinsche Ringe;
VII. Ringe linearer Transformationen;
VIII. Halbeinfache Ringe und vollständig primäre Ringe;
IX. Moduln über halbeinfachen Ringen;
X. Die additive Struktur der Artinschen Ringe.

Das Buch gründet sich auf Vorlesungen, die Verf. 1962/63 an der Universität zu Halle gehalten hat. Dementsprechend verfolgt der Verf. ein zweifaches Ziel: einerseits wünscht er eine Einführung in die allgemeine Ringtheorie zu bieten (wobei im Kapitel V unter dem Radikal stets das Jacobsonsche Radikal verstanden wird), andererseits will er die wichtigsten Resultate der Theorie der Artinschen Ringe darstellen. Dabei strebt Verf. nicht nach Vollständigkeit, jedoch werden einige Arbeiten bezüglich der im Buch nicht betrachteten Theorie der einreihigen Ringe, der Quasi-Frobeniusschen Ringe, der distributiv darstellbaren Ringe, des Basisunterringes, der linear kompakten Ringe, der Ringe mit Minimalbedingung für Hauptrechtsideale und bezüglich der im Buch nicht betrachteten Galoisschen Theorie für Artinsche Ringe im Literaturverzeichnis angeführt. Fast zweidrittel Teil des Buches diskutiert Resultate über Mengen, algebraische Strukturen und beliebige (assoziative) Ringe, und umgefähr eindrittel Teil des Textes enthält Ergebnisse über Artinsche Ringe. Hiernach kann dieses vorzügliche, gut lesbare und nützbare Buch auch als ein Einführungsbuch in die Algebra bzw. in die Ringtheorie angesehen werden. Das Buch ist also keine Monographie. Bei der Ausbearbeitung des Stoffes sind einige Resultate ungarischer Algebraiker, insbesondere mehrere Resultate des Verfassers betrachtet.

Im allgemeinen wird ein großes Gewicht auf den streng systematischen Aufbau des Gegenstandes gelegt. Die Definitionen und Sätze sind klar formuliert, die Beweise sind poliert und möglichst kurz ausgeführt. Das Studium des interessanten Buches erfordert keine Vorkenntnisse. Nach jedem Kapitel gibt es zahlreiche Übungsaufgaben; am Ende des Buches sind die Anleitungen zur Lösung der schwierigeren Aufgaben angegeben. Weiterhin enthält das Buch eine Zusammenfassung der ständigen Bezeichnungen, einen Namen- und Sachregister, ein Literaturverzeichnis und es werden sechs offene Probleme formuliert. Aus diesen Problemen hat der Referent inzwischen das Problem 3 vollständig gelöst (vgl. *Acta Math. Acad. Sci. Hung.*, **18** (1967), 261—272) und gewisse Resultate auch im Zusammenhang mit dem Problem 1 erhalten (vgl. *Acta Sci. Math.*, **28** (1967), 31—37).

F. Szász (Budapest)

G. Pólya, Vom Lösen mathematischer Aufgaben, Einsicht und Entdeckung, Lernen und Lehren, Bd. 1—2 (Sammlung „Wissenschaft und Kultur", Bd. 20—21), 315+286 Seiten, Basel—Stuttgart, Birkhäuser Verlag, 1966—67. Ins deutsche übersetzt von Lulu Bechtolsheim.

„Jede Aufgabe, die ich löste, wurde zu einer Regel, die später zur Lösung anderer Aufgaben diente." Dieses Zitat von DESCARTES gibt das wesentliche Ziel des Buches. Das vorliegende Werk ist die Fortsetzung von zwei früheren („Schule des Denkens" und „Mathematik und plausibles Schliessen").

Die in diesem Buch vorgelegten Aufgaben verlangen kaum über das Mittelschulniveau hinausgehende Kenntnisse, aber sie verlangen einen gewissen und zuweilen einen hohen Grad der Konzentration und Einsicht. Die Themen des ersten Teils (Lösungsschemata) sind die folgenden: das Schema zweier, geometrischer Örter, das Descartessche Schema („man reduziere jede Art von Problem auf ein mathematisches Problem, zweitens auf ein algebraisches Problem und drittens man reduziere jedes algebraische Problem auf die Lösung einer einzigen Gleichung"), das Rekursionsverfahren und das Superpositionsverfahren.

Der zweite Teil hat mehr einen theoretischen Charakter. Hier finden wir solche Abschnitte: Aufgaben („was ist eine Aufgabe?"), Geometrische Darstellung des Werdegangs der Lösung, Pläne und Programme, Aufgaben in Aufgaben, Die Geburt der Idee, So sollten wir denken, Lernen lehren und lehren lernen, Erraten und wissenschaftliche Methode. Verf. formuliert die folgenden drei Prinzipien des Lehrens: (1) Man lasse die Schüler selbst so viel wie unter den gegebenen Umständen irgend tunlich ist, entdecken; (2) Das Lehren der Mathematik sei interessant. (3) Wir sollen in aufeinanderfolgenden Phasen lehren.

Die Ausstattung des Buches ist vorzüglich.

J. Berkes (Szeged)

A. I. Markushevich, Series (Fundamental concepts with historical exposition), 175 pages, Delhi, Hindustan Publishing Corporation (India), 1967.

The aim of the book is to acquaint the reader, possessing mathematical background up to an undergraduate level, in an easy and independent manner, with the concept of series, the fundamental properties of series, and the representation of elementary functions by series, without using TAYLOR's formula. It is not intended to serve as a text book, but is suitable to assist the young readers in getting interested in mathematical analysis in its early period. In fact, the book follows the path laid down by NEWTON, LEIBNIZ, EULER, CAUCHY, ABEL, D'ALEMBERT, and contains many historical comments which may be of interest to the adult mathematician readers too.

J. Németh (Szeged)

Raphaël Salem, Oeuvres mathématiques, 645 pages, Paris, Hermann, 1967.

Les travaux de RAPHAËL SALEM ont influencé d'une manière considérable le développement modernedes théories des séries de Fourier et des séries trigonométriques. Son oeuvre comprend, outre ses articles apparus dans divers périodiques, trois livres: „*Essais sur les séries trigonométriques*" (Paris, 1940), „*Algebraic numbers and Fourier series*" (Boston, Mass., 1963), „*Ensembles parfaits et séries trigonométriques*" (avec J. P. KAHANE, Paris, 1963).

Dans ce volume on a reproduit ses articles et son premier livre et cela dans le groupement suivant:

A. théorie générale des séries trigonométriques,
B. séries trigonométriques lacunaires et aléatoires,
C. ensembles parfaits et entiers algébriques,
D. ensembles parfaits et mesures singulières,
E. résultats isolés.

Une préface, par A. ZYGMUND, expose les données biographiques principales et l'activité en directions diverses de R. SALEM. Puis, dans une introduction de 21 pages, J. P. KAHANE et A. ZYGMUND donnent une revue sur ses résultats mathématiques.

Cette collection bien rédigée et de belle présentation contribuera beaucoup de faire connaître toute l'oeuvre de ce savant remarquable que fût le regretté RAPHAËL SALEM.

K. Tandori (Szeged)

Béla Sz.-Nagy—Ciprian Foiaş, Analyse harmonique des opérateurs de l'espace de Hilbert, XI+373 pages, Budapest, Akadémiai Kiadó, 1967.

After the basic general concepts of the dilation theory of operators in Hilbert space have been brought to light there appeared the question how useful the methods of this theory are in the study of the nature of general bounded operators. Most of the attention has been paid to the study of contraction operators by using their unitary dilations. The book presents what has been done up till now in this domain, mostly for a single contraction. The authors develop the method which can be called the Fourier operator analysis. The leading role in the inner contents and applications of this analysis is played by the theory of unitary dilations. It has been much influenced by the abstract prediction theory of weakly stationary stochastic processes on the one hand, and by the theory related to the concept of characteristic functions of operators, originated by Livšic (and enclosing thereby investigations of M. G. Krein's school), on the other hand. The theory of the authors gives an ingenious synthesis and simplification of both theories mentioned above, as well as a huge variety of other results, by presenting the true reasons why they succeed and work on. The point is however that Fourier operator analysis when combined with unitary dilation methods goes considerably further and creates methods which made possible the discovery and the characterization of intrinsic and essentially new geometrical and functional features of quite general contraction operators. Special attention should be paid here to the completely general functional models of contractions (Chapter VI) and the analysis of analytic contraction valued functions (Chapters V and VII), both in separable spaces, which is not an essential restriction. These achievements of the authors stand for an essentially new step forward in the general operator theory and certainly are one of the most powerful and fruitful means we get to work with.

The subtle methods developed by the authors will be certainly used with success in the future for penetrating into new domains of operator theory. One should point out the unified and elegant manner of developing the theory in the book, in contrary to some often observable tendencies of treating modern operator theory in Hilbert space as a collection of separated tricks or artificially arranged computations.

The main part of the book is a unified exposition (with many simplified proofs and new details) of original results of the authors published before in a series of papers in these *Acta* since 1953. The book contains also some entirely new results of the authors not published elsewhere or only announced in a short form. Each chapter is concluded by "Notices" in which there are included detailed references, historical remarks and short comments, and some other problems concerning the contents of the corresponding chapter. The writing of the book is precise and the exposition is clear and careful. The geometrical methods are developed in a nice harmony with the analytical means.

The first four chapters of the book deal with unitary dilations of contraction operators (mostly of a single contraction or of a one-parameter contraction semi-group) and with some of their direct applications. After giving the basic existence results (Chapter I) the authors study the geometry of the dilation space, the basic function theoretic properties of the harmonic-spectral measure of a single contraction and the interplay between the properties of a contraction and its unitary dilation (Chapter II). All this is used to build up a functional calculus for a single contraction: bounded functions are investigated in Chapter III and the unbounded ones in Chapter IV. Several applications are given concerning one-parameter contraction semigroups, fractional powers of dissipative operators and related topics. A nice part of Chapter III is devoted to a detailed study of an interesting class of contractions, called class C_0.

The most important part of the book starts with Chapter V. After giving there some basic definitions and preliminary properties of operator valued analytic bounded functions on the unit disc the authors prove two crucial lemmas about the Fourier representation of contractions which behave in a prescribed way with respect to some shift operators. The rest of Chapter V is devoted to some almost direct applications of these lemmas. Herewith are included the general Beurling invariance theorem, the Szegő type factoring of positive operator valued functions, and the factoring of contraction valued analytic functions. There is introduced a very important class of analytic contraction valued functions, the functions which admit a scalar multiple. All the above includes among others an extremely nice and general exposition of abstract prediction theory not to be found elsewhere in such a compact and elegant form.

Chapters VI and VII form the central part of the book. The first one begins with the description of the characteristic function of a contraction. Then, using the previous theorems the authors construct for an arbitrary (completely non-unitary) contraction its functional model. The true and explicit characterization of this model is performed with the help of the isometric and unitary dilations of the contraction in question. It becomes clear that the use of these dilations is in the essence of the matter. There is discussed in detail the interplay between the given contraction, its functional model and its characteristic function. The chapter finishes with a complete description of unitary dilations in terms of spectral multiplicities. In Chapter VII the authors introduce the concept of the regular factoring of contraction valued analytic functions in the unit disc. Then there are established the subtle relations between such factorings of the characteristic function of the contraction T and the invariant subspaces for T. The results are in some sense definitive and close the question how far the generalizations of Beurling type invariance theorems may be continued within the frames of the developed theory. The functional models are essentially involved and the general results are illustrated by investigations of contractions T of class C_{11} (i.e. such that $T^n h$ and $T^{*n} h$ do not tend to 0 as $n \to \infty$, for $h \neq 0$).

The two last chapters, VIII and IX, are of a more special character. Among others, they illustrate some applications of the whole theory developed before. Chapter VIII deals with the so called weak contractions, and establishes a general spectral decomposition theorem for them. Chapter IX concerns criteria for the similarity of a contraction to a unitary operator, and problems of ,,quasi-similarity" and unicellularity for contraction operators.

The book concludes with a long list of references*) and with indexes of authors quoted and of the more important terms and symbols.

The reading of the book requires merely the knowledge of the operator theory as presented in the ,,Leçons d'analyse fonctionnelle" of F. Riesz and Sz.-Nagy and some knowledge about special Hardy classes of functions analytic in the unit disc. This together with a clear and direct exposition makes the book accessible for a wide variety of mathematicians who would like to become acquainted with modern operator theory in Hilbert space. The book is an excellent help for specialists in the field and certainly will be a rich source of inspiration in further research.

W. Mlak (Kraków)

A. Rényi, Calcul des probabilités avec un appendice sur la théorie de l'information. XIII+620 pages, Paris, Dunod, 1966.

Ce livre de belle présentation est la traduction autorisée du même ouvrage publié en langue allemande en 1962. Au cours de la traduction on a corrigé quelques fautes d'impression et quelques imprécisions.

K. Tandori (Szeged)

---

*) This opportunity should be used to mention that the following two papers (although cited on pp. 54 and 224) are missing from the list of references: Nakano, H. [1] On unitary dilations of bounded operators, *Acta Sci. Math.*, **22** (1961), 286—288, Lowdenslager, D. B. [1] On factoring matrix valued functions, *Annals of Math.*, **78** (2) (1963), 450—454. Also, the exact title of Foiaş, C.—Mlak, W. [1] is: "The extended spectrum of completely non-unitary contractions and the spectral mapping theorem".

INDEX — TARTALOM

ACTA SCIENTIARUM MATHEMATICARUM
SZEGED (HUNGARIA), ARADI VÉRTANÚK TERE 1

On peut s'abonner à l'entreprise de commerce des livres et journaux „Kultúra" (Budapest, I. Fő utca 32).

INDEX: 26 024

68-5826 — Szegedi Nyomda

Felelős szerkesztő és kiadó: Szőkefalvi-Nagy Béla
A kézirat nyomdába érkezett: 1968. január hó
Megjelenés: 1968. június hó

Példányszám: 1000. Terjedelem: 19,6 (A/5) ív
Készült monószedéssel, íves magasnyomással, az MSZ 5601-24 és az MSZ 5602-55 szabvány szerint