Studia Scientiarum Mathematicarum Hungarica

AUXILIO
CONSILII INSTITUTI MATHEMATICI
ACADEMIAE SCIENTIARUM HUNGARICAE

TOMUS XIV.
FASC. 1—3.
1979

AKADÉMIAI KIADÓ, BUDAPEST

9

Studia Scientiarum Mathematicarum Hungarica

A Magyar Tudományos Akadémia matematikai folyóirata

Szerkesztőség: 1053 Budapest V., Reáltanoda u. 13—15.

Technikai szerkesztő: Merza József

Kiadja az Akadémiai Kiadó, 1054 Budapest V., Alkotmány u. 21.

A *Studia Scientiarum Mathematicarum Hungarica* angol, német, francia vagy orosz nyelven közöl eredeti értekezéseket a matematika tárgyköréből. Félévenként jelenik meg, évi egy kötetben.

Előfizetési ára belföldre 120,— Ft, külföldre 165,— Ft. Megrendelhető a belföld számára az Akadémiai Kiadónál, a külföld számára pedig a Kultúra Könyv és Hírlap Külkereskedelmi Vállalatnál (1011 Budapest I., Fő u. 32.).

Cserekapcsolatok felvétele ügyében kérjük az MTA Matematikai Kutató Intézete Könyvtárához (1053 Budapest V., Reáltanoda u. 13—15.) fordulni.

Közlésre szánt dolgozatokat kérjük két példányban a szerkesztőség címére küldeni.

Studia Scientiarum Mathematicarum Hungarica is a journal of the Hungarian Academy of Sciences publishing original papers on mathematics, in English, German, French or Russian.

It is published semiannually, making up one volume per year.

Editorial Office: 1053 Budapest V., Reáltanoda u. 13—15, Hungary.
Technical Editor: J. Merza

Orders may be placed with *Kultura* Trading Co. for Books and Newspapers, Budapest 62, P.O.B. 149 or with its representatives abroad.

For establishing exchange relations please write to the Library of the Mathematical Institute (1053 Budapest V., Reáltanoda u. 13—15.)

Papers intended for publication should be sent to the Editor in 2 copies.

Studia Scientiarum Mathematicarum Hungarica

Auxilio
Consilii Instituti Mathematici
Academiae Scientiarum Hungaricae

Redigit

L. Fejes Tóth

Adiuvantibus

Á. Császár, I. Csiszár, A. Hajnal, P. Révész, O. Steinfeld, E. T. Schmidt, J. Szabados, D. Szász, I. Vincze

Tomus XIV

Akadémiai Kiadó, Budapest
1979

STUDIA SCIENTIARUM MATHEMATICARUM HUNGARICA

Tomus XIV

INDEX

81-4202 — Szegedi Nyomda — F. v.: Dobó József igazgató

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 1—9

INTERPOLATION IN SPACES OF WEIGHTED MAXIMUM NORM

by

BARBARA HÁY and P. VÉRTESI

To Professor O. KIS *on his 50-th birthday*

1. Introduction

Hermite—Fejér interpolation are investigated in spaces of weighted maximum norm. A general convergence criterium is established, which will be applied for the Jacobi roots. We give some estimations for the corresponding Lagrange interpolations, too.

2. Preliminaries. Definitions

Considering the pointmatrix $X=\{x_{kn}\}$

$$(2.1)\qquad -1 \leqq x_{nn} < x_{n-1,n} < \ldots < x_{2n} < x_{1n} \leqq 1,$$

we can investigate the convergence behaviour of the Hermite—Fejér parabolas

$$(2.2)\qquad \begin{cases} H_n(f, x) = H_n(f, X, x) = \sum_{k=1}^{n} f(x_{kn})\, h_{kn}(x), \\ H_n(f, d, x) = H_n(f, d, X, x) = H_n(f, x) + \sum_{k=1}^{n} d_{kn} \mathfrak{h}_{kn}(x) \qquad (n = 1, 2, \ldots) \end{cases}$$

(where, as usual,

$$h_{kn}(x) = h_{kn}(X, x) = v_{kn}(x)\, l_{kn}^2(x), \quad v_{kn}(x) = v_{kn}(X, x) = 1 - 2l'_{kn}(x_{kn})(x - x_{kn}),$$

$$l_{kn}(x) = l_{kn}(X, x) = \Omega_n(x)\, [\Omega'_n(x_{kn})(x - x_{kn})]^{-1},$$

$$\Omega_n(x) = \Omega_n(X, x) = c_n \prod_{k=1}^{n} (x - x_{kn}), \quad \mathfrak{h}_{kn}(x) = \mathfrak{h}_{kn}(X, x) = (x - x_k)\, l_{kn}^2(x),$$

d_{kn} are real numbers). If $X = X^{(\alpha, \beta)}$, i.e., $x_{kn} = x_{kn}^{(\alpha, \beta)}$ is the k-th root of the n-th Jacobi polynomials $P_n^{(\alpha, \beta)}(x)$ $(\alpha, \beta > -1)$ then, if $-1 < \alpha, \beta < 0$

$$(2.3)\qquad \lim_{n \to \infty} \|H_n(f, x) - f(x)\| = 0 \quad \text{if} \quad f \in C,$$

where C denotes the continuous functions on $[-1, 1]$, $\|\cdot\|$ is the max-norm on $[-1, 1]$ (see e.g. [1], 14.6 and 4.1).

Our aim is to prove an estimation similar to (2.3) for any $\alpha, \beta > -1$ in another (semi) normed spaces of continuous functions.

DEFINITIONS. Let $w(x)\in C$, $w(x)\geqq 0$, $w(x)\not\equiv 0$ and

(2.4) $$\|f\|_w = \max_{-1\leqq x\leqq 1} w(x)|f(x)| \quad \text{if} \quad f\in C.$$

Let us denote by C_w the linear seminormed space of continuous functions endowed by the seminorm (2.4). If $u_{kn}(x)\in C$, further $w(x)>0$ whenever $x\in X$, one can consider the linear bounded operators $U_n=U_n(X, w)$ from C_w to C_w defined by $\{u_{kn}\}$ as follows

(2.5) $$U_n f(x) = \sum_{k=1}^{n} f(x_{kn})u_{kn}(x) \qquad (n=1, 2, \ldots).$$

U_n are bounded indeed, because by usual arguments

(2.6) $$\|U_n\|_w = \sup_{\|f\|_w=1} \|U_n f\|_w = \max_{-1\leqq x\leqq 1} w(x)\sum_{k=1}^{n}\frac{|u_k(x)|}{w(x_k)}$$

(compare [2], 1). Using the obvious fact, that

(2.7) $$\|g\|_w \leqq \|g\|\cdot\|w\|,$$

we immediately obtain that

(2.8) $$\lim_{n\to\infty} \|H_n(f, x)-f(x)\|_w = 0 \qquad (f\in C),$$

whenever (2.3) is true. In the next we prove (2.8) for certain X when (2.3) is not valid.

3. Results

3.1. Let us consider a matrix X, a function $w(x)\in C$, $w(x)\geqq 0$, $w(x)>0$ if $x\in X$, and the operator-sequences H_n and D_n of type (2.5) defined by $\{h_{kn}(X, x)\}$ and $\{\mathfrak{h}_{kn}(X, x)\}$, respectively (see (2.2)). It is easy to prove the next

THEOREM 3.1. *If for the above X and $w(x)$*

(3.1) $$\mu_n(w, X) \stackrel{\text{def}}{=} \|H_n\|_w = O(1),$$

moreover, with $M_n=\max\left(1, \max_{1\leqq k\leqq n} |d_{kn}|\right)$ *and* $v_n(w, X) \stackrel{\text{def}}{=} \|D_n\|_w$

(3.2) $$M_n\cdot v_n(w, X) = o(1),$$

then

(3.3) $$\lim_{n\to\infty} \|H_n(f, d, X, x)-f(x)\|_w = 0 \quad \textit{when} \quad f\in C.$$

3.2. To apply this theorem we prove the following estimations which are in teresting in themselves.

THEOREM 3.2. *If* $X=X^{(\alpha,\beta)}$ *and*

$$w(x) = w(\alpha+1/2, \beta+1/2, x) = (1-x)^{\alpha+1/2}(1+x)^{\beta+1/2} \qquad (\alpha, \beta\geqq -1/2),$$

then we have as follows:

(3.4) $$\mu_n(w, X) = O(1),$$

(3.5) $$\nu_n(w, X) = (1+o(1))\frac{2}{\pi}\frac{\ln n}{n} \qquad (n \geqq 2).$$

By these statements we immediately obtain the analogous of (2.3).

COROLLARY 3.3. *Supposing* $X = X^{(\alpha,\beta)}$, $w(x) = w(\alpha+1/2, \beta+1/2, x)$ $(\alpha, \beta \geqq -1/2)$ *and* $|d_{kn}| = o(n/\ln n)$, *we have* (3.3).

3.3. Let us consider now the corresponding problem for the Lagrange interpolation

(3.6) $$L_n(f, x) = L_n(f, X, x) = \sum_{k=1}^{n} f(x_k)\, l_k(x).$$

By considerations similar to 4.1—4.7, we shall prove

THEOREM 3.4. *If* $X = X^{(\alpha,\beta)}$ *and* $w(x) = w(\alpha/2+1/4, \beta/2+1/4, x)$ $(\alpha, \beta \geqq -1/2)$, *then*

(3.7) $$\lambda_n(w, X) \overset{\text{def}}{=\!=} \|L_n\|_w = \begin{cases} (1+o(1))\dfrac{2}{\pi}\ln n & \text{if} \quad -1/2 \leqq \alpha, \beta \leqq 1/2, \\ (1+o(1))\delta(\eta)\ln n & \text{if} \quad \eta = \max(\alpha, \beta) > 1/2, \end{cases}$$

where $\delta(\eta) > 2/\pi$.

3.4. In her paper [2] I. MELINDER claimed that $\lambda_n(w, X) = (1+o(1))(2/\pi)\ln n$ for *any* $\alpha, \beta \geqq -1/2$. Unfortunately, the proof is not correct (see e.g. [2], page 41, rows 12—18). Let us remark that the constant $\delta(\eta)$ comes from the fact that (4.11), contrary to (4.16), holds on the *whole* interval $[0, \pi]$.

4. Proofs

4.1. PROOF of Theorem 3.1. By a suitable polynomial $p_s(x)$ of degree $\leqq 2n-1$ we can write

$$\|H_n(f, d, x) - f(x)\|_w =$$

$$= \|H_n(f, x) - H_n(p_s, x) + \sum_{k=1}^{n} [d_k - p_s'(x_k)]\eta_k(x) + p_s(x) - f(x)\|_w \leqq$$

$$\leqq \mu_n \|f - p_s\|_w + [\|p_s'\|_w + M_n]\nu_n + \|f - p_s\|_w \leqq \varepsilon,$$

which was stated (compare [3]).

4.2. PROOF of Theorem 3.2. First we quote some relations. If $x = \cos\vartheta$, $x_{kn}^{(\alpha,\beta)} = \cos\vartheta_{kn}^{(\alpha,\beta)}$, $x_{0n} = 1$, $x_{n+1,n} = -1$, $\vartheta_{0n} = 0$, $\vartheta_{n+1,n} = \pi$ then omitting the superfluous notations

$$\vartheta_{k+1}-\vartheta_k \sim \frac{1}{n}, \quad \vartheta_k \sim \frac{k}{n} \qquad (0 \leqq \vartheta_k < \pi - \varepsilon), \tag{4.1}$$

$$|x-x_k| \sim \frac{|j-k|(j+k)}{n^2}, \qquad k=1,2,\dots,n,\ k \neq j, \tag{4.2}$$

where x_j is the nearest root to x,

$$\vartheta_{k,n} = \frac{\left(k-s_n-\frac{1}{2}\right)\pi-\gamma}{N}+\varepsilon_{k,n} \tag{4.3}$$

if $k>|s_n|$, $0<\vartheta_k \leqq \frac{\pi}{2}$, $n \geqq n_0$, $|s_n|<M_0$ and $|\varepsilon_{kn}|<\frac{c_0}{kn}$,

$$|P_n(x)| \sim |x-x_j|\,\vartheta_j^{-\alpha-3/2}n^{1/2} \sim |\vartheta-\vartheta_j|\,\vartheta_j^{-\alpha-1/2}n^{1/2} \tag{4.4}$$

uniformly in $x \in [-1+\varepsilon, 1]$.

$$P_n(x) = \frac{\cos(N\vartheta+\gamma)+O\left((n\sin\vartheta)^{-1}\right)}{\sqrt{n\pi}\left(\sin\frac{\vartheta}{2}\right)^{\alpha+1/2}\left(\cos\frac{\vartheta}{2}\right)^{\beta+1/2}}, \qquad \frac{c}{n} \leqq \vartheta \leqq \pi-\frac{c}{n}, \tag{4.5}$$

$c>0$ is arbitrary,

$$|P_n'(x_k)| \sim k^{-\alpha-3/2}n^{\alpha+2}, \qquad 0 \leqq \vartheta_k \leqq \pi-\varepsilon, \tag{4.6}$$

$$P_n'(x_k) = (-1)^{k-1}\sqrt{\frac{n}{\pi}}\,\frac{1+O(k^{-1})}{2\left(\sin\frac{\vartheta_k}{2}\right)^{\alpha+3/2}\left(\cos\frac{\vartheta_k}{2}\right)^{\beta+3/2}} \tag{4.7}$$

uniformly in k if $k>M_0$ and $0<\vartheta_k \leqq \pi-\varepsilon$,

$$\lim_{n\to\infty}\sqrt{n}\,R_n = \begin{cases} 1/\sqrt{\pi} & \text{if } \ -1/2 \leqq \alpha \leqq 1/2, \\ \varrho(\alpha) & \text{where } \ \varrho(\alpha) > 1/\sqrt{\pi} \ \text{ if } \ \alpha > 1/2. \end{cases} \tag{4.8}$$

Here

$$R_n = \max_{0 \leqq \vartheta \leqq \pi/2}\left|\left(\sin\frac{\vartheta}{2}\right)^{\alpha+1/2}\left(\cos\frac{\vartheta}{2}\right)^{\beta+1/2}P_n(\cos\vartheta)\right|,$$

moreover $\alpha,\beta>-1$, $P_n(x)=P_n^{(\alpha,\beta)}(x)$, $\varepsilon>0$, is fixed, s_n is integer, c_0 and M_0 are suitable constants depending on α and β, $N=n+(\alpha+\beta+1)/2$ and $\gamma=-(\alpha+1/2)\pi$ (see [1], (8.21.18), (8.9.2), (7.32.9) ((4.5), (4. 6), (4. 8)), [4] ((4.1), (4.2)); [5], Lemma 3 ((4.4)) and [6] Lemma 4.1 ((4.3), (4.7)); for the symbol "$\sim$" see [1], 1.1). In connection with (4.8) see further [7], pp. 59. By $P_n^{(\alpha,\beta)}(x)=(-1)^nP_n^{(\beta,\alpha)}(-x)$ we can use the corresponding formulae in $[\varepsilon, \pi]$, too.

4.3. First we prove (3.5). For this aim let $\varphi_n=[\ln n]$ and $\tau_k=\vartheta_k-\varepsilon_k$. We state

LEMMA 4.1. *We have when* $n \to \infty$

$$(4.9) \qquad \frac{\sin \vartheta_k}{\cos \vartheta - \cos \vartheta_k} = \big(1+o(1)\big) \frac{\sin \tau_k}{\cos \vartheta - \cos \tau_k} \quad \text{if} \quad \varphi_n \leqq k \leqq n-\varphi_n, \ k \neq j$$

uniformly in k and $0 \leqq \vartheta \leqq \pi$.

Indeed, by $\sin \vartheta_k = \sin \tau_k + \cos \xi_k (\vartheta_k - \tau_k)$ and

$$\cos \vartheta_k = \cos \tau_k - \sin \eta_k (\vartheta_k - \tau_k), \quad \xi_k, \eta_k \in [\vartheta_k, \tau_k],$$

$$\frac{\dfrac{\sin \vartheta_k}{\cos \vartheta - \cos \vartheta_k}}{\dfrac{\sin \tau_k}{\cos \vartheta - \cos \tau_k}} = \frac{1 + \dfrac{\cos \xi_k}{\sin \tau_k} (\vartheta_k - \tau_k)}{1 + \dfrac{\sin \eta_k}{\cos \vartheta - \cos \tau_k} (\vartheta_k - \tau_k)}.$$

Let, e.g., $0 < \vartheta_k \leqq \pi/2$. Then $\cos \xi_k (\sin \tau_k)^{-1} = O(n/k)$ and

$$(4.10) \qquad \frac{\sin \eta_k}{|\cos \vartheta - \cos \tau_k|} \leqq \frac{\sin \eta_k + \sin \vartheta}{|\cos \vartheta - \cos \tau_k|} = \frac{O(1)}{\sin \dfrac{|\vartheta - \tau_k|}{2}} = O\left(\frac{n}{|j-k|}\right).$$

Using these and $|\vartheta_k - \tau_k| = O(n^{-1} k^{-1})$, we obtain (4.9).
By this lemma we can prove the following. Let

$$S_p(\vartheta) \overset{\text{def}}{=} \sum_{k=\varphi_n}^{n-\varphi_n}{}' \frac{\sin^p \vartheta_k}{|\cos \vartheta - \cos \vartheta_k|} \qquad (p = 1, 2),$$

where Σ' means that $k \neq j$.

LEMMA 4.2. *We have uniformly in* $0 \leqq \vartheta \leqq \pi$

$$(4.11) \qquad S_1 = \big(1+o(1)\big) \frac{2}{\pi} n \ln n.$$

Indeed, by (4.9) we have to prove (4.10) with τ_k instead of ϑ_k.
By symmetry we can suppose $0 \leqq \vartheta \leqq \pi/2$. Let

$$(4.12) \qquad g(t) = \frac{\sin t}{|\cos \vartheta - \cos t|}.$$

It is easy to see that $g(t)$ is monotone increasing if $0 \leqq t \leqq \vartheta \leqq \pi/2$. So

$$g\left(t - \frac{\pi}{N}\right) \leqq g(\tau_k) \leqq g(t) \quad \text{if} \quad t \in [\tau_k, \tau_{k+1}], \ \varphi_n \leqq k < j-1$$

from where

$$(4.13) \qquad \frac{N}{\pi} \int_{\tau_{k-1}}^{\tau_k} g(t)\, dt \leqq g(\tau_k) \leqq \frac{N}{\pi} \int_{\tau_k}^{\tau_{k+1}} g(t)\, dt \quad \text{if} \quad \varphi_n \leqq k < j-1.$$

If $k>j+1$, by $\sin\eta=\sin\tau_k+(\eta-\tau_k)\cos\xi$ we have

$$g\left(t+\frac{\pi}{N}\right)+\frac{\left(\tau_k-t-\dfrac{\pi}{N}\right)\cos\xi_1}{\cos\vartheta-\cos\left(t+\dfrac{\pi}{N}\right)}\leqq g(\tau_k)\leqq g(t)+\frac{(\tau_k-t)\cos\xi_2}{\cos\vartheta-\cos t}\quad\text{if}\quad t\in[\tau_{k-1},\tau_k],$$

$j+1<k\leqq n-\varphi_n$, from where uniformly for these k

$$\frac{N}{\pi}\int_{\tau_k}^{\tau_{k+1}}g(t)\,dt+O\left(\frac{N}{k^2-j^2}\right)\leqq g(\tau_k)\leqq\frac{N}{\pi}\int_{\tau_{k-1}}^{\tau_k}g(t)\,dt+O\left(\frac{N}{k^2-j^2}\right).\tag{4.14}$$

By (4.13) and (4.14), using that $g(\tau_k)=O(N)$ $(k\neq j)$, we obtain by $c_n\sim 1$

$$\sum_{k=\varphi_n}^{n-\varphi_n}{}' g(\tau_k)=\frac{N}{\pi}\ln\frac{\cos\tau_{\varphi_n}-\cos\vartheta}{\cos\tau_{j-1}-\cos\vartheta}\cdot\frac{\cos\vartheta-\cos\tau_{n-\varphi_n}}{\cos\vartheta-\cos\vartheta_{j+1}}+O(N)=$$

$$=\frac{N}{\pi}\ln c_n\frac{j^2-\varphi_n^2}{j}\cdot\frac{(n-\varphi_n)^2-j^2}{j}+O(N)=\big(1+o(1)\big)\frac{2}{\pi}n\ln n$$

uniformly in n if $j\geqq\varphi_n+2$. Similar considerations hold for the remaining values of j.

4.4. Let us consider the case $p=2$.

LEMMA 4.3. *We have*

$$S_2(\vartheta)\leqq\big(1+o(1)\big)\frac{2}{\pi}n\ln n\quad\textit{uniformly in}\quad 0\leqq\vartheta\leqq\pi,\tag{4.15}$$

moreover,

$$S_2(\vartheta)=\big(1+o(1)\big)\frac{2}{\pi}n\ln n\tag{4.16}$$

uniformly in ϑ *if* $n|\vartheta-\pi/2|=o(\varphi_n)$.

Indeed, the upper estimation is trivial by $s_2\leqq s_1$ and (4.11). On the other hand, we can write for the mentioned ϑ

$$S_2(\vartheta)\geqq\left(\sum_{k=\varphi_n}^{j-\varphi_n}+\sum_{k=j+\varphi_n}^{n-\varphi_n}\right)\frac{\sin^2\vartheta_k}{|\cos\vartheta-\cos\vartheta_k|}=$$

$$=\big(1+o(1)\big)\left(\sum_{k=\varphi_n}^{j-\varphi_n}+\sum_{k=j+\varphi_n}^{n-\varphi_n}\right)\frac{\sin^2\vartheta_k}{|\cos\vartheta_k|}=\big(1+o(1)\big)(I_1+I_2).$$

Here by a suitable $M_1>0$ (see (4.3))

$$I_1=\sum_{k=\varphi_n}^{j-\varphi_n}\left(\frac{1}{\cos\vartheta_k}-\cos\vartheta_k\right)\geqq\sum_{k=\varphi_n}^{j-\varphi_n}\frac{1}{\cos\vartheta_k}+O(n)\geqq$$

$$\geqq\sum_{k=\varphi_n}^{j-\varphi_n}\frac{1}{\cos\left(\frac{k}{N}\pi-\frac{M_1}{N}\right)}+O(n)\geqq\sum_{k=\varphi_n}^{j-\varphi_n}\frac{1}{\frac{\pi}{2}-\frac{k\pi}{N}+\frac{M_1}{N}}+O(n)=$$

$$=\frac{N}{\pi}\sum_{k=\varphi_n}^{j-\varphi_n}\frac{1}{\frac{N}{2}-k+\frac{M_1}{\pi}}+O(n)=\big(1+o(1)\big)\frac{1}{\pi}n\ln n.$$

Using similar argument for I_2, we obtain (4.16).

4.5. To prove (3.5), we have by (2.6)

$$v_n=\max_{-1\leqq x\leqq 1}w(\alpha+1/2,\beta+1/2,x)\sum_{k=1}^{n}\frac{|x-x_k|\,l_k^2(x)}{w(\alpha+1/2,\beta+1/2,x_k)}\overset{\text{def}}{=\!=}\max_{-1\leqq x\leqq 1}B_n(x),$$

where $x_k=x_k^{(\alpha,\beta)}$, $l_k(x)=l_{kn}^{(\alpha,\beta)}(x)$ and $\alpha,\beta\geqq-1/2$. We can suppose $0\leqq\vartheta\leqq\pi/2$. First we state

$$B_n(x)=o\left(\frac{\ln n}{n}\right)\quad\text{uniformly in }\vartheta\text{, if}\quad 0\leqq\vartheta\leqq\varphi_n. \tag{4.17}$$

Indeed, by (4.1), (4.2), (4.4) and (4.6)

$$B_n(x)=O(1)\left\{\frac{1}{n^2}\sum_{k=1}^{\left[\frac{3n}{n}\right]}{}'\frac{k^2}{|k+j|\,|k-j|}+\frac{1}{n^4}\sum_{k=\left[\frac{3n}{4}\right]+1}^{n}(n-k)^2+\frac{j}{n^2}\right\}=o\left(\frac{\ln n}{n}\right).$$

On the other hand, let $\vartheta_{\varphi_n}\leqq\vartheta\leqq\pi/2$. Then by (4.5) and (4.7)

$$\begin{aligned}B_n(x)&=\frac{[\cos(N\vartheta+\gamma)+O((n\sin\vartheta)^{-1})]^2}{n^2}\left\{\sum_{k=\varphi_n}^{\left[\frac{n}{2}\right]}{}'\frac{\sin^2\vartheta_k}{|\cos\vartheta-\cos\vartheta_k|}\cdot\right.\\&\left.\cdot\frac{1}{[1+O(k^{-1})]^2}+\sum_{k=\left[\frac{n}{2}\right]+1}^{n-\varphi_n}\frac{\sin^2\vartheta_k}{|\cos\vartheta-\cos\vartheta_k|}\,\frac{1}{[1+O(n-k)^{-1}]^2}\right\}+\\&+\sum_{k=1}^{\varphi_n}{}'\ldots+\sum_{k=n-\varphi_n+1}^{n}\ldots+\sum_{k=j}\ldots=I_1+I_2+I_3+I_4+I_5.\end{aligned}\tag{4.18}$$

It is easy to see that now

$$I_3+I_4+I_5=O(n^{-1}),$$

moreover

$$I_1+I_2=\frac{[\cos(N\vartheta+\gamma)+o(1)]^2}{n^2}S_2(\vartheta).$$

I.e., by (4.15), $I_1+I_2\leqq(1+o(1))(2/\pi)(\ln n/n)$, further, choosing the values of ϑ such that $|\cos(N\vartheta+\gamma)|=1$ and $0\leqq(\pi/2-\vartheta)n\leqq K$ if $n\geqq n_0$ (K is fixed; obviously $\vartheta=\vartheta(n)$), by (4.16) we obtain that

$$I_1+I_2=\frac{1+o(1)}{n^2}S_2(\vartheta)=\big(1+o(1)\big)\frac{2}{\pi}\frac{\ln n}{n}. \tag{4.19}$$

By (4.17)—(4.19) we obtain (3.5):

4.6. To prove (3.4), we use (4.1), (4.2), (4.4), (4.6) and [1] (14.5.2) which says

$$v_k(x)=\frac{1-xx_k+(\alpha-\beta)(x_k-x)+(\alpha+\beta+1)x_k(x_k-x)}{1-x_k^2}.$$

We obtain, if e.g. $0\leqq\vartheta\leqq\pi/2$, as follows:

$$\mu_n=O(1)\left[\frac{1}{n^2}\sum_{k=1}^{n}{}'\frac{1-xx_k}{(x-x_k)^2}+\frac{1}{n^2}\sum_{k=1}^{n}{}'\frac{\alpha-\beta+(\alpha+\beta+1)x_k}{x_k-x}\right]+O(1)=$$
$$=O(1)(I_1+I_2)+O(1).$$

Here by $1-xx_k\leqq1-\cos\vartheta\cos\vartheta_k+\sin\vartheta\sin\vartheta_k=2\sin^2(\vartheta+\vartheta_k)/2$,

$$I_1=O\left(\frac{1}{n^2}\right)\sum_{k=1}^{n}{}'\frac{1}{\sin^2\dfrac{\vartheta-\vartheta_k}{2}}=O(1)$$

and

$$I_2=O\left(\frac{1}{n^2}\right)\sum_{k=1}^{n}{}'\frac{1}{\sin\dfrac{\vartheta-\vartheta_k}{2}\sin\dfrac{\vartheta+\vartheta_k}{2}}=O(1),$$

which give (3.4).

4.7. Proof of Theorem 3.4. The proof is analogous to the above one so we only sketch it.

By definition

$$\lambda_n=\max_{-1\leqq x\leqq1}\sum_{k=1}^{n}\frac{w(x)|P_n(x)|}{|x-x_k|\,|P_n'(x_k)|\,w(x_k)}\overset{\text{def}}{=\!=}\max_{-1\leqq x\leqq1}Q_n(x).$$

We write

$$Q_n(x)=\sum_{k=1}^{\varphi_n-1}{}'+\sum_{k=\varphi_n}^{n-\varphi_n}{}'+\sum_{k=\varphi_n+1}^{n}{}'+\sum_{k=j}=I_1+I_2+I_3+I_4.$$

By (4.1), (4.4) and (4.6)

$$I_1=O\left(\frac{1}{n^2}\right)\sum_{k=1}^{\varphi_n}{}'\frac{k}{|\cos\vartheta-\cos\vartheta_k|}=O(1)\sum_{k=1}^{\varphi_n}\frac{k}{|k+j|\,|k-j|+1}=O(\ln\varphi_n).$$

Similar estimation holds for I_3, further $I_4=O(1)$. To handle $I_2=I_2(x)$, let us denote by $z_n=\cos\xi_n$, $0\leqq\xi_n\leqq\pi/2$, that point for which

$$w(z_n)|P_n(z_n)| = 2^{\frac{\alpha+\beta+1}{2}} R_n \tag{4.20}$$

(see (4.8)). Then, by (4.7), (4.8) and (4.11), we get for $0\leqq x\leqq 1$

$$I_2(x) \leqq w(z_n)|P_n(z_n)| \max_{0\leqq x\leqq 1} \sum_{k=\varphi_n}^{n-\varphi_n}{}' \frac{1}{|P_n'(x_k)|\,|x-x_k|\,w(x_k)} =$$

$$= (1+o(1))\sqrt{\pi}\,\frac{\sqrt{n}\,R_n}{n} \max_{0\leqq x\leqq 1} \sum_{k=\varphi_n}^{n-\varphi_n}{}' \frac{\sin\vartheta_k}{|\cos\vartheta-\cos\vartheta_k|} =$$

$$= (1+o(1))\sqrt{\pi}\sqrt{n}\,R_n\,\frac{2}{\pi}\ln n = \begin{cases} (1+o(1))\dfrac{2}{\pi}\ln n & \text{if} \quad -\dfrac{1}{2}\leqq\alpha\leqq 1/2 \\ (1+o(1))\,\delta(\alpha)\ln n & \text{if} \quad \alpha>1/2. \end{cases}$$

Here $\delta(t)=2\varrho(t)/\sqrt{\pi}$.

Using that the right sides are attained if $x=z_n$, we obtain

$$\max_{0\leqq x\leqq 1} I_2(x) = (1+o(1))\,I_2(z_n).$$

Applying similar argument for $-1\leqq x\leqq 0$, we obtain (3.7).

REFERENCES

[1] Szegő, G., *Orthogonal Polynomials,* AMS Coll. Publ. XXIII, 1974.
[2] Melinder, Ingrid, Accurate Approximation in Weighted Maximum Norm by Interpolation, *J. Approximation Theory* **22** (1978), 33—45.
[3] Grünwald, G., On the theory of interpolation, *Acta Math.* **75** (1942), 219—245.
[4] Natanson, G. I., Two sided estimate for the Lebesgue function of the Lagrange interpolation with Jacobi nodes, *Izv. Vysš. Učebn. Zaved. Matematika,* **11** (1967), 67—74 (in Russian).
[5] Vértesi, P., On Lagrange interpolation, *Period. Math. Hungar.* **12** (1981), 103—112.
[6] Vértesi, P., Lagrange interpolation for functions of bounded variations, *Acta Math. Acad. Sci. Hungar.* **35** (1980), 23—31.
[7] Bernstein, S. N., *Collected Works* II, Izd. AN SSSR, 1954 (in Russian).

(Received January 7, 1980)

MTA SZÁMÍTÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI
KUTATÓ INTÉZETE
KENDE U. 13.
H—1502 BUDAPEST
HUNGARY

MTA MATEMATIKAI KUTATÓ INTÉZETE
REÁLTANODA U. 13—15.
H—1053 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 11—14

ON WAGNER SPACES OF W-SCALAR CURVATURE

by

MASAO HASHIGUCHI and TÜNDE VARGA

Dedicated to Professor Dr. Makoto Matsumoto on the occasion of his sixtieth birthday

Introduction. S. NUMATA [4] and T. VARGA [5] have independently proved that if an $n(\geqq 3)$-dimensional Finsler space is a Berwald space of scalar curvature K, then it is a Riemannian space of constant curvature K, or a locally Minkowski space according as $K \neq 0$, or $K=0$. The purpose of the present paper is to examine Wagner spaces of W-scalar curvature and to prove a similar theorem.

The authors' attention was drawn to this problem by Professor Dr. M. MATSUMOTO.

Throughout the present paper we shall use the terminology and notations of Matsumoto's monograph [3].

§ 1. Wagner connections and Wagner spaces. V. WAGNER [6] called a Finsler space a generalized Berwald space if it is possible to introduce a generalized Cartan connection such that the connection coefficients $F_j{}^i{}_k$ depend on the position (x^i) alone. Nowadays these spaces are called Wagner spaces [1]. Wagner connections and Wagner spaces are defined as follows:

DEFINITION 1. Let $s_i(x)$ be a covariant vector field on a Finsler space F. The *Wagner connection* $W\Gamma$ of the Finsler space F is a Finsler connection $(F_j{}^i{}_k, N^i{}_k, C_j{}^i{}_k)$ uniquely determined by the five axioms:

(C1) $W\Gamma$ is v-metrical, i.e., $g_{ij}|_k=0$.

(C2) The $(v)v$-torsion tensor S^1 of $W\Gamma$ vanishes, i.e.,

$$C_j{}^i{}_k = C_k{}^i{}_j.$$

(C3) $W\Gamma$ is h-metrical, i.e., $g_{ij|k}=0$.

(C4) The $(h)h$-torsion tensor T of $W\Gamma$ is given by

$$T_j{}^i{}_k \equiv F_j{}^i{}_k - F_k{}^i{}_j = \delta^i_j s_k - \delta^i_k s_j.$$

(C5) The deflection tensor D of $W\Gamma$ vanishes, i.e.,

$$D^i{}_k \equiv y^j F_j{}^i{}_k - N^i{}_k.$$

(C1) and (C2) show that C_{ijk} are the same as in the Cartan connection $C\Gamma$.

DEFINITION 2. A Finsler space is called a *Wagner space* if there exists a vector field s_i such that $F_j{}^i{}_k$ of the Wagner connection with respect to s_i are functions of x^i only.

§ 2. Wagner spaces of W-scalar curvature. We consider one of the Bianchi identities ([3] (11.1′)) of a $W\Gamma$:

$$\sigma_{(ijk)}\{T_i{}^h{}_r T_j{}^r{}_k + T_i{}^h{}_{j|k} + C_i{}^h{}_r R^r{}_{jk} - R_i{}^h{}_{jk}\} = 0, \tag{2.1}$$

where

$$R^r{}_{jk} \equiv \delta_k N^r{}_j - \delta_j N^r{}_k \quad \left(\delta_k \equiv \frac{\partial}{\partial x^k} - \frac{\partial}{\partial y^a} N^a{}_k\right) \tag{2.2}$$

are components of the $(v)h$-torsion tensor, $R_i{}^h{}_{jk}$ are components of the h-curvature tensor, and $\sigma_{(ijk)}$ means the cyclic permutation of indices and summation for the expression in the brackets behind it.

First we derive the equation (2.4) which holds for some special Wagner connection of W-scalar curvature K, and will be needed in our subsequent proof. From (C4) we obtain

$$T_i{}^h{}_r T_j{}^r{}_k = -s_i T_j{}^h{}_k,$$

$$s_h T_j{}^h{}_k = 0.$$

and thus

$$\sigma_{(ijk)}\{T_i{}^h{}_r T_j{}^r{}_k\} = 0,$$

$$\sigma_{(ijk)}\{T_i{}^h{}_{j|k}\} = \sigma_{(ijk)}\{\delta_j^h(s_{k|i} - s_{i|k})\},$$

$$s_{i|j} - s_{j|i} = \frac{\partial s_i}{\partial x^j} - \frac{\partial s_j}{\partial x^i}.$$

Therefore, if s_i is a gradient vector, i.e., $s_i = \dfrac{\partial s}{\partial x^i}$, then (2.1) gets the form

$$\sigma_{(ijk)}\{C_i{}^h{}_r R^r{}_{jk} - R_i{}^h{}_{jk}\} = 0.$$

Since $W\Gamma$ is h-metrical, we get $R_{ihjk} = -R_{hijk}$. Thus (2.1) is brought to the form

$$\sigma_{(ijk)}\{C_{hir} R^r{}_{jk} + R_{hijk}\} = 0,$$

which implies $\sigma_{(ijk)}\{R_{0ijk}\} = 0$. Then on account of [3] Theorem 13.3 we obtain $\sigma_{(ijk)}\{R_{ijk}\} = 0$, which implies $R_{i0j} = R_{j0i}$.

Now we consider the equation

$$R_{i0k} X^i X^k = KL^2 h_{ik} X^i X^k, \tag{2.3}$$

where R_{i0k} is formed from the coefficients of the Wagner connection, L is the metric function, $h_{ik} = g_{ik} - l_i l_k$ is the angular metric tensor, and K, called the *W-sectional curvature,* is a function of x, y and of the vector field X. ((2.3) has the same form as (26.1) in [3] for the Cartan connection $C\Gamma$.)

DEFINITION 3. If the W-sectional curvature $K(x, y, X)$ in (2.3) is a scalar field which does not depend on X^i, then the Wagner connection is called of *W-scalar curvature K.*

Consequently, if we consider a Wagner connection $W\Gamma$ with a gradient vector field and of W-scalar curvature K, then the symmetric property of R_{i0j} leads us to the equation

$$R_{i0k} = KL^2 h_{ik}, \tag{2.4}$$

which is the same as the well-known equation in a Finsler space of scalar curvature with the Cartan connection.

§ 3. Berwald spaces conformal to Wagner spaces. Let a Finsler space F_n be a Wagner space with respect to a gradient vector field $s_i(x)=\dfrac{\partial s}{\partial x^i}$. The metric function of F_n is L, the metric tensor is g_{ij}, and the Wagner connection is given by $(F_j{}^i{}_k, N^i{}_k, C_j{}^i{}_k)$. Let us consider the Finsler space F_n^* with the metric function $L^*=e^{-s(x)}L$. We define

$$F_j^{*i}{}_k = F_j{}^i{}_k - \delta_j^i s_k, \tag{3.1}$$

$$N^{*i}{}_k = N^i{}_k - y^i s_k, \tag{3.2}$$

$$C_j^{*i}{}_k = C_j{}^i{}_k. \tag{3.3}$$

These $(F_j^{*i}{}_k, N^{*i}{}_k, C_j^{*i}{}_k)$ form the Cartan connection on F_n^*. Namely, denoting quantities in F_n^* by an asterisk we get

$$g_{ij}^* = e^{-2s} g_{ij}, \tag{3.4}$$

and then we can easily check the validity of the relations $g^*_{ij|k}=0$ with respect to $F_j^{*i}{}_k$ and $g^*_{ij}|_k=0$ with respect to $C_j^{*i}{}_k$, $D^{*i}{}_k=0$, and moreover the symmetry of $F_j^{*i}{}_k$ and $C_j^{*i}{}_k$ in j and k. These together mean that $(F_j^{*i}{}_k, N^{*i}{}_k, C_j^{*i}{}_k)$ is the Cartan connection on F_n^*. (See [3] Definition and Theorem 17.2.)

Since the $F_j{}^i{}_k$ of the Wagner connection are functions of x^i alone, so are the $F_j^{*i}{}_k$. Hence, F_n^* is a Berwald space. (See also Theorem B of M. Hashiguchi and Y. Ichijyō [2].)

We show that the Berwald space F_n^* is of scalar curvature if the Wagner space F_n is of W-scalar curvature. By virtue of (3.4) we have

$$R^*_{i0k} = g^*_{ir} R^{*r}{}_{0k} = e^{-2s} g_{ir} R^{*r}{}_{0k}.$$

Since s_i is gradient and $N^i{}_k$ is $(1)p$-homogeneous, from (2.2) and (3.2) we get

$$R^{*r}{}_{jk} = R^r{}_{jk},$$

and thus

$$R^*_{i0k} = e^{-2s} R_{i0k}.$$

Assuming that F_n is of W-scalar curvature, (2.3) yields (2.4), and we obtain

$$R^*_{i0k} = e^{-2s} KL^2 h_{ik} = KL^{*2} h_{ik}.$$

Furthermore

$$h^*_{ik} = g^*_{ik} - \frac{\partial L^*}{\partial y^i}\frac{\partial L^*}{\partial y^k} = e^{-2s}(g_{ik} - l_i l_k) = e^{-2s} h_{ik}.$$

Thus we have

$$R^*_{i0k} = e^{2s} KL^{*2} h^*_{ik},$$

namely, the Berwald space F_n^* is of scalar curvature $K^*=e^{2s}K$.

Thus, by the theorem of S. Numata ([4] Theorem 2) and T. Varga ([5] Theorem 11), this Berwald space F_n^* for $n\geqq 3$ is a Riemannian space of constant curvature K^*, or a locally Minkowski space, according as $K\neq 0$, or $K=0$. Furthermore the Wagner space F_n is conformal to this F_n^*, so we have the following

THEOREM. *If an $n(\geqq 3)$-dimensional Finsler space is a Wagner space with respect to a gradient vector field, and of W-scalar curvature K, then the space is conformal to a Riemannian space of constant curvature, or conformal to a locally Minkowski space, according as $K\neq 0$ or $K=0$.*

It should be noticed that the above scalar $K\neq 0$ does not depend on the supporting element y^i, but it is not necessarily constant.

REFERENCES

[1] HASHIGUCHI, M., On Wagner's generalized Berwald spaces, *J. Korean Math. Soc.* **12** (1975), 51—61.
[2] HASHIGUCHI, M. and ICHIJYŌ, Y., On conformal transformations of Wagner spaces, *Rep. Fac. Sci. Kagoshima Univ. (Math. Phys. Chem.)* **10** (1977), 19—25.
[3] MATSUMOTO, M., *Foundation of Finsler geometry and special Finsler spaces* (to appear).
[4] NUMATA, S., On Landsberg spaces of scalar curvature, *J. Korean Math. Soc.* **12** (1975), 97—100.
[5] VARGA, T., Über Berwaldsche Räume I, *Publ. Math. Debrecen* **25** (1978), 213—223.
[6] WAGNER, V., On generalized Berwald spaces, *C. R. (Doklady) Acad. Sci. URSS (N. S.)* **39** (1943), 3—5.

(Received January 19, 1980)

DEPARTMENT OF MATHEMATICS
KAGOSHIMA UNIVERSITY
KAGOSHIMA
JAPAN

KOSSUTH LAJOS TUDOMÁNYEGYETEM
MATEMATIKAI INTÉZETE
POSTAFIÓK 12
H—4010 DEBRECEN
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 15—26

REGULAR EXTENSIONS OF QUASI-UNIFORMITIES

by

ÁKOS CSÁSZÁR

0. Introduction

In a recent paper [1], I investigated the following problem: Let $(X, \mathcal{U})$ be a quasi-uniform space, $\mathcal{T}=\mathcal{T}(\mathcal{U})$ the topology induced by $\mathcal{U}$, $(Y, \mathcal{T}')$ an extension of the topological space $(X, \mathcal{T})$; does it exist a quasi-uniformity $\mathcal{U}'$ *compatible* with $(\mathcal{U}, \mathcal{T}')$, i.e., such that $\mathcal{T}(\mathcal{U}')=\mathcal{T}'$ and $\mathcal{U}'$ induces on X the subspace quasi-uniformity $\mathcal{U}$?

For the terminology used above and in the sequel, the reader has to consult [1]. In particular, for $t\in Y$, $\mathfrak{s}(t)$ will denote the trace in X of the $\mathcal{T}'$-neighbourhood filter of t, and, for $A\subset X$,

$$s(A) = \{t\in Y\colon A\in\mathfrak{s}(t)\}. \tag{0.1}$$

A quasi-uniformity $\mathcal{U}'$ on Y, compatible with $(\mathcal{U}, \mathcal{T}')$, is said to be a *strict extension* of $\mathcal{U}$ iff there exists, for any $U'\in\mathcal{U}'$, a $U_1'\in\mathcal{U}'$ such that

$$s\big(U_1'(t)\cap X\big) \subset U'(t) \quad \text{for} \quad t\in Y. \tag{0.2}$$

A quasi-uniformity $\mathcal{U}$ on X is said to be *regular* iff there is, for $U\in\mathcal{U}$, a $U_1\in\mathcal{U}$ such that $\overline{U_1(x)}\subset U(x)$ for $x\in X$ (the closure is understood with respect to $\mathcal{T}(\mathcal{U})$).

The main purpose of this paper is to give a solution for Problem 9.11 in [1]: Look for a necessary and sufficient condition for the existence of a compatible regular quasi-uniformity. The solution will be based on the results in [1] concerning the existence of compatible strict extensions. These can be summarized as follows ([1], 7.2 and 7.3, cf. also 6.2):

If $\mathcal{U}'$ is a strict extension of $\mathcal{U}$ compatible with $(\mathcal{U}, \mathcal{T}')$, then consider, for $U'\in\mathcal{U}'$, the set

$$U = U'\cap(X\times X) \tag{0.3}$$

and the map $S\colon Y-X\to 2^X$ defined by

$$S(p) = U'(p)\cap X. \tag{0.4}$$

AMS (MOS) subject classifications (1979). Primary 54E15.

Key words and phrases. Quasi-uniformity, extension, strict extension, regular quasi-uniformity.

Then the set Φ of the pairs (U, S) obtained in this way satisfies

(0.5) $U \in \mathcal{U}$, $S(p) \in \mathfrak{s}(p)$ for $p \in Y-X$, $(U, S) \in \Phi$,

(0.6) for $U_0 \in \mathcal{U}$ there is $(U, S) \in \Phi$ such that $U \subset U_0$,

(0.7) for $p \in Y-X$, $S_0 \in \mathfrak{s}(p)$ there is $(U, S) \in \Phi$ such that $S(p) \subset S_0$,

(0.8) for a given $(U, S) \in \Phi$ there exists $(U^*, S^*) \in \Phi$ such that

(a) $U^{*2} \subset U$,

(b) $U^*(S^*(p)) \subset S(p)$ for $p \in Y-X$,

(c) $U^*(x) \in \mathfrak{s}(p)$ implies $S^*(p) \subset U(x)$ for $x \in X$, $p \in Y-X$,

(d) $S^*(p) \in \mathfrak{s}(q)$ implies $S^*(q) \subset S(p)$ for $p, q \in Y-X$.

Moreover, the topology $\mathcal{T}'$ is a strict extension of $\mathcal{T}$.

In connection with the above results, let us denote by Σ the set of all maps S: $Y-X \to 2^X$. Let us further call a family $\Phi \subset \mathcal{U} \times \Sigma$ satisfying (0.5)—(0.8) an *extensor* for the pair $(\mathcal{U}, \mathcal{T}')$. Hence we can say that if $\mathcal{U}'$ is a strict extension of $\mathcal{U}$ compatible with $(\mathcal{U}, \mathcal{T}')$, then (0.3) and (0.4) furnish, if U' runs over $\mathcal{U}'$, an extensor for $(\mathcal{U}, \mathcal{T}')$.

Conversely, if $\mathcal{T}'$ is a strict extension of $\mathcal{T}$ and Φ is an extensor for $(\mathcal{U}, \mathcal{T}')$, then define, for $(U, S) \in \Phi$, $W = W(U, S) \subset Y \times Y$ by

$$W(x) = U(x) \cup s(U(x)) \quad \text{for} \quad x \in X, \tag{0.9}$$

$$W(p) = S(p) \cup s(S(p)) \quad \text{for} \quad p \in Y-X. \tag{0.10}$$

The set of these $W(U, S)$ constitutes a quasi-uniform subbase on Y generating a quasi-uniformity $\mathcal{U}'(\Phi)$ that is a strict extension of $\mathcal{U}$ compatible with $(\mathcal{U}, \mathcal{T}')$.

Finally, if $\mathcal{U}'$ is a strict extension of $\mathcal{U}$ compatible with $(\mathcal{U}, \mathcal{T}')$ and Φ is defined with the help of (0.3) and (0.4) then $\mathcal{U}'(\Phi) = \mathcal{U}'$.

1. Necessary and sufficient conditions for the existence of regular compatible quasi-uniformities

Now we can formulate first a necessary condition:

THEOREM 1.1. *Let $\mathcal{U}'$ be a regular quasi-uniformity compatible with $(\mathcal{U}, \mathcal{T}')$. Define, for $U' \in \mathcal{U}'$, the set $U \in \mathcal{U}$ and the map $S \in \Sigma$ according to* (0.3) *and* (0.4), *and denote by Φ the set of all pairs (U, S). Then Φ satisfies* (0.5), (0.6), (0.7), *and*

(1.1.1) *for a given $(U, S) \in \Phi$, there is $(U^*, S^*) \in \Phi$ satisfying*

(a) $U^*(\overline{U^*(x)}) \subset U(x)$ *for* $x \in X$,

(b) $U^*(\overline{S^*(p)}) \subset S(p)$ *for* $p \in Y-X$,

(c) *if $U^*(x)$ intersects each member of $\mathfrak{s}(p)$ then $S^*(p) \subset U(x)$ for $x \in X$, $p \in Y-X$,*

(d) *if $S^*(p)$ intersects each member of $\mathfrak{s}(q)$ then $S^*(q) \subset S(p)$ for $p, q \in Y-X$.*

PROOF. By [1], 8.7, $\mathcal{U}'$ is a strict extension of $\mathcal{U}$ so that Φ satisfies (0.5)—(0.7). In order to verify (1.1.1), let $(U, S)\in\Phi$ be obtained from $U'\in\mathcal{U}'$ with the help of (0.3) and (0.4). Choose $U_1'\in\mathcal{U}'$ such that $U_1'^2\subset U'$ and $U_2'\in\mathcal{U}'$ such that $\operatorname{cl} U_2'(t)\subset \subset U_1'(t)$ for $t\in Y$ (where $\operatorname{cl} A$ denotes the $\mathcal{T}'$-closure of A while $\bar{A}$ is the $\mathcal{T}$-closure for $A\subset X$). Put

$$U^* = U_2'\cap(X\times X), \quad S^*(p) = U_2'(p)\cap X.$$

Then $(U^*, S^*)\in\Phi$,

$$U^*\big(\overline{U^*(x)}\big) \subset U_1'\big(\operatorname{cl} U_2'(x)\big)\cap X \subset U_1'^2(x)\cap X \subset U'(x)\cap X = U(x) \quad \text{for} \quad x\in X;$$

$$U^*\big(\overline{S^*(p)}\big) \subset U_1'\big(\operatorname{cl} U_2'(p)\big)\cap X \subset U_1'^2(p)\cap X \subset U'(p)\cap X = S(p) \quad \text{for} \quad p\in Y-X.$$

If $U^*(x)$ intersects each member of $\mathfrak{s}(p)$ then $p\in\operatorname{cl} U^*(x)\subset\operatorname{cl} U_2'(x)\subset U_1'(x)$, hence

$$S^*(p) \subset U_1'(p)\cap X \subset U_1'^2(x)\cap X \subset U'(x)\cap X = U(x)$$

for $x\in X, p\in Y-X$.

If $S^*(p)$ intersects each member of $\mathfrak{s}(q)$ then $q\in\operatorname{cl} S^*(p)\subset\operatorname{cl} U_2'(p)\subset U_1'(p)$, hence

$$S^*(q) \subset U_1'(q)\cap X \subset U_1'^2(p)\cap X \subset U'(p)\cap X = S(p)$$

for $p, q\in Y-X$. ∎

Let us call a family $\Phi\subset\mathcal{U}\times\Sigma$ satisfying (0.5)—(0.7) and (1.1.1) a *regular extensor* for $(\mathcal{U}, \mathcal{T}')$. The terminology is motivated by 1.1 and the following

THEOREM 1.2. *Let $\mathcal{T}'$ be a strict extension of $\mathcal{T}$ and Φ be a regular extensor for $(\mathcal{U}, \mathcal{T}')$. Define $W=W(U, S)$ by* (0.9) *and* (0.10) *for $(U, S)\in\Phi$. Then*

$$\{W(U, S)\colon (U, S)\in\Phi\}$$

is a quasi-uniform subbase and generates a regular quasi-uniformity $\mathcal{U}'(\Phi)$ compatible with $(\mathcal{U}, \mathcal{T}')$.

PROOF. Each of the conditions (1.1.1) (a)—(d) is obviously stronger than the respective condition in (0.8). Hence $\mathcal{U}'(\Phi)$ is a strict extension of $\mathcal{U}$ compatible with $(\mathcal{U}, \mathcal{T}')$. We show that it is regular. For this purpose, it suffices to verify that, if $(U^*, S^*)\in\Phi$ is chosen, for a given $(U, S)\in\Phi$, according to (1.1.1), then $W^*= =W(U^*, S^*)$ satisfies

(1.2.1) $$\operatorname{cl} W^*(t) \subset W(t) \quad \text{for} \quad t\in Y.$$

In fact, then

$$\operatorname{cl}\Big(\bigcap_i^n W_i^*\Big)(t) = \operatorname{cl}\Big(\bigcap_1^n W_i^*(t)\Big) \subset \bigcap_1^n \operatorname{cl} W_i^*(t) \subset \bigcap_1^n W_i(t) = \Big(\bigcap_1^n W_i\Big)(t)$$

whenever $t\in Y$, $W_i=W(U_i, S_i)$, $W_i^*=W(U_i^*, S_i^*)$, and $(U_i, S_i)\in\Phi$ satisfies (1.1.1) together with $(U_i^*, S_i^*)\in\Phi$.

Now, in order to prove (1.2.1), we need the following simple

LEMMA 1.2.2. *If $\mathcal{T}'$ is a strict extension of $\mathcal{T}$, then*

(1.2.3) $$\operatorname{cl} \mathfrak{s}(A) \subset \operatorname{cl} A \quad \textit{for} \quad A \subset X.$$

PROOF of the lemma. The sets $s(G)$, where G is $\mathcal{T}$-open, constitute a base for $\mathcal{T}'$. Hence if $t\in \operatorname{cl} s(A)$, then $t\in s(G)$ implies $s(G)\cap s(A)\neq\emptyset$ for every $\mathcal{T}$-open G, whence $G\cap A\neq\emptyset$ and, *a fortiori*, $s(G)\cap A\neq\emptyset$ for all these G, i.e., $t\in \operatorname{cl} A$. ∎

By hypothesis, $\mathcal{T}'$ is a strict extension of $\mathcal{T}$. Hence, for $x\in X$, (0.9) implies by 1.2.2

$$\operatorname{cl} W^*(x) = \operatorname{cl} U^*(x)\cup \operatorname{cl} s\big(U^*(x)\big) \subset \operatorname{cl} U^*(x),$$

and if $y\in X$ then $y\in \operatorname{cl} W^*(x)$ implies

$$y\in \operatorname{cl} U^*(x)\cap X = \overline{U^*(x)} \subset U(x) \subset W(x)$$

by (1.1.1) (a) and (0.9); on the other hand, $p\in Y-X$, $p\in \operatorname{cl} W^*(x)$ implies that $U^*(x)$ intersects every member of $\mathfrak{s}(p)$, hence $S^*(p)\subset U(x)$ and $U(x)\in\mathfrak{s}(p)$,

$$p\in s\big(U(x)\big) \subset W(x)$$

by (1.1.1) (c) and (0.9).

Now let $p\in Y-X$. Then by 1.2.2 and (0.10)

$$\operatorname{cl} W^*(p) = \operatorname{cl} S^*(p)\cup \operatorname{cl} s\big(S^*(p)\big) \subset \operatorname{cl} S^*(p).$$

Hence $x\in X$, $x\in \operatorname{cl} W^*(p)$ implies

$$x\in \operatorname{cl} S^*(p)\cap X = \overline{S^*(p)} \subset S(p) \subset W(p)$$

by (1.1.1) (b) and (0.10), and $q\in Y-X$, $q\in \operatorname{cl} W^*(p)$ implies that $S^*(p)$ intersects each member of $\mathfrak{s}(q)$ so that, by (1.1.1) (d),

$$S^*(q) \subset S(p), \quad S(p)\in\mathfrak{s}(q), \quad q\in s\big(S(p)\big) \subset W(p)$$

by (0.10). Thus (1.2.1) is established. ∎

COROLLARY 1.3. *For the existence of a regular quasi-uniformity compatible with* $(\mathcal{U}, \mathcal{T}')$ *it is necessary and sufficient that* $\mathcal{T}'$ *be a strict extension of* $\mathcal{T}$ *and there exist a regular extensor for* $(\mathcal{U}, \mathcal{T}')$.

PROOF. 1.1 and 1.2. ∎

COROLLARY 1.4. *If* $\mathcal{U}'$ *is a regular quasi-uniformity compatible with* $(\mathcal{U}, \mathcal{T}')$, Φ *is defined as in* 1.1, *and* $\mathcal{U}'(\Phi)$ *is constructed as in* 1.2, *then* $\mathcal{U}'(\Phi)=\mathcal{U}'$.

PROOF. By [1], 8.7, $\mathcal{U}'$ is a strict extension of $\mathcal{U}$, hence the statement is a consequence of the corresponding result on strict extensions (cf. also [1], 8.8). ∎

2. Almost Cauchy filters

From the above results, it is easy to deduce a necessary condition concerning the trace filters $\mathfrak{s}(p)$. For this purpose, let us call *almost Cauchy* a filter base $\mathfrak{r}$ in a quasi-uniform space $(X, \mathcal{U})$ iff, for every $U\in\mathcal{U}$, there exist $U_0\in\mathcal{U}$ and $R_0\in\mathfrak{r}$ such that $R_0\subset U(x)$ whenever $x\in X$ and each member of $\mathfrak{r}$ intersects $U_0(x)$.

THEOREM 2.1. *If there is a regular quasi-uniformity* $\mathcal{U}'$ *compatible with* $(\mathcal{U}, \mathcal{T}')$ *then every trace filter* $\mathfrak{s}(p)$ *is almost Cauchy for* $p\in Y-X$.

PROOF. By 1.1 there is a regular extensor Φ for $(\mathcal{U}, \mathcal{T}')$. For a given $U \in \mathcal{U}$, select by (0.6) $(U_1, S_1) \in \Phi$ such that $U_1 \subset U$ and by (1.1.1) (c) $(U_0, S_0) \in \Phi$ such that $S_0(p) \subset U_1(x) \subset U(x)$ whenever $U_0(x)$ intersects each member of $\mathfrak{s}(p)$. Clearly, $U_0 \in \mathcal{U}$ and $S_0(p) \in \mathfrak{s}(p)$ show that $\mathfrak{s}(p)$ is almost Cauchy for an arbitrary $p \in Y - X$. ∎

We investigate the properties of almost Cauchy filter bases.

THEOREM 2.2. *Every Cauchy filter base is almost Cauchy, and every almost Cauchy filter base is weakly Cauchy.*

PROOF. Let $\mathfrak{r}$ be Cauchy, $U \in \mathcal{U}$, and $U_0 \in \mathcal{U}$, $U_0^2 \subset U$. Select $R_0 \in \mathfrak{r}$ such that $R_0 \times R_0 \subset U_0$. Then if $U_0(x)$ intersects each member of $\mathfrak{r}$, then $R_0 \cap U_0(x) \neq \emptyset$, hence $R_0 \subset U_0^2(x) \subset U(x)$.

The second statement is obvious by the respective definitions (cf. [1]). ∎

There is no such relation between almost Cauchy and stable (see [1]) filter bases. In fact, if $X = \mathbf{R}$ and $\mathcal{U}$ is the euclidean uniformity, then

$$\mathfrak{r}_1 = \{(c, +\infty) \colon c \in \mathbf{R}\}$$

is almost Cauchy without being stable, while

$$\mathfrak{r}_2 = \{(-1-\varepsilon, 1+\varepsilon) \colon \varepsilon > 0\}$$

is stable but fails to be almost Cauchy because, if

$$U = \{(x, y) \colon |x-y| < 1\}$$

and $x=0$, then each member of $\mathfrak{r}_2$ meets each neighbourhood of x but $U(x)$ does not contain any member of $\mathfrak{r}_2$. Observe that both $\mathfrak{r}_1$ and $\mathfrak{r}_2$ are round.

For almost Cauchy filter bases, we have a much stronger result than [1], 4.9:

THEOREM 2.3. *If $\mathfrak{r}$ is an almost Cauchy filter base and $\mathfrak{r}'$ is a filter base finer than $\mathfrak{r}$ then $\mathfrak{r}'$ is almost Cauchy.*

PROOF. For $U \in \mathcal{U}$, let $U_0 \in \mathcal{U}$ and $R_0 \in \mathfrak{r}$ be chosen according to the definition, and $R_0 \supset R_0' \in \mathfrak{r}'$. If $U_0(x)$ meets every member of $\mathfrak{r}'$, then the same holds for the members of $\mathfrak{r}$, and $R_0' \subset R_0 \subset U(x)$. ∎

The following lemma corresponds to [1], 4.10:

LEMMA 2.4. *Let $(X, \mathcal{U})$ be a subspace of $(Y, \mathcal{U}')$. If $\mathfrak{r}$ is a filter base in X, almost Cauchy with respect to $\mathcal{U}'$, then it is almost Cauchy with respect to $\mathcal{U}$.*

PROOF. For $U \in \mathcal{U}$, choose $U' \in \mathcal{U}'$ such that $U = U' \cap (X \times X)$, and $U_0' \in \mathcal{U}'$, $R_0 \in \mathfrak{r}$ such that $R_0 \subset U'(t)$ whenever $t \in Y$ and $U_0'(t)$ intersects every member of $\mathfrak{r}$. Define $U_0 = U_0' \cap (X \times X)$. Now if $U_0(x)$ meets each member of $\mathfrak{r}$ for an $x \in X$, then the same is true for $U_0'(x)$ so that $R_0 \subset U'(x) \cap X = U(x)$. ∎

COROLLARY 2.5. *Let $(X, \mathcal{U})$ be a subspace of $(Y, \mathcal{U}')$ and $\mathfrak{r}'$ be an almost Cauchy filter base in $(Y, \mathcal{U}')$ having a trace $\mathfrak{r}$ in X. Then $\mathfrak{r}$ is almost Cauchy with respect to $\mathcal{U}$.*

PROOF. 2.3 and 2.4. ∎

We know from [1], 4.6 that every neighbourhood filter is stable. The question whether neighbourhood filters are almost Cauchy is more delicate.

THEOREM 2.6. *In a regular quasi-uniform space $(X, \mathcal{U})$, every neighbourhood filter is almost Cauchy.*

PROOF. Let $\mathfrak{r}$ be the neighbourhood filter of $y \in X$ and $U \in \mathcal{U}$ be given. Choose $U_1 \in \mathcal{U}$ such that $U_1^2 \subset U$ and $U_0 \in \mathcal{U}$ such that $\overline{U_0(x)} \subset U_1(x)$ for $x \in X$. Define $R_0 = U_1(y) \in \mathfrak{r}$.

If $U_0(x)$ intersects each member of $\mathfrak{r}$, then $y \in \overline{U_0(x)} \subset U_1(x)$, hence $R_0 \subset \subset U_1^2(x) \subset U(x)$. ∎

Observe that 2.1 can be deduced from 2.6 and 2.5.

In general, neighbourhood filters need not be almost Cauchy. This is shown by

EXAMPLE 2.7. Let $(X, \mathcal{U})$ be the quasi-uniform space defined in [1], 8.2, and suppose that there is a $\mathcal{T} = \mathcal{T}(\mathcal{U})$-open set G that is not closed. Let $y \in \bar{G} - G$, $U = U_G$, and assume that $U_0 = \bigcap_1^n U_{G_i}$ and a neighbourhood R_0 of y correspond to U according to the almost Cauchy property of the neighbourhood filter of y. Consider all sets of the form

$$G \cap \bigcap_1^n H_i \tag{2.7.1}$$

where $H_i = G_i$ or $H_i = X - G_i$ for $i = 1, \dots, n$. Clearly, y belongs to the closure of at least one of the sets (2.7.1), say $y \in \bar{A}$, and select $x \in A$. Now $U_0(x)$ is the intersection of those G_i for which $x \in G_i$, hence $A \subset U_0(x)$, and every member of $\mathfrak{r}$ intersects $U_0(x)$. On the other hand, R_0 is not contained in $U(x) = G$. Thus $\mathfrak{r}$ cannot be almost Cauchy. ∎

Since a filter base of the form $\{\{y\}\}$ is Cauchy, we see from 2.7 that a statement similar to [1], (4.6) does not hold for Cauchy or almost Cauchy filter bases.

3. Finite, regular extensions

Let us now consider the case when $Y - X$ is finite or, more generally, of the form occurring in [1], 5.1 and 7.6. In this case we can replace the conditions in 1.3 by much simpler ones.

THEOREM 3.1. *Let* $Y - X = \bigcup_1^n Z_i$, $\mathfrak{s}(p) = \mathfrak{s}_i$ *for* $p \in Z_i$. *There exists a regular quasi-uniformity compatible with* $(\mathcal{U}, \mathcal{T}')$ *iff* $\mathcal{U}$ *is a regular quasi-uniformity,* $\mathcal{T}'$ *is a regular topology, and every* $\mathfrak{s}_i$ *is round and almost Cauchy.*

PROOF. The necessity follows from [1], 8.5, 8.1, 1.1, and Theorem 2.1 of the present paper.

Suppose now that the conditions are fulfilled. We can assume $\mathfrak{s}_i \neq \mathfrak{s}_j$ for $i \neq j$. We construct a regular extensor for $(\mathcal{U}, \mathcal{T}')$.

Let Φ consist of the pairs (U, S) such that $U \in \mathcal{U}$, $S_i \in \mathfrak{s}_i$ are arbitrarily chosen, and $S(p) = S_i$ for $p \in Z_i$. (0.5)—(0.7) are obviously true.

In order to verify (1.1.1), let $U \in \mathcal{U}$ and $S_i \in \mathfrak{s}_i$ be given ($i = 1, \dots, n$). Select $U_1 \in \mathcal{U}$ such that $U_1^2 \subset U$, then by the regularity of $\mathcal{U}$ a $U_2 \in \mathcal{U}$ such that $\overline{U_2(x)} \subset$

$\subset U_1(x)$ for $x\in X$. Since $\mathfrak{s}_i$ is round, there are $U_{3i}\in\mathcal{U}$ and $S_{3i}\in\mathfrak{s}_i$ such that $U_{3i}(S_{3i})\subset S_i$, then by the regularity of $\mathcal{T}'$ choose $S_{4i}\in\mathfrak{s}_i$ such that $\bar{S}_{4i}\subset S_{3i}$ (S_{4i} is the trace of a $\mathcal{T}'$-neighbourhood of a $p\in Z_i$ having a $\mathcal{T}'$-closure contained in a $\mathcal{T}'$-neighbourhood whose trace is S_{3i}). Again by the regularity of $\mathcal{T}'$, $\mathfrak{s}_i\neq\mathfrak{s}_j$ implies, for $i\neq j$, the existence of $\hat{S}_{ij}\in\mathfrak{s}_i$ such that $\hat{S}_{ij}$ does not meet all members of $\mathfrak{s}_j$. Finally, the almost Cauchy property of $\mathfrak{s}_i$ implies the existence of $U_{5i}\in\mathcal{U}$ and $S_{5i}\in\mathfrak{s}_i$ such that $S_{5i}\subset U(x)$ whenever $x\in X$ and $U_{5i}(x)$ intersects every member of $\mathfrak{s}_i$.

Define

$$U^* = U_1\cap U_2\cap \bigcap_1^n U_{3i}\cap \bigcap_1^n U_{5i}\in\mathcal{U},$$

$$S_i^* = S_{3i}\cap S_{4i}\cap \bigcap_{j\neq i} \hat{S}_{ij}\cap S_{5i}\in\mathfrak{s}_i,$$

and $S^*(p)=S_i^*$ for $p\in Z_i$.

Then, for $x\in X$, $p\in Z_i$, $q\in Z_j$,

$$U^*\big(\overline{U^*(x)}\big)\subset U_1\big(\overline{U_2(x)}\big)\subset U_1^2(x)\subset U(x),$$

$$U^*\big(\overline{S^*(p)}\big)\subset U_{3i}(\bar{S}_{4i})\subset U_{3i}(S_{3i})\subset S_i = S(p),$$

further if $U^*(x)$ meets each member of $\mathfrak{s}_i$ then the same holds for $U_{5i}(x)$, hence $S^*(p)\subset S_{5i}\subset U(x)$. Finally, $i=j$ clearly implies $S^*(q)=S_i^*\subset S_{3i}\subset S_i=S(p)$; on the other hand, if $i\neq j$, then $S_i^*\subset\hat{S}_{ij}$ implies that $S^*(p)$ does not intersect all members of $\mathfrak{s}_j=\mathfrak{s}(q)$.

Since $\mathcal{T}'$ is regular, it is a strict extension of $\mathcal{T}$, hence the statement follows from 1.2. ∎

The following example shows that the almost Cauchy property of the trace filters cannot be omitted from 3.1:

EXAMPLE 3.2. Let $X=\mathbf{R}$ and Π be the collection of all finite partitions of X (i.e., $\mathfrak{P}\in\Pi$ iff $\mathfrak{P}=\{A_1, \dots, A_n\}$, $X=\bigcup_1^n A_i$, $A_i\cap A_j=\emptyset$ for $i\neq j$). For $\mathfrak{P}\in\Pi$, $\mathfrak{P}=\{A_1, \dots, A_n\}$, set

$$U(\mathfrak{P}) = \bigcup_1^n (A_i\times A_i).$$

Then

$$\{U(\mathfrak{P}):\ \mathfrak{P}\in\Pi\}$$

is a uniform base that generates a uniformity $\mathcal{U}$ on X. Clearly, $\mathcal{T}(\mathcal{U})=\mathcal{T}$ is the discrete topology of X.

Choose $p\notin X$, let $Y=X\cup\{p\}$ and $\mathfrak{s}(p)$ be the filter in X generated by the filter base

$$\mathfrak{r} = \{(c, +\infty):\ c\in\mathbf{R}\}.$$

Denote by $\mathcal{T}'$ the strict extension of $\mathcal{T}$ on Y corresponding to the trace filter $\mathfrak{s}(p)$.

$\mathcal{U}$ is regular by [1], 8.3. The topology $\mathcal{T}'$ is regular because $\{x\}$ is a $\mathcal{T}'$-closed neighbourhood of $x\in X$ and $(c, +\infty)\cup\{p\}$ is a similar neighbourhood of p. Observe that $\mathfrak{s}(p)$ is round and weakly Cauchy so that, by [1], 7.6, there exists a strict extension $\mathcal{U}'$ of $\mathcal{U}$ compatible with $(\mathcal{U}, \mathcal{T}')$. In fact, if $\mathfrak{P}_1$ denotes the partition consisting of $(c, +\infty)$ and $(-\infty, c]$, then $U(\mathfrak{P}_1)((c, +\infty)) \subset (c, +\infty)$, and if $\mathfrak{P}_2 = \{\mathbf{Q}, \mathbf{R}-\mathbf{Q}\}$, then $U(\mathfrak{P}_2)(x)$ never belongs to $\mathfrak{s}(p)$.

On the other hand, $\mathfrak{s}(p)$ is not almost Cauchy. Indeed, for $U = U(\mathfrak{P}_2)$, there is no $U_1 = U(\mathfrak{P}_3)$ with the property occurring in the definition, because, if $\mathfrak{P}_3 = \{A_1, \ldots, A_n\}$, then at least one of the sets A_i, say A_1, is not bounded from above so that, for $x \in A_1$, $U(\mathfrak{P}_3)(x) = A_1$ intersects each member of $\mathfrak{s}(p)$ but $U(\mathfrak{P}_2)(x)$ does not belong to $\mathfrak{s}(p)$. ∎

4. Questions of unicity

We shall show that, under the conditions of 3.1, the regular compatible quasi-uniformity $\mathcal{U}'$ is unique. In fact, a similar statement is true, more generally, for strict compatible extensions in the case examined in [1], 7.6.

In order to see this, let us first consider an arbitrary quasi-uniform space $(X, \mathcal{U})$ and an extension $(Y, \mathcal{T}')$ of $(X, \mathcal{T})$ where, as usually, $\mathcal{T} = \mathcal{T}(\mathcal{U})$. Let Φ_1 and Φ_2 be two subsets of $\mathcal{U} \times \Sigma$. We shall say that Φ_1 is *coarser* than Φ_2 or Φ_2 is *finer* than Φ_1 iff, for $(U_1, S_1) \in \Phi_1$, there is $(U_2, S_2) \in \Phi_2$ satisfying

$$(4.1) \qquad U_2 \subset U_1 \quad \text{and} \quad S_2(p) \subset S_1(p) \quad \text{for} \quad p \in Y-X.$$

Φ_1 and Φ_2 are said to be *equivalent* iff each of them is finer than the other.

LEMMA 4.1. *If $\mathcal{T}'$ is a strict extension of $\mathcal{T}$, Φ_1 and Φ_2 are extensors for $(\mathcal{U}, \mathcal{T}')$, and Φ_1 is coarser than Φ_2, then $\mathcal{U}'(\Phi_1) \subset \mathcal{U}'(\Phi_2)$.*

PROOF. By (0.9) and (0.10), (4.1) implies

$$W(U_2, S_2) \subset W(U_1, S_1). \quad ∎$$

LEMMA 4.2. *If Φ_2 is an extensor for $(\mathcal{U}, \mathcal{T}')$ and Φ_2 is equivalent to Φ_1, then Φ_2 is an extensor as well.*

PROOF. For $(U_2, S_2) \in \Phi_2$, choose $(U_1, S_1) \in \Phi_1$ such that

$$(4.2.1) \qquad U_1 \subset U_2, \quad S_1(p) \subset S_2(p) \quad \text{for} \quad p \in Y-X.$$

Then $S_1(p) \in \mathfrak{s}(p)$ implies $S_2(p) \in \mathfrak{s}(p)$ so that Φ_2 satisfies (0.5).

Now, for a given $U_0 \in \mathcal{U}$, choose $(U_1, S_1) \in \Phi_1$ such that $U_1 \subset U_0$ and $(U_2, S_2) \in \Phi_2$ such that

$$(4.2.2) \qquad U_2 \subset U_1, \quad S_2(p) \subset S_1(p) \quad \text{for} \quad p \in Y-X.$$

Then $U_2 \subset U_1 \subset U_0$ and Φ_2 fulfils (0.6).

For $p \in Y-X$, $S_0 \in \mathfrak{s}(p)$, choose $(U_1, S_1) \in \Phi_1$ such that $S_1(p) \subset S_0$ and $(U_2, S_2) \in \Phi_2$ satisfying (4.2.2). Then $S_2(p) \subset S_1(p) \subset S_0$ and Φ_2 fulfils (0.7).

For $(U_2, S_2) \in \Phi_2$, select $(U_1, S_1) \in \Phi_1$ satisfying (4.2.1), then $(U_1^*, S_1^*) \in \Phi_1$ such that (0.8) be fulfilled by (U_1, S_1) and (U_1^*, S_1^*) instead of (U, S) and (U^*, S^*), finally $(U_2^*, S_2^*) \in \Phi_2$ such that

$$(4.2.3) \qquad U_2^* \subset U_1^*, \quad S_2^*(p) \subset S_1^*(p) \quad \text{for} \quad p \in Y-X.$$

Then

$$U_2^{*2} \subset U_1^* \subset U_1 \subset U_2,$$

$$U_2^*(S_2^*(p)) \subset U_1^*(S_1^*(p)) \subset S_1(p) \subset S_2(p),$$

$$U_2^*(x) \in \mathfrak{s}(p) \quad \text{implies} \quad U_1^*(x) \in \mathfrak{s}(p), \quad \text{hence}$$

$$S_2^*(p) \subset S_1^*(p) \subset U_1(x) \subset U_2(x),$$

$$S_2^*(p) \in \mathfrak{s}(q) \quad \text{implies} \quad S_1^*(p) \in \mathfrak{s}(q), \quad \text{hence}$$

$$S_2^*(q) \subset S_1^*(q) \subset S_1(p) \subset S_2(p)$$

for $x \in X$, $p, q \in Y-X$. Therefore Φ_2 satisfies (0.8). ∎

COROLLARY 4.3. *If $\mathcal{T}'$ is a strict extension of $\mathcal{T}$, Φ_1 is an extensor for $(\mathcal{U}, \mathcal{T}')$, and Φ_2 is equivalent to Φ_1, then $\mathcal{U}'(\Phi_1) = \mathcal{U}'(\Phi_2)$.*

PROOF 4.2 and 4.1. ∎

LEMMA 4.4. *Let $\mathcal{T}'$ be a strict extension of $\mathcal{T}$ and Φ be an extensor for $(\mathcal{U}, \mathcal{T}')$. Then there exists an extensor $\hat{\Phi}$, equivalent to Φ, and such that*

(4.4.1) $\mathfrak{s}(p) = \mathfrak{s}(q)$ *implies* $\hat{S}(p) = \hat{S}(q)$ *for* $p, q \in Y-X$, $(\hat{U}, \hat{S}) \in \hat{\Phi}$.

PROOF. Clearly, $Y-X = \bigcup_{i \in I} Z_i$ where $\mathfrak{s}(p) = \mathfrak{s}_i$ for $p \in Z_i$ and $\mathfrak{s}_i \neq \mathfrak{s}_j$ for $i \neq j$. Select a point $p_i \in Z_i$ for every i.

For every $(U, S) \in \Phi$, we define a pair $(\hat{U}, \hat{S}) \in \mathcal{U} \times \Sigma$ in the following way. We choose a pair $(U^*, S^*) \in \Phi$ satisfying (0.8) and then we set

$$\hat{U} = U^*, \quad \hat{S}(p) = S^*(p_i) \quad \text{for} \quad p \in Z_i. \tag{4.4.2}$$

We show that

$$\hat{\Phi} = \{(\hat{U}, \hat{S}) \colon (U, S) \in \Phi\} \tag{4.4.3}$$

is the family we are looking for.

In fact, $\hat{\Phi}$ clearly fulfils (4.4.1) so that, by 4.2, it suffices to show that $\hat{\Phi}$ is equivalent to Φ. Now, by (0.8) (a), applied for Φ, we have $\hat{U} = U^* \subset U$, and, by (0.8) (d), $\hat{S}(p) = S^*(p_i) \subset S(p)$ if $p \in Z_i$ since then $S^*(p) \in \mathfrak{s}(p) = \mathfrak{s}(p_i)$. Hence Φ is coarser than $\hat{\Phi}$. On the other hand, for a given $(U, S) \in \Phi$, consider $(U^*, S^*) \in \Phi$ and $(\hat{U}, \hat{S}) \in \hat{\Phi}$ as above, and then choose $(U^{**}, S^{**}) \in \Phi$ satisfying (0.8) together with (U^*, S^*) (i.e., by replacing (U, S) and (U^*, S^*) by (U^*, S^*) and (U^{**}, S^{**}), respectively). Then $U^{**} \subset U^* = \hat{U}$ and, for $p \in Z_i$,

$$S^{**}(p) \subset S^*(p) = \hat{S}(p_i) \tag{4.4.4}$$

since $S^{**}(p_i) \in \mathfrak{s}(p_i) = \mathfrak{s}(p)$ implies (4.4.4) by (0.8) (d). ∎

THEOREM 4.5. *Let $Y-X = \bigcup_{1}^{n} Z_i$ and $\mathfrak{s}(p) = \mathfrak{s}_i$ for $p \in Z_i$. If $\mathcal{T}'$ is a strict extension of $\mathcal{T}$ and every $\mathfrak{s}_i$ is round and weakly Cauchy then there exists a unique strict extension of $\mathcal{U}$ compatible with $(\mathcal{U}, \mathcal{T}')$.*

PROOF. By [1], 7.6, there exists a strict extension $\mathcal{U}'$ of $\mathcal{U}$ compatible with $(\mathcal{U}, \mathcal{T}')$. Let $\mathcal{U}_1'$ and $\mathcal{U}_2'$ be two such extensions. By [1], 7.3, $\mathcal{U}_1'=\mathcal{U}'(\Phi_1)$, $\mathcal{U}_2'=\mathcal{U}'(\Phi_2)$ for two extensors Φ_1 and Φ_2. By 4.4 there are two extensors $\hat{\Phi}_1$ and $\hat{\Phi}_2$, equivalent to Φ_1 and Φ_2, respectively, and satisfying

$$\hat{S}_1(p) = \hat{S}_1(q), \quad \hat{S}_2(p) = \hat{S}_2(q) \tag{4.5.1}$$

for $(\hat{U}_1, \hat{S}_1)\in\hat{\Phi}_1$, $(\hat{U}_2, \hat{S}_2)\in\hat{\Phi}_2$, $p, q\in Z_i$. By 4.3

$$\mathcal{U}_1' = \mathcal{U}'(\hat{\Phi}_1), \quad \mathcal{U}_2' = \mathcal{U}'(\hat{\Phi}_2).$$

Consider $(\hat{U}_1, \hat{S}_1)\in\hat{\Phi}_1$. By (4.5.1) there are sets $S_i\in\mathfrak{s}_i$ such that $\hat{S}_1(p)=S_i$ for $p\in Z_i$. Choose points $p_i\in Z_i$ and apply (0.6) for $\hat{\Phi}_2$. It follows that there is $(\hat{U}_2, \hat{S}_2)\in\hat{\Phi}_2$ such that $\hat{U}_2\subset\hat{U}_1$. Then apply (0.7) for $\hat{\Phi}_2$ and p_i; we obtain $(\hat{U}_{2i}, \hat{S}_{2i})\in\hat{\Phi}_2$ such that

$$\hat{S}_{2i}(p_i) \subset S_i\in\mathfrak{s}_i = \mathfrak{s}(p_i).$$

By (4.5.1) $\hat{S}_{2i}(p)=\hat{S}_{2i}(p_i)$ for $p\in Z_i$, hence if we put

$$W_2 = W(\hat{U}_2, \hat{S}_2)\cap\bigcap_1^n W(\hat{U}_{2i}, \hat{S}_{2i})$$

then $W_2\in\mathcal{U}'(\hat{\Phi}_2)$ and

$$\begin{aligned} W_2(x) \subset W(\hat{U}_2, \hat{S}_2)(x) &= \hat{U}_2(x)\cup s(\hat{U}_2(x)) \subset \\ &\subset \hat{U}_1(x)\cup s(\hat{U}_1(x)) = W(\hat{U}_1, \hat{S}_1)(x) \end{aligned} \tag{4.5.2}$$

for $x\in X$ by (0.9), further $p\in Z_i$ implies

$$\begin{aligned} W_2(p) \subset W(\hat{U}_{2i}, \hat{S}_{2i})(p) &= \hat{S}_{2i}(p)\cup s(\hat{S}_{2i}(p)) = \\ &= \hat{S}_{2i}(p_i)\cup s(\hat{S}_{2i}(p_i)) \subset S_i\cup s(S_i) = \\ &= \hat{S}_1(p)\cup s(\hat{S}_1(p)) = W(\hat{U}_1, \hat{S}_1)(p) \end{aligned} \tag{4.5.3}$$

by (0.10). From (4.5.2) and (4.5.3) we obtain $W_2\subset W(\hat{U}_1, \hat{S}_1)$ and this clearly implies $\mathcal{U}'(\hat{\Phi}_1)\subset\mathcal{U}'(\hat{\Phi}_2)$. The converse inclusion can be proved in the same manner. Therefore $\mathcal{U}_1'=\mathcal{U}_2'$. ∎

COROLLARY 4.6. *Under the conditions of* 3.1, *there exists a unique regular quasi-uniformity compatible with* $(\mathcal{U}, \mathcal{T}')$.

PROOF. The existence is contained in 3.1. The uniqueness follows from 4.5 because a regular extension is strict by [1], 8.7, the regular topology $\mathcal{T}'$ is a strict extension of $\mathcal{T}$, and the filters $\mathfrak{s}_i$ are weakly Cauchy by 2.2. ∎

Observe that [1], 7.12 is an obvious consequence of 4.5.

In the general case, it can happen that there are infinitely many regular quasi-uniformities compatible with a given pair $(\mathcal{U}, \mathcal{T}')$. This is shown by

EXAMPLE 4.7. Let $X=\mathbf{R}^2$, $\mathcal{U}$ be the euclidean uniformity,

$$U_\varepsilon = \{(x, y)\colon d(x, y) < \varepsilon\} \quad (\varepsilon > 0)$$

with the euclidean distance d. Let $Y \supset X$ be chosen in a manner that there exist a bijection h from $Y-X$ onto the unit circle $\{x: d(x, a)=1\}$ where $a=(0, 0)$. For $p \in Y-X$, $\alpha \geqq 1$, $\varepsilon > 0$, let $R(p)$ denote the ray issuing from a and passing through the point $h(p)$, and let $A(p, \alpha, \varepsilon)$ be the unbounded rectangle whose boundary consists of a segment perpendicular to $R(p)$, intersecting it at a distance α from a, and of two rays having the same direction as $R(p)$ and lying at a distance ε from $R(p)$. Let $\mathfrak{s}(p)$ be the filter generated in X by the filter base

$$\{A(p, \alpha, \varepsilon): \alpha \geqq 1, \varepsilon > 0\},$$

and $\mathcal{T}'$ be the strict extension of $\mathcal{T}=\mathcal{T}(\mathcal{U})$ corresponding to the trace filter system $\{\mathfrak{s}(p): p \in Y-X\}$.

Let $f: Y-X \to [1, +\infty)$ be an arbitrary function, and consider the family Φ_f consisting of the pairs

$$(U_\varepsilon, S(f, c, \varepsilon))$$

where $\varepsilon > 0$, $c \geqq 0$, and

$$S(f, c, \varepsilon)(p) = A\big(p, f(p)+c, \varepsilon\big).$$

The family Φ_f clearly satisfies (0.5)—(0.7). In order to see that it fulfils (1.1.1), consider, for a given $(U_\varepsilon, S(f, c, \varepsilon)) \in \Phi_f$, the pair $\left(U_{\varepsilon/2}, S\left(f, c+\frac{\varepsilon}{2}, \frac{\varepsilon}{2}\right)\right)$. Then (1.1.1) (a) and (b) are obviously fulfilled, and the same holds for (c) and (d) because $U_{\varepsilon/2}(x)$ never intersects every member of $\mathfrak{s}(p)$, and similarly $S\left(f, c+\frac{\varepsilon}{2}, \frac{\varepsilon}{2}\right)(p)$ meets each member of $\mathfrak{s}(q)$ only if $q=p$, in which case $S\left(f, c+\frac{\varepsilon}{2}, \frac{\varepsilon}{2}\right) \subset S(f, c, \varepsilon)$.

Hence $\mathcal{U}'(\Phi_f)=\mathcal{U}'_f$ is a regular quasi-uniformity compatible with $(\mathcal{U}, \mathcal{T}')$ whenever f is a function from $Y-X$ to $[1, +\infty)$. Clearly, $\mathcal{U}'_f=\mathcal{U}'_g$ can hold only if $f-g$ is bounded so that there are infinitely many distinct $\mathcal{U}'_f$. ∎

5. A problem

It would be interesting to know whether the simple conditions in 3.1 are sufficient in the general case. In other words, it would be important to solve

PROBLEM 5.1. Look for $X, Y, \mathcal{U}, \mathcal{T}'$ such that $\mathcal{U}$ and $\mathcal{T}'$ are regular, every filter $\mathfrak{s}(p)$ is round and almost Cauchy for $p \in Y-X$, and yet there is no regular $\mathcal{U}'$ compatible with $(\mathcal{U}, \mathcal{T}')$.

In solving 5.1, it would be perhaps useful the following

LEMMA 5.2. *If there exists a regular $\mathcal{U}'$ compatible with $(\mathcal{U}, \mathcal{T}')$ then, for $U_0 \in \mathcal{U}$, there is $U_1 \in \mathcal{U}$ such that $U_0(x) \in \mathfrak{s}(p)$ whenever $x \in X, p \in Y-X$, and $U_1(x)$ intersects every member of $\mathfrak{s}(p)$.*

PROOF. By 1.1 there exists a regular extensor Φ for $(\mathcal{U}, \mathcal{T}')$. For $U_0 \in \mathcal{U}$, choose $(U, S) \in \Phi$ such that $U \subset U_0$ and $(U^*, S^*) \in \Phi$ satisfying (1.1.1). Then $U_1=U^*$ will do according to (1.1.1) (c). ∎

Now one can imagine that, if $Y-X$ is infinite, the condition in 5.2 is not fulfilled in spite of the fact that, for a given $p\in Y-X$, there is a $U_1\in\mathcal{U}$ (depending on p) such that $U(x)\in\mathfrak{s}(p)$ whenever $U_1(x)$ intersects each member of $\mathfrak{s}(p)$ (moreover $U(x)$ contains a member of $\mathfrak{s}(p)$ independent of x). However, I did not succeed in finding an example of this kind.

REFERENCE

[1] CSÁSZÁR, Á., Extensions of quasi-uniformities, *Acta Math. Acad. Sci. Hungar.* **37** (1981), 121—145.

(Received April 2, 1980)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
ANALÍZIS I. TANSZÉK
MÚZEUM KRT. 6—8
H—1088 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 27—36

ON TURÁN—RAMSEY TYPE THEOREMS II

by

P. ERDŐS and VERA T. SÓS

This paper is a continuation of our papers [5], [10]. We investigated the following problem:

Let the edges of K_n be coloured by r colours, G_i, $1 \leqq i \leqq r$ be the graph formed by the i'th colour. Let $f(n;\, k_1, \ldots, k_r)$ be the largest integer for which there is an r-colouring of K_n such that

$$K_{k_i} \not\subset G_i, \quad 1 \leqq i \leqq r$$

and

$$\sum_{i=1}^{r-1} e(G_i) = f(n;\, k_1, \ldots, k_r). \tag{1}$$

(Here $e(G)$ denotes the number of edges of G.)

Due to Ramsey's theorem for fixed $k_1, \ldots, k_r$, $n > N(k_1, \ldots, k_r)$ such a graph does not exist. Therefore the problem makes sense only in the case when at least one of the $k_i \to \infty$ with $n \to \infty$.

It is trivial that $f(n;\, 3, l) \leqq \frac{1}{2} nl$. We proved in [2] that if $l = o(n)$ then

$$f(n;\, 2k+1, l) = \frac{1}{2}\left(1 - \frac{1}{k}\right) n^2 + o(n^2). \tag{2}$$

Bollobás—Erdős [1] and Szemerédi [11] proved that $f(n;\, 4, l) = \frac{n^2}{8} + o(n^2)$ for $l = o(n)$. No asymptotic formula is known for $f(n;\, 2k, l)$ when $l = o(n)$ and $k > 2$.

Here we start to investigate $f(n;\, k_1, \ldots, k_r)$ for $r = 3$.

Notation. $G_n(V; E)$ is a graph with $|V| = n$, $e(G_n) = |E|$, $K(k_1, \ldots, k_r)$ is a complete r-partite graph with k_i vertices in the i'th class, K_n is the complete graph on n vertices.

Let V be the vertex set of the complete graph K_n. If we consider an r-colouring of the edges of K_n, let E_i be the set of edges of K_n having the ith colour for $1 \leqq i \leqq r$. Put $G_i = G(V; E_i)$ and

$$V_i(x) = \{y\colon (x, y) \in E_i\}, \quad d_i = |V_i(x)|,$$

$$V_i(x; U) = \{y\colon (x, y) \in E_i,\ y \in V - U\},$$

$$d_i(x; U) = |V_i(x; U)|.$$

For the case $r = 3$ we prove the following theorems:

THEOREM 1.

$$f(n;\,3,\,3,\,\varepsilon n)<\frac{n^2}{4}+c_2\varepsilon n^2 \tag{3}$$

and for $n>n_0(\varepsilon)$

$$\frac{n^2}{4}+c_1\varepsilon n^2<f(n;\,3,\,3,\,\varepsilon n),$$

where $c_1>0$, $c_2>0$ *are absolute constants.*

THEOREM 2. *Let* $G_i(V;E_i)$, $1\leqq i\leqq 3$ *be graphs belonging to a 3-colouring of* K_n *with the property*

$$K_3\not\subset G_i \qquad i=1,2, \tag{4}$$

$$K_{\varepsilon n}\not\subset G_3 \tag{5}$$

and

$$|E_1|\geqq|E_2|>cn^2. \tag{6}$$

Then

$$|E_1\cup E_2|<n^2\left(\frac{1}{4}-\sqrt{c}+2c\right)+\eta n^2 \tag{7}$$

where $\eta\to 0$ *with* $\varepsilon\to 0$.

REMARK. We obtain the lower bound in Theorem 1 by a colouring in which G_1 is the complete bipartite graph $K\left(\left[\frac{n}{2}\right],\left[\frac{n+1}{2}\right]\right)$ and G_2 formed by two copies of a trianglefree graph with maximum independent set of size $o(n)$ and $|E_2|=o(n^2)$. Theorem 2 shows that this extremum is sharp; by the condition (6) we have the stronger inequality (7) instead of (3).

PROOF of Theorem 1.

(a) *The upper bound.*

We shall use the simple observation that

$$K_3\not\subset G_i \qquad i=1,2$$

$$K_{\varepsilon n}\not\subset G_3$$

implies

$$|V_1(x)\cap V_2(y)|<\varepsilon n \tag{8}$$

for any $x\neq y$, $x,y\in V$.

Assume $|E_1|\geqq|E_2|$. Let x_0 be a vertex for which $d_1(x)$ is maximal. Let

$$d_1(y_0)=\max_{y\in V_1(x_0)} d_1(y), \qquad y_0\in V_1(x_0).$$

Since $K_3\not\subset G_1$

$$V_1(x)\cap V_1(y)=\emptyset.$$

Let $U=V-(V_1(x_0)\cup V_1(y_0))$. Put

$$E_2^*=\{(x,y)\colon (x,y)\in E_2,\ x\notin U \text{ or } y\notin U\}.$$

First we prove

(9) $$|E_2^*| < \sqrt{2}\varepsilon n^2.$$

By (8), obviously, any point $z \in V$ can be joined in G_2 to at most $2\varepsilon n$ points of $V_1(x_0) \cup V_1(y_0)$. This gives (9). Thus we only have to consider the set of edges

$$E_2^{**} = \{(x, y)\colon\ (x, y) \in E_2,\ x \in U,\ y \in U\}.$$

Put

$$|U| = \delta n$$

and

$$\delta^* n = \max_{x \in U} d_2(x; V-U) = d_2(x^*; V-U) \qquad (x^* \in U).$$

As before, by (8) we get that the number of edges in G_1 incident to a vertex in $V_2(x^*)$ is at most εn^2. Since $K_3 \not\subset G_1$, the number of the remaining edges of G_1 is less than $\frac{n^2}{4}(1-\delta^*)^2$. By all of these we obtain

(11) $$|E_1 \cup E_2| < \frac{n^2}{4}(1-\delta^*)^2 + \delta\delta^* \frac{n^2}{2} + 3\varepsilon n^2.$$

If $\delta < \frac{2}{3}$ $\left(\text{and consequently } \delta^* < \frac{2}{3}\right)$ then (11) gives

$$|E_1 \cup E_2| < \frac{n^2}{4} + 3\varepsilon n^2.$$

So all we have to show is $\delta < \frac{2}{3}$.

We assumed $|E_1| \geqq |E_2|$, thus we may suppose

(12) $$|E_1| > \frac{n^2}{8}, \quad |V_1(x_0)| > \frac{n}{4}.$$

Put $|V_1(x_0)| = \frac{n}{4} + t$. If $|V_1(y_0)| > \frac{n}{12} - t$ then

$$|V_1(x_0) \cup V_1(y_0)| > \frac{n}{3},$$

i.e., $\delta < \frac{2}{3}$.

If $|V_1(y_0)| \leqq \frac{n}{12} - t$, then

$$d_1(x) \leqq \frac{n}{12} - t \quad \text{for} \quad x \in V_1(x_0).$$

This gives

$$|E_1| \leqq \frac{1}{2}\left(\frac{3n}{4} - t\right)\left(\frac{n}{4} + t\right) + \left(\frac{n}{4} + t\right)\left(\frac{n}{12} - t\right) =$$

$$= \frac{1}{2}\left(\frac{n}{4} + t\right)\left(\frac{5}{6}n - 2t\right) \leqq \frac{1}{2}\left(\frac{5}{24}n^2 + \frac{2}{3}nt - 2t^2\right) < \frac{n^2}{8},$$

which contradicts to (12).

This completes the proof of the upper bound of (3).

(b) *The lower bound* in (3) follows by the adaptation of a construction in P. ERDŐS [2]:

Let l be an integer which will be determined later, let the vertices of G be the $0-1$ sequences of length $3l+1$. Two vertices of G are joined by an edge in G if the Hamming-distance of the corresponding two sequences is at least $2l+1$ (i.e., if the sequences differ in at least $2l+1$ places). This graph has no triangle and it follows from a theorem of KLEITMAN [9] that the size of the maximum independent set equals the common degree of the vertices.

Now from this graph G we construct the graph G^* as follows: we replace each vertex by a set of vertices of size $\left[\frac{m}{2^{3l+1}}\right]$, where l is the smallest integer for which

$$\sum_{i=0}^{l+1} \binom{3l+1}{i} \frac{m}{2^{3l+1}} < \varepsilon m.$$

It is easy to see, that this graph has no triangles and the maximum independent set has $<\varepsilon m$ vertices. The number of edges in G^* is $>c\varepsilon m^2$ where $c>0$ is an absolute constant.

Now we consider the following three-colouring of K_{2m}:

Let $V=V_1\cup V_2$ with $|V_1|=|V_2|=m$. Let $G^*(V_1)$, $G^{**}(V_2)$ be two graphs isomorphic to the above constructed G^* and

$$E_2 = E\big(G^*(V_1)\big)\cup E\big(G^{**}(V_2)\big),$$

$G_1(V)$ be the complete bipartite graph $K(V_1, V_2)$.

This construction gives the proof of the lower bound in (3).

REMARK. Very likely the following stronger result holds: There is an absolute constant c such that $(\varepsilon\to 0)$

$$f(n;\, 3, 3, \varepsilon n) = \frac{n^2}{4} + \big(c+o(1)\big)\,\varepsilon n^2$$

but at the moment we do not know how to prove this.

PROOF of Theorem 2.

Now we construct a sequence of points $x_1, \dots, x_k$ and a corresponding sequence of indices $i_1, \dots, i_k$ where $i_\nu \in \{1, 2\}$, with the following property: for $\lambda=\sqrt{\varepsilon}$ let

$$\lambda_{i_1}(x_1) > \lambda n,$$

$$\lambda_{i_\nu}(x_\nu; U_\nu) > \lambda n \quad \text{if} \quad \nu > 1$$

where for $\nu>1$

$$U_\nu = V - \bigcup_{l=1}^{\nu} V_{i_l}(x_l).$$

Let $x_1, \ldots, x_k$ be maximal in the sense that for any $x \in V - \{x_1, \ldots, x_k\}$

$$d_i(x; U_k) < \lambda n.$$

Obviously, $k < \frac{1}{\lambda}$. Put

$$V_1 = \bigcup_{\substack{1 \le l \le k \\ i_l = 1}} V_{i_l}(x_l; U_l), \quad V_2 = \bigcup_{\substack{1 \le l \le k \\ i_l = 2}} V_{i_l}(x_l; U_l)$$

$$(V_1 \cap V_2 = \emptyset) \quad \text{and} \quad V_3 = V - (V_1 \cup V_2),$$

$n_i = |V_i|$, $1 \le i \le 3$.

Consider now the edges in $E_1 \cup E_2$ of the following type:

$$F_{ji,l} = \{(x, y)\colon\ x \in V_j,\ y \in V_l,\ (x, y) \in E_i\},$$

$1 \le j \le 3$, $1 \le l \le 3$, $i = 1, 2$.

(a) $$|F^1_{1,1}| \le \frac{1}{4} n_1^2, \quad |F^2_{2,2}| \le \frac{1}{4} n_2^2$$

since G_1 and G_2 are triangle-free.

(b) $$|F^1_{1,1}| < \lambda n^2, \quad |F^2_{1,2}| < \lambda n^2$$

$$|F^2_{1,1}| < \lambda n^2, \quad |F^1_{2,2}| < \lambda n^2.$$

Otherwise we would have two points x_ν and x_μ with

$$|V_1(x_\nu) \cap V_2(x_\mu)| > \frac{\lambda}{k} n > \lambda\lambda n = \varepsilon n$$

which contradicts (8).

(c) $$|F^i_{j,3}| < \lambda n^2 \quad \text{for} \quad j = 1, 2, 3, \quad i = 1, 2.$$

Otherwise we would have an $x \in V$ with

$$\max_{i=1,2} d_i(x; V_3) \ge \lambda n.$$

But since

$$d_i(x; V_3) = d_i(x; U_k)$$

this would contradict the maximality of the sequence $x_1, \ldots, x_k$.

By (a)—(c) we obtain

(13) $$|E_1| \le \frac{1}{4} n_1^2 + 10\lambda n^2$$

(14) $$|E_2| \le \frac{1}{4} n_2^2 + 10\lambda n^2.$$

Now by the assumption

$$|E_1| \ge |E_2| \ge cn^2$$

we get

$$n_i \ge 2n\sqrt{c - 10\lambda} \qquad i = 1, 2.$$

Hence by (13) and (14)

$$|E_1|+|E_2| \leqq n^2\left(\frac{1}{4}-\sqrt{c}+2c\right)+\eta(\varepsilon)n^2$$

where, as a simple computation shows, $\eta(\varepsilon)\to 0$ with $\varepsilon\to 0$.

REMARK. If $c>\frac{1}{16}$, there does not exist a three colouring of K_n, for which

$$K_3 \not\subset G_i, \qquad i=1,2$$

$$K_{\varepsilon n} \not\subset G_3$$

and

$$|E_1| \geqq |E_2| \geqq cn^2.$$

REMARK. First observe that the constant $\frac{1}{4}-\sqrt{c}+2c$ in Theorem 2 is best possible. To see this, let

$$V=V_1\cup V_2, \quad V_1\cap V_2=\emptyset,$$

$$|V_1|=[2\sqrt{c}n], \quad |V_2|=[(1-2\sqrt{c})n],$$

$$V_i=A_i\cup B_i, \quad |A_i|=\left[\frac{1}{2}|V_i|\right], \quad |B_i|=\left[\frac{1}{2}|V_i|+1\right], \quad i=1,2.$$

Join every vertex of A_i to every vertex of B_i in G_i (for $i=1,2$). Let the further edges of G_1, resp. G_2 form a graph on A_2 and on B_2, resp. on A_1 and on B_1 which has no triangle and the number of independent points is $o(n)$. (It is well-known, that such a graph exists and in fact we used this method in P. ERDŐS—V. T. SÓS [1] or in (6) of the proof of Theorem 1. Obviously this colouring has the required properties.

To get the exact result for $f(n;\,3,3,\varepsilon n)$ is rather hopeless because of its close connection with the Ramsey-numbers. This close connection is shown already by the following

PROPOSITION 1. *Let* $\varepsilon(n)\to 0$ *with* $n\to\infty$. *Then*

(15) $$R(3,\varepsilon(n)n)=o(n)$$

implies

(16) $$f(n;\,3,3,\varepsilon(n)n)=o(n^2).$$

(Here $R(k,l)$ is the Ramsey-number.)

PROOF. (a) Suppose $R(3,\varepsilon(n)n)=o(n)$ and that with a constant $c>0$

$$f(n;\,3,3,\varepsilon(n)n)>cn^2$$

holds. This means, that we have a three-colouring of K_n, for which

$$K_3 \not\subset G_i \qquad i=1,2$$

$$K_{\varepsilon(n)n} \not\subset G_3$$

and, e.g., $|E_1|>\frac{c}{2}n^2$. Thus we have a vertex x with $d_1(x)>cn$. Since $K_3\not\subset G_1$, in $V_1(x)$ we have only edges of E_2 and E_3.

But this means, that we have a two-colouring of the edges of K_{cn}, where in the first colour class there is no K_3 and in the second there is no $K_{\varepsilon(n)n}$. This contradicts (16).

The converse statement, that (16) implies (15) is probably true, too, but we could only prove the following weaker result:

Assume that

$$R(3, \varepsilon(n)n) > cn.$$

Then

$$f\left(n;\, 3, 3, \frac{\varepsilon(n)n}{2c}\right) > cn^2.$$

We hope to return to this subject later.

Some remarks on the Ramsey-numbers

As it is well-known, ERDŐS and SZEKERES [7] proved

$$R(k, l) \leqq \binom{k+l-2}{k-1}. \tag{17}$$

Probably (17) is not very far from being best possible, in particular

$$c_2\frac{n^2}{(\log n)^2} < R(3, n) < c_1\frac{n^2\log\log n}{\log n}.$$

It seems certain that

$$R(4, n) > n^{3-\varepsilon}. \tag{18}$$

The probability method surely must give (18) but so far technical difficulties prevented success.

GREENWOOD and GLEASON [8] proved

$$R_1(k_1+1, \ldots, k_r+1) \leqq \frac{(k_1+\ldots+k_r)!}{k_1!\ldots k_r!}.$$

This gives for example

$$R_3(3, 3, n) \leqq cn^4$$

and more generally

$$R_r(\underbrace{3, 3, \ldots, 3}_{1 \ \ldots \ r}, n) \leqq c_r n^{2r}.$$

A simple observation leads to the following improvement:

PROPOSITION 2.

$$R(3, 3, n) = o(n^3) \tag{19}$$

and more generally

(20) $$R_r(\underbrace{3, 3, \ldots, 3}_{r}, \underbrace{n}_{1}) \leqq rnR_{r-1}(\underbrace{3, \ldots, 3}_{r-1}, \underbrace{n}_{1}) = o(n^{r+1}).$$

PROOF. Let us consider a "good" r-colouring of K_m for $k_1 = \ldots = k_{r-1} = 3$, $k_r = n$. Let G_i, $1 \leqq i \leqq r$ the graph formed by the edges of the ith colour-class. Put

$$V_i(x) = \{y\colon (x, y) \in E_i\}, \quad 1 \leqq i \leqq r.$$

Let $U = \{x_1, \ldots, x_v\}$ be the vertex-set of a maximal-sized complete graph in G_r. We have $v \leqq n-1$. By the maximality of $|U|$ we have

$$\bigcup_{j=1}^{v} \bigcup_{i=1}^{r-1} V_i(x_j) = V - U.$$

Since G_i, $1 \leqq i \leqq r-1$ is triangle-free,

$$|V_i(x_j)| < R_{r-1}(3, \ldots, 3, n) \quad \text{for} \quad j = 1, \ldots, v.$$

Now taking into consideration $R(3, n) = o(n^2)$, this proves (20).

REMARK. We have no nontrivial lower bound for $R(3, 3, n)$. It is trivially true, that

$$R(3, 3, n) \geqq 2R(3, n).$$

We expect that

$$R(3, 3, n)/R(3, n) \to \infty$$

$$R(3, 3, n)n^{-2} \to \infty$$

or even more,

$$R(3, 3, n) > n^{3-\varepsilon}.$$

Some remarks on the two-colourings of K_n

The following problem belongs to the questions we considered in [5]. Let $f(n; G)$ be the smallest integer for which every graph of n vertices and of $f(n; G)$ edges contains a subgraph isomorphic to G and $f(n; G, \varepsilon n)$ be the smallest integer for which every graph of n vertices and $f(n; G, \varepsilon n)$ edges either contains a subgraph isomorphic to G or has an independent set of size εn.

First we investigate conditions which imply

(21) $$f(n; G, \varepsilon n) \leqq \eta n^2$$

where $\eta \to 0$ with $\varepsilon \to 0$ or

(22) $$f(n; G, \varepsilon n) < f(n; G)(1-c)$$

with a $c > 0$.

We prove some preliminary results about (21) and (22) and state without proof a few more results.

PROPOSITION 1. (21) *holds for* $G \sim K(1, r, r)$.

PROOF. We need the following result of Erdős:

For every l there exists a constant $c_l>0$ such that if $n>n_0$ and $e(G_n)>cn^2$ then G_n contains a $K(l, c_l, n)$.

Using this it is easy to show that if for G_n $e(G_n)=cn^2$ and the largest independent set in G_n has size less than $\varepsilon(c)n$, then G_n contains a $K(1, r, r)$.

PROPOSITION.

$$f(n; K(3, 3, 3), \varepsilon n) = \frac{n^2}{4}(1+\eta)$$

where $\eta\to 0$ *with* $\varepsilon\to 0$.

PROOF. The stronger

$$f(n; K(3, 3, 3)) \leqq \frac{n^2}{4}(1+\eta)$$

follows from ERDŐS—STONE [6].

We can prove the lower bound as follows:

Let $|V_1|=\left[\frac{n}{2}\right]$, $|V_2|=\left[\frac{n+1}{2}\right]$. We join every vertex of V_1 to every vertex of V_2. Additionally on V_1 resp. on V_2 we consider a graph whose largest independent set has size εn and which contains no circuit C_r with $3\leqq r\leqq 5$. (We know the existence of such a graph from [3], [4].) This graph contains no $K(3, 3, 3)$ since the vertex set of $K(3, 3, 3)$ cannot be decomposed into two sets neither of which spans a graph without a circuit.

In a forthcoming paper we prove the more general

THEOREM A. *Let G be a graph which is k-chromatic and the vertex-set can be decomposed into $k-1$ sets which span graphs without circuits. Then there is a* $c>0$ *such that*

$$f(n; G, \varepsilon n) \leqq \frac{n^2}{2}\left(1-\frac{1}{k-1}-c\right)$$

for $\varepsilon<\varepsilon_0, n>n_0$.

As to (22) we prove

THEOREM B. *Let G be a graph which is k-chromatic and the vertex-set of G cannot be decomposed into $k-1$ sets such that the subgraphs spanned by these sets have no circuit. Then for every* $\eta>0$

$$f(n; G, \varepsilon n) \geqq \frac{n^2}{2}\left(1-\frac{1}{k-1}-\eta\right)$$

if $\varepsilon<\varepsilon_0(\eta), n>n_0(\eta)$.

Added in proof (December, 1981). We proved with A. Hajnal and E. Szemerédi that

$$f(n; 2k, l) = \frac{1}{2}\left(\frac{3k-5}{3k-2}\right)n^2+o(n^2) \quad \text{for} \quad k \geqq 2$$

when $l=o(n)$. The proof will appear in a quadruple paper in *Combinatorica.*

REFERENCES

[1] Bollobás, B. and Erdős, P., On a Ramsey—Turán type problem, *J. C. J.* (B) **21** (1976), 166—168.
[2] Erdős, P., On the construction of certain graphs, *J. Combinatorial Theory* **1** (1966), 149—153.
[3] Erdős, P., Graph theory and probability I, *Canadian J. Math.* **11** (1959), 34—38.
[4] Erdős, P.: On circuits and subgraphs of chromatic graphs, *Math.* **9** (1962), 170—175.
[5] Erdős, P. and T. Sós, V., Some remarks on Ramsey's and Turán's theorem, *Combinatorial Theory and its Applications*, Coll. Math. Soc. J. Bolyai, Balatonfüred, Hungary, 1969, 395—404.
[6] Erdős, P. and Stone, E. H., On the structure of linear graphs, *Bull. Amer. Math. Soc.* **52** (1946), 1087—1091.
[7] Erdős, P. and Szekeres, G., A combinatorial problem in geometry, *Compositio Math.* **2** (1935), 463—470.
[8] Greenwood, A. M. and Gleason, A. M., Combinatorial relations and chromatic graphs, *Canadian J. Math.* **7** (1955), 1—7.
[9] Kleitman, D. J., Families of non-disjoint subsets, *J. Combinatorial Theory* **1** (1966), 153—155.
[10] T. Sós, V., On extremal problems in graph theory, *Proc. Calgary Internat. Conf. on Combinatorial Structures*, 1969, 407—410.
[11] Szemerédi, E., Graphs without complete quadrilaterals, *Mat. Lapok* **23** (1973), 113—116 (in Hungarian).

(Received April 10, 1980)

MTA MATEMATIKAI KUTATÓ INTÉZETE
REÁLTANODA U. 13—15.
H—1053 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 37—41

SUFFICIENCY AND f-DIVERGENCES

by

D. MUSSMANN

0. Introduction and summary

We shall characterize sufficiency of a sub-σ-field for two probability measures by f-divergences, where f is a convex function which is not affine. This extends a theorem of CSISZÁR. As a corollary we obtain a criterion for sufficiency in terms of total variations. This criterion may be applied to prove PFANZAGL's characterization of sufficiency by power functions. An essential tool is a result on the attainment of equality in Jensen's inequality for conditional expectations for convex functions which are not necessarily strictly convex.

1. Preliminaries

The first part of this section is concerned with strict inequality in some well-known inequalities for convex functions. In the second part we state Neyman's criterion for sufficiency in a form suitable for our purposes.

Let f be a convex (continuous) function defined on $I:=]a,b[$, where $-\infty\leqq\leqq a<b\leqq+\infty$. Furthermore we shall suppose f to be not affine, i.e. not of type $\alpha x+\beta$. We denote the right derivative of f at $x\in I$ by $D^+(f;x)$ and define $[xy](f):=(f(x)-f(y))/(x-y)$, where $x,y\in I$ such that $x\neq y$. If a, resp. b, is finite, we define $f(a):=\lim\limits_{r\downarrow a}f(r)$, resp. $f(b):=\lim\limits_{r\uparrow b}f(r)$; $f(a)$, $f(b)$ may be $+\infty$.

The assumption that f is not affine is equivalent to each of the following two conditions:

(1.1) There exists an $x_0\in I$ such that

$$f(x_0)<\frac{y-x_0}{y-x}f(x)+\frac{x_0-x}{y-x}f(y)\quad\text{whenever}\quad x,y\in I\quad\text{and}\quad x<x_0<y.$$

(1.2) There exists an $x_0\in I$ such that

$$f\big(tx+(1-t)\,y\big)<tf(x)+(1-t)f(y)$$

whenever

$$x,y\in I,\quad x<x_0<y,\quad\text{and}\quad t\in]0,1[.$$

Otherwise every $x_0\in I$ would be contained in an open interval on which f is affine (see [1], p. 232, Satz 2), hence f would be affine on I.

By the same reasoning as in [1], p. 197, 8.3.4.3 Satz, we get from (1.2):

(1.3) $[rx](f)<[yx](f)<[yr](f)$ whenever $x,y,r\in I$, $x<x_0<y$, and $x<r<y$ (same x_0 as in (1.2)).

Again by the same reasoning as in [1], p. 198, Satz (3a), we obtain from (1.3):

(1.4) $f(y) > D^+(f; x)(y-x)+f(x)$ whenever $x, y \in I$ and $x < x_0 < y$ or $y < x_0 < x$.

The strict inequality in (1.4) is even valid if a, resp. b, is finite and $y=a$, resp. $y=b$, because $[yx](f)$ is isotone in y.

If $b=\infty$, we have $\lim_{y\to\infty}[ry](f)=\lim_{y\to\infty}\frac{f(y)}{y}$. Combining this fact with (1.3), we have:

(1.5) $f(r) < f(x)+(r-x)\lim_{y\to\infty}\frac{f(y)}{y}$ whenever $r, x \in I$, $x < x_0$, and $x < r$.

We close this section with a variant of Neyman's criterion. Let P and Q be probability measures defined on a measurable space $(\Omega, \mathscr{A})$ and let $P=\mu_1+\mu_2$ be the Lebesgue decomposition of P with respect to Q, $\mu_1 \ll Q$ and $\mu_2 \perp Q$.

A sub-σ-field $\mathscr{S}$ of $\mathscr{A}$ is sufficient for P, Q iff

(1.6) there is an $\mathscr{S}$-measurable $\frac{d\mu_1}{dQ}$ and a $T\in\mathscr{S}$ such that $\mu_2(T)=0$ and $Q(T)=1$.

This assertion seems to be known. Concerning the "if" part, observe that the density with respect to μ_2+Q of P, resp. Q, is $(1-1_T)+1_T\cdot\frac{d\mu_1}{dQ}$, resp. 1_T.

2. Equality in Jensen's inequality

We employ the same notation as in 1; the point x_0 is of special importance. In the following, expectations are understood with respect to the probability measure Q. Let X be a Q-integrable function with values in $\bar{I}$ (closure of I in $]-\infty, +\infty[$).

(2.1) THEOREM. *If there is equality in Jensen's inequality* $f(\mathbf{E}(X|\mathscr{S}))\leqq \mathbf{E}(f\cdot X|\mathscr{S})[Q]$, *then* $Q\{X<x_0<\mathbf{E}(X|\mathscr{S})\}=0$ *and* $Q\{\mathbf{E}(X|\mathscr{S})<x_0<X\}=0$.

PROOF. It is known that there exists a Markov kernel φ from $(\Omega, \mathscr{A})$ to $(\bar{I}, \mathscr{B}_1\cap\bar{I})$ such that $\mathbf{E}(g\circ X|\mathscr{S})=\int g(r)\varphi(\cdot, dr)[Q]$ whenever $\int g\circ X dQ$ exists.

Define $h(\omega):=\int r\varphi(\omega, dr)$, we may suppose $-\infty<h(\omega)<\infty$ for all $\omega\in\Omega$. Our premise can be rewritten in the form:

(2.2) There exists an $N\in\mathscr{A}$ of Q-measure zero such that $f(h(\omega))=\int f(r)\varphi(\omega, dr)$ unless $\omega\in N$.

Now suppose ω to be not in N. We shall show

(2.3) $\varphi(\omega,]x_0, b])=0$ whenever $a\leqq h(\omega)<x_0$, and $\varphi(\omega, [a, x_0[)=0$ whenever $x_0<h(\omega)\leqq b$.

First we assume $h(\omega)\notin I$, i.e. $-\infty<a=h(\omega)$ or $h(\omega)=b<+\infty$, then $\varphi(\omega, \cdot)$ equals Dirac measure in a, resp. in b, hence (2.3) is valid.

Suppose next that $h(\omega)\in I$. Replacing x with $h(\omega)$ in (1.4) we obtain $f(y)> D^+(f; h(\omega))(y-h(\omega))+f(h(\omega))$ whenever $y\in\bar{I}$ and $h(\omega)<x_0<y$ or $y<x_0<h(\omega)$ ($\geqq$ holds for every $y\in\bar{I}$). Integrating with respect to $\varphi(\omega, \cdot)$ the assertion (2.3) follows.

To conclude the proof observe that for $A\in\mathcal{A}$ we have $\mathbf{E}(1_A|\mathcal{S})=\varphi(\cdot, A)[Q]$ and then using (2.3) we infer

$$Q\{X<x_0<\mathbf{E}(X|\mathcal{S})\} = \int 1_{\{h>x_0\}}\cdot 1_{\{X<x_0\}}\,dQ = \int 1_{\{h>x_0\}}\varphi(\cdot, [a, x_0[)\,dQ = 0.$$

In an analogous manner we get $Q\{\mathbf{E}(X|\mathcal{S})<x_0<X\}=0$.

REMARK. Pfanzagl ([6], p. 493, Theorem 2) proved that for strictly convex functions equality in Jensen's inequality implies $\mathbf{E}(X|\mathcal{S})=X[Q]$. This is easily obtained from (2.1). Indeed, for a strictly convex f the $x_0\in I$ is arbitrary, hence, if M denotes the rational numbers of I, then

$$Q\{X\neq\mathbf{E}(X|\mathcal{S})\}\leqq\sum_{r\in M}Q(\{X<r<\mathbf{E}(X|\mathcal{S})\}\cup\{\mathbf{E}(X|\mathcal{S})<r<X\})=0.$$

(2.4) COROLLARY. *Suppose $a=0$ and $b=+\infty$. Let D denote a countable dense subset of I.*

If $f(\mathbf{E}(\beta X|\mathcal{S}))=\mathbf{E}(f\circ(\beta X)|\mathcal{S})[Q]$ for all $\beta\in D$, then $\mathbf{E}(X|\mathcal{S})=X[Q]$.

PROOF. If D is dense in I, so is $\left\{\frac{x_0}{\beta}\colon \beta\in D\right\}$. Replacing X with βX in (2.1) we get

$$Q\{\mathbf{E}(X|\mathcal{S})\neq X\}\leqq\sum_{\beta\in D}Q\left(\left\{\mathbf{E}(X|\mathcal{S})<\frac{x_0}{\beta}<X\right\}\cup\left\{X<\frac{x_0}{\beta}<\mathbf{E}(X|\mathcal{S})\right\}\right)=0.$$

REMARK. If X is bounded, $X<M<\infty$, then in (2.4) we only need $\left\{\frac{x_0}{\beta}\colon \beta\in D\right\}$ to be dense in $]0, M[$.

Now we shall investigate equality in another inequality. Suppose $a=0$ and $b=+\infty$ and let Y and Z be measurable functions on Ω with values in $[0,\infty[$. Replacing x with $Y(\omega)$ and r with $Y(\omega)+Z(\omega)$ in (1.5) we obtain:

(2.5) *If $f(Y+Z)=f(Y)+Z\lim\limits_{s\to\infty}\frac{f(s)}{s}[Q]$, then $Q\{Y<x_0,\ Z>0\}=0$.*

Furthermore, let D denote a countable subset of I with cluster point zero.

(2.6) *If $f(\beta Y+\beta Z)=f(\beta Y)+\beta Z\lim\limits_{s\to\infty}\frac{f(s)}{s}$ for all $\beta\in D$, then $Q\{Z>0\}=0$.*

This holds because of (2.5) and $\Omega=\bigcup\limits_{\beta\in D}\{\beta Y<x_0\}$.

REMARK. If f is strictly convex, x_0 is arbitrary, and therefore $Q\{Z>0\}=0$ follows even from the premise of (2.5). If Y is bounded, we need only one β in (2.6).

3. The main result and applications

In this section we present a condition in terms of f-divergences which is equivalent to sufficiency. This is an extension of a result obtained by Csiszár for a strictly convex f. Our condition is easily seen to be fulfilled if a certain condition for total variations or Pfanzagl's condition for power functions hold.

We employ the same notation as in 1. We suppose $a=0$ and $b=\infty$. For $\beta>0$ the Lebesgue decomposition of βP with respect to Q is $\beta P=\beta\mu_1+\beta\mu_2$.

(3.1) DEFINITION. $I_f(\beta P, Q):=\int f\circ\left(\beta\frac{d\mu_1}{dQ}\right)dQ+\beta\mu_2(\Omega)\lim_{s\to\infty}\frac{f(s)}{s}$ is called the *f-divergence* of βP with respect to Q (see [2], [3], [4]).

Set $P':=P|\mathcal{S}$ and in the same manner Q', μ_1', and μ_2'.

(3.2) THEOREM. *Let D denote a countable dense subset of I. Suppose $I_f(\beta P', Q')<<\infty$ for every $\beta\in D$. $\mathcal{S}$ is sufficient for P, Q iff $I_f(\beta P', Q')=I_f(\beta P, Q)$ for every $\beta\in D$.*

PROOF. The "only if" part is known (and follows at once from (1.6)), and the "if" part is known for strictly convex f with $\beta=1$ ([2], p. 90, Satz 1; [3], p. 310; [4], p. 141, Satz 17.2).

By the same reasoning as in [4], p. 145, one shows:

$$(3.3)\quad I_f(\beta P', Q')=I_f(\beta P, Q)\quad\text{iff}\quad f\circ\mathbf{E}\left(\beta\frac{d\mu_1}{dQ}\bigg|\mathcal{S}\right)=\mathbf{E}\left(f\circ\left(\beta\frac{d\mu_1}{dQ}\right)\bigg|\mathcal{S}\right)[Q]\quad\text{and}$$

$$f\left(\mathbf{E}\left(\beta\frac{d\mu_1}{dQ}\bigg|\mathcal{S}\right)+\beta\frac{dm_1}{dQ'}\right)=f\left(\mathbf{E}\left(\beta\frac{d\mu_1}{dQ}\bigg|\mathcal{S}\right)\right)+\beta\frac{dm_1}{dQ'}\lim_{s\to\infty}\frac{f(s)}{s}[Q]$$

where m_1 denotes with respect to Q' absolutely continuous part of μ_2'.

Now, replacing X with $\frac{d\mu_1}{dQ}$ in (2.4) we get $\mathbf{E}\left(\frac{d\mu_1}{dQ}\bigg|\mathcal{S}\right)=\frac{d\mu_1}{dQ}[Q]$, hence there exists an $\mathcal{S}$-measurable $\frac{d\mu_1}{dQ}$.

Replacing Y with $\mathbf{E}\left(\frac{d\mu_1}{dQ}\bigg|\mathcal{S}\right)$ and Z with $\frac{dm_1}{dQ'}$ in (2.6) we obtain $\frac{dm_1}{dQ'}=0\,[Q]$, whence $Q'\perp\mu_2'$, i.e. there exists a $T\in\mathcal{S}$ such that $Q(T)=Q'(T)=1$ and $\mu_2(T)==\mu_2'(T)=0$. Then the sufficiency of $\mathcal{S}$ follows by (1.6).

(3.4) COROLLARY. *$\mathcal{S}$ is sufficient for P, Q iff $\|\beta P-Q\|=\|\beta P'-Q'\|$ for all $\beta\in D$ (D as in (3.2); $\|\ \|$ denotes the total variation).*

PROOF. Take $f(x)=|x-1|$. Let p and q be densities of P, resp. Q, with respect to $\nu:=P+Q$. It is straightforward to show that $I_f(\beta P, Q)=\int|\beta p-q|\,d\nu$ and

$$\|\beta P-Q\|=\max\left\{\frac{1}{2}\int|\beta p-q|\,d\nu+\frac{1}{2}(\beta-1),\ \frac{1}{2}\int|\beta p-q|\,d\nu-\frac{1}{2}(\beta-1)\right\}.$$

Hence $\|\beta P-Q\|=\|\beta P'-Q'\|$ iff

$$I_f(\beta P, Q)=I_f(\beta P', Q').$$

In order to characterize sufficiency, Pfanzagl ([5], p. 197) presents the following condition:

(3.5) For every $A\in\mathcal{A}$ there exists an $\mathcal{S}$-measurable test φ such that $P(A)\geqq\geqq\int\varphi\,dP$ and $Q(A)\leqq\int\varphi\,dQ$.

If (3.5) holds, then obviously for every $A \in \mathscr{A}$ there exists an $\mathscr{S}$-measurable test ψ such that $P(A) \leqq \int \psi \, dP$ and $Q(A) \geqq \int \psi \, dQ$.

The inequalities may be multiplied by $\beta > 0$. It follows that for every $A \in \mathscr{A}$ there exists an $\mathscr{S}$-measurable test χ such that

$$|\beta P(A) - Q(A)| \leqq \left|\beta \int \chi \, dP - \int \chi \, dQ\right| \leqq \|\beta P' - Q'\|,$$

hence $\|\beta P - Q\| = \|\beta P' - Q'\|$ and (3.4) can be applied.

The following example shows that (3.4) is not valid with $D = \{1\}$.

EXAMPLE. $\Omega = \{1, 2, 3\}$, $\mathscr{A} = \mathscr{P}(\Omega)$. Let P, resp. Q, be given by $(p_1, p_2, p_3) = \left(\frac{1}{8}, \frac{1}{4}, \frac{5}{8}\right)$, resp. $(q_1, q_2, q_3) = \left(\frac{5}{8}, \frac{1}{4}, \frac{1}{8}\right)$. Let $\mathscr{S}$ denote the σ-field generated by $\{3\}$. $\frac{dP}{dQ}$ takes the values $\frac{1}{5}$, 1, 5, whence $\frac{dP}{dQ}$ is not $\mathscr{S}$-measurable and $\mathscr{S}$ not sufficient for P, Q. But $\|P - Q\| = \frac{4}{8} = \|P' - Q'\|$ and furthermore $\|\lambda - \mu\| = \|\lambda' - \mu'\|$ when λ and μ are in the convex hull of $\{P, Q\}$.

I wish to thank D. PLACHKY who suggested Corollary 3.4 and its application to Pfanzagl's characterization of sufficiency by power functions.

REFERENCES

[1] AUMANN, G. und HAUPT, D.: *Einführung in die reelle Analysis I*, De Gruyter, Berlin—New York, (1974).

[2] CSISZÁR, I.: Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten, *Publ. Math. Inst. Hung. Acad. Sci.* **8** Ser. A (1963), 85—107.

[3] CSISZÁR, I.: Information-type measures of difference of probability distributions and indirect observations, *Studia Sci. Math. Hung.* **2** (1967), 299—318.

[4] HEYER, H.: *Mathematische Theorie statistischer Experimente*, Springer, Berlin—Heidelberg—New York, (1973).

[5] PFANZAGL, J.: A characterization of sufficiency by power function, *Metrika* **21** (1974), 197—199.

[6] PFANZAGL, J.: Convexity and conditional expectations, *Ann. Prob.* **2** (1974), 490—494.

(Eingegangen am 28. Juni 1977)

INSTITUT FÜR MATHEMATISCHE STATISTIK DER
WESTFÄLISCHEN WILHELMS-UNIVERSITÄT MÜNSTER
ROXELER STRASSE 64
D—4400 MÜNSTER
WEST GERMANY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 43—46

SUBBASE AND DIMENSION II*

by

J. DEÁK

Inequalities between topological dimension functions investigated in the first part of this series (*Studia Sci. Math. Hungar.* **11** (1976), 389—397) will be proved here for larger classes of spaces. The reader is supposed to have the first part at hand. A reference like (I.3.2) will mean (3.2) of the first part. All results of the present paper are consequences of Theorem (2.1).

§ 1. Definitions

Before introducing a new dimension function, we define some generalizations of total paracompactness. This notions are (with the possible exception of A-TPC) not new.

(1.1) DEFINITION. Let $\mathscr{Y}$ and $\mathscr{Z}$ be systems of sets in a space.

a) $\mathscr{Y}$ is a *tight refinement* of $\mathscr{Z}$ if for each $Y\in\mathscr{Y}$ there is a $Z\in\mathscr{Z}$ with $Y\subset Z$ and $\operatorname{Fr} Y\subset \operatorname{Fr} Z$.

b) $\mathscr{Y}$ is a *boundary-refinement* of $\mathscr{Z}$ if for each $Y\in\mathscr{Y}$ there is a locally finite collection of closed sets covering $\operatorname{Fr} Y$ and refining the system of the boundaries of the elements of $\mathscr{Z}$. (A boundary-refinement need not be a refinement!)

Recall that a space X is *totally paracompact* (TPC) if each of its open bases has a locally finite subsystem covering X.

(1.2) DEFINITION. A space is *almost totally paracompact* (A-TPC) if each of its open bases has a locally finite open tight refinement covering the space.

(1.3) DEFINITION. (NAGAMI [13].) A space is *σ-totally paracompact* (σ-TPC) if each of its open bases has a σ-locally finite open tight refinement covering the space.

(1.4) DEFINITION. (FITZPATRICK and FORD [11]**.) A space is *order-totally paracompact* (O-TPC) if each of its open bases has a totally ordered (not necessarily

---

* *Added in proof.* Since our paper „On the generalizations of total paracompactness" submitted for publication more than a year later has already appeared [*Studia Sci. Math. Hungar.* **13** (1978), 393—409], we delete the proof of Theorem (2.1) for it is an immediate consequence of Theorem 3H of the above-mentioned paper where some of the results in §3 of the present paper are improved on as well, but this overlapping cannot be remedied just by omitting parts of §3, so we leave it in its original form.

** Fitzpatrick and Ford give a somewhat different definition; FRENCH [12], however, uses O-TPC in this sense (cf. the Remark at the end of [11]).

well-ordered) open tight refinement $\mathscr{V}$ covering the space such that for each $V \in \mathscr{V}$, the elements of $\mathscr{V}$ preceding V form a system locally finite at the points of V.

(1.5) Definition. (French [12].) A space X is *closure-totally paracompact* (C-TPC) if an arbitrary open base $\mathscr{B}$ of X has a locally finite closed boundary-refinement covering X and refining the system of the closures of the elements of $\mathscr{B}$.

Note that the word "closed" can be omitted from Definition (1.5). The following diagram shows the relations between these notions (some of the implications are trivial, others are proved in French [12]):

$$\begin{array}{ccc} & \sigma\text{-TPC} \Rightarrow & \text{O-TPC} \\ & \Uparrow & \Downarrow \\ \text{TPC} \Rightarrow & \text{A-TPC} & \text{C-TPC} \\ & \Downarrow & \Downarrow S_3 \\ & \multicolumn{2}{c}{\text{paracompact}} \end{array}$$

The theorems of the first part are all corollaries to Theorem (I.3.3). In order to generalize the results of the first part, we prove a theorem similar to Theorem (I.3.3) with ind^+ replaced by ind° defined below.

(1.6) Definition. For a space X, $\operatorname{ind}^\circ X = -1$ iff $X = \emptyset$; $\operatorname{ind}^\circ X \leqq n$ iff X has a base $\mathscr{B}$ with

$$\operatorname{ind}^\circ \operatorname{Fr} \bigcup \mathscr{L} \leqq n-1$$

for an arbitrary locally finite boundary-refinement $\mathscr{L}$ of $\mathscr{B}$ $(n=0, 1, 2, \ldots)$; $\operatorname{ind}^\circ X = n$ iff $\operatorname{ind}^\circ X \leqq n$ but $\operatorname{ind}^\circ X \leqq n-1$ is not true.

The only difference between the definitions of ind^+ and ind° is that the word "subsystem" is replaced by "boundary-refinement". Since a subsystem is a boundary-refinement, $\operatorname{ind}^+ X \leqq \operatorname{ind}^\circ X$ (proof by induction).

§ 2

(2.1) Theorem. *For an arbitrary space X,*

$$\operatorname{ind}^\circ X \leqq \operatorname{sbd} X.$$

As $\operatorname{ind}^+ X \leqq \operatorname{ind}^\circ X$, Theorem (I.3.3) is a corollary to this theorem. The proof is similar to that of Theorem (I.3.3).

§ 3

In this section we prove theorems similar to those of the first part.

(3.1) Proposition. *If all closed subspaces of a normal space X are* C-TPC, *then*

$$\operatorname{Ind} X \leqq \operatorname{ind}^\circ X.$$

Proof. The inequality is evidently true if $\operatorname{ind}^\circ X = -1$. Let now $\operatorname{ind}^\circ X = n \geqq 0$ and suppose that $\operatorname{Ind} Y \leqq \operatorname{ind}^\circ Y$ holds for any normal space Y with the closed subspaces of it C-TPC satisfying $\operatorname{ind}^\circ Y \leqq n-1$. Let $\mathscr{B}$ be a base of X as described in Definition (1.6). Let F be a closed and G an open subset of X with $F \subset G$. Take open sets V and V_1 with

$$F \subset V \subset \overline{V} \subset V_1 \subset \overline{V}_1 \subset G.$$

We may suppose that

$$\mathscr{B} < \{V_1, X-V\}.$$

Since X is C-TPC, there is a locally finite closed covering $\mathscr{F}$ of X such that $\mathscr{F}$ is a boundary-refinement of $\mathscr{B}$ and

$$\mathscr{F} < \overline{\mathscr{B}} = \{\overline{B}\colon\ B \in \mathscr{B}\}.$$

Let now $\mathscr{C}$ be the system of those elements of $\mathscr{F}$ which intersect V. $\mathscr{C}$ is a locally finite boundary-refinement of $\mathscr{B}$, so

$$\operatorname{ind}^\circ \operatorname{Fr} \bigcup \mathscr{C} \leqq n-1.$$

$\bigcup \mathscr{C}$ is a closed subset of X, thus it is normal C-TPC and

$$\operatorname{Ind} \operatorname{Fr} \bigcup \mathscr{C} \leqq n-1,$$

according to the induction assumption. Further,

$$V \subset \bigcup \mathscr{C} \subset G,$$

thus with

$$A = \operatorname{int} \bigcup \mathscr{C}$$

we have an open set A satisfying

$$F \subset A \subset G.$$

$\operatorname{Fr} A$ is a closed subset of $\operatorname{Fr} \bigcup \mathscr{C}$, thus

$$\operatorname{Ind} \operatorname{Fr} A \leqq \operatorname{Ind} \operatorname{Fr} \bigcup \mathscr{C} \leqq n-1$$

and $\operatorname{Ind} X \leqq n$.

(3.2) PROPOSITION. *For a* σ-TPC S_3*-space* X,

$$\operatorname{Ind} X \leqq \operatorname{ind}^\circ X.$$

PROOF. A σ-TPC S_3-space is paracompact, so X is normal. The closed subspaces of X are σ-TPC, thus they are C-TPC and Proposition (3.1) can be applied.

(3.3) PROPOSITION. *For an* A-TPC *space* X,

$$\operatorname{Ind} X \leqq \operatorname{ind}^\circ X.$$

Note that the regularity of X is not required here.

PROOF. Suppose that the proposition is true for A-TPC spaces Y with $\operatorname{ind}^\circ Y \leqq n-1$, $n \geqq 0$ and let $\operatorname{ind}^\circ X = n$. Let F be a closed and G an open set, $F \subset G$. Take a base $\mathscr{B}$ according to Definition (1.6). We may suppose that

$$\mathscr{B} < \{G, X-F\}.$$

X is A-TPC, so there is a locally finite tight open refinement $\mathscr{V}$ of $\mathscr{B}$ covering X. Let $\mathscr{W}$ be the system of those elements of $\mathscr{V}$ which intersect F. Then

$$F \subset \bigcup \mathscr{W} \subset G.$$

$\mathscr{W}$ is a locally finite tight refinement of $\mathscr{B}$, so it is a locally finite boundary-refinement of $\mathscr{B}$ and

$$\operatorname{ind}^\circ \operatorname{Fr} \bigcup \mathscr{W} \leqq n-1.$$

The induction assumption can now be applied to $\operatorname{Fr} \bigcup \mathscr{W}$, since a closed subset of an A-TPC space is A-TPC as well.

Now we generalize Theorems (I.2.2), (I.4.9) and (I.4.7).

(3.4) THEOREM. *If X is a σ*-TPC *S_3-space or an* A-TPC *space, then*

$$\operatorname{Ind} X \leqq \operatorname{Dim} X.$$

PROOF. Propositions (3.2) and (3.3), Theorem (2.1) and Proposition (I.4.3).

(3.5) THEOREM. *If X is a σ*-TPC *S_3-space or an* A-TPC *S_1-space, then*

$$\operatorname{Ind} X \leqq \operatorname{DIM} X.$$

PROOF. Propositions (3.2) and (3.3), Theorem (2.1) and Proposition (I.4.4).

(3.6) COROLLARY. *If X is a regular Lindelöf space, then*

$$\operatorname{Ind} X \leqq \operatorname{Dim} X, \quad \operatorname{Ind} X \leqq \operatorname{DIM} X.$$

PROOF. A Lindelöf space is evidently σ-TPC.

(3.7) THEOREM. *If X is a σ*-TPC *S_3-space or an* A-TPC *S_1-space and it has an* ind-*nice subbase, then*

$$\operatorname{ind} X = \operatorname{Ind} X.$$

PROOF. Propositions (3.2) and (3.3), Theorem (2.1) and Proposition (I.4.5).

(3.8) PROPOSITION. *For a separable metric space M,*

$$\operatorname{ind}^\circ M = \operatorname{sbd} M = \operatorname{Ind} M = \operatorname{ind}^+ M$$

if any of these dimensions is finite.

PROOF. As the proof of Proposition (I.5.2) (note that M is Lindelöf).

REFERENCES

[1]...[10]: see at the end of the first part.
[11] FITZPATRICK, B. jr and FORD, R. M.: On the equivalence of small and large inductive dimension in certain metric spaces, *Duke Math. J.* **34** (1967), 33—38.
[12] FRENCH, J. A.: Coincidence of small and large inductive dimension, *Topo 72, Lecture Notes* **378** (1974), 132—139.
[13] NAGAMI, K.: A note on the large inductive dimension of totally normal spaces, *J. Math. Soc. Japan* **21** (1969), 282—290.

(Received October 1, 1977)

MTA MATEMATIKAI KUTATÓ INTÉZETE
REÁLTANODA U. 13—15
H—1053 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 47—49

IF THE INTERSECTION OF ANY r SETS HAS A SIZE $\neq r-1$

by

P. FRANKL and G. O. H. KATONA

Working on problems connected with data base systems, suggested by J. DEMETROVICS, we observed the following simple but interesting

THEOREM 1. *Let $A_1, \ldots, A_m$ be a family of not necessarily distinct subsets of an n element set X. Suppose that*

$$\left|\bigcap_{j=1}^{r} A_{i_j}\right| \neq r-1 \tag{1}$$

holds for any r $(1 \leqq r \leqq m)$ and any distinct indices $i_1, \ldots, i_r$ $(1 \leqq i_j \leqq m)$. Then

$$m \leqq n. \tag{2}$$

PROOF. We use induction over n. $d(x)$ is the degree of $x \in X$: $d(x) = |\{j\colon x \in A_j\}|$.

1. We prove first that $d(x) \leqq |A_i|$ follows from $x \in A_i$.

Fix an i and $x \in A_i$. Take the sets $A_i \cap A_j - \{x\}$ for all $j \neq i$ such that $x \in A_j$. If these sets do not satisfy (1), there are indices $j_1, \ldots, j_r$ such that $x \in A_{j_l}$ $(1 \leqq l \leqq r)$ and

$$\left|\bigcap_{l=1}^{r} (A_i \cap A_{j_l} - \{x\})\right| = r-1.$$

Hence $\left|A_i \cap \bigcap_{l=1}^{r} (A_i \cap A_{j_l})\right| = r$ would follow contradicting (1). The sets $A_i \cap A_j - \{x\}$ satisfy (1) on a set of size $|A_i| - 1 \leqq n-1$: so we may use the inductional hypothesis: the number of sets $x \in A_j$, $j \neq i$ is $\leqq |A_i| - 1$. Thus the number of sets $x \in A_j$ is $\leqq |A_i|$.

2. It follows from the induction hypothesis that the union of any r of the sets A_i has a size at least r if $r \leqq n$. By Hall's theorem we obtain elements $x_i \in A_i$ $(1 \leqq i \leqq r)$, where $x_1, \ldots, x_n$ lists all the elements of X. The first section gives

$$d(x_i) \leqq |A_i|.$$

Hence

$$\sum_{i=1}^{n} |A_i| \geqq \sum_{i=1}^{n} d(x_i) = \sum_{i=1}^{m} |A_i|$$

and

$$\sum_{i=n+1}^{m} |A_i| = 0$$

follows. $|A_i| \neq 0$ by (1), consequently the sum must be empty. We have $m \leqq n$, and the proof is complete.

The n different one-element sets give equality in (2).

COROLLARY. *If $A_1, \ldots, A_m$ are non-empty subsets of a set of n elements, no two have an intersection equal to* 1 *and no three have an intersection* >1 *then* $m \leqq n$.

PROOF. It is easy to see that the sets satisfy (1).

THEOREM 2. *Let $A_1, \ldots, A_m$ be a family of not necessarily distinct subsets of an n element set X, and let $t>0$ be a fixed integer. Suppose that*

$$\left|\bigcap_{j=1}^{r} A_{i_j}\right| \neq r-1-t \tag{3}$$

holds for any r $(1 \leqq r \leqq m)$ and any distinct indices $i_1, \ldots, i_r$ $(1 \leqq i_j \leqq m)$. Then

$$m \leqq n+t \tag{4}$$

moreover

$$m \leqq n \tag{5}$$

with the additional conditions $A_i \neq A_j$ $(i \neq j)$ and $2^{t-1} \leqq n$.

PROOF. Take the sets $\left(\bigcap_{i=1}^{t} A_i\right) \cap A_j$ $(t<j \leqq m)$. The intersection of any r different ones cannot be of size $r-1$ by (3). Apply Theorem 1 for these sets: $m-t \leqq n$. The choice $A_i = X$ $(1 \leqq i \leqq n+t)$ gives equality in (4).

The proof of (5) proceeds in a similar way. The only difference is that we have to choose some distinct sets $A_{i_1}, \ldots, A_{i_t}$ with $\left|\bigcap_{j=1}^{t} A_{i_j}\right| \leqq n-t$. It can be proved by induction over t (with fixed n) that this can be done if $m \geqq n+1$: By the inductional hypothesis we can find $t-1$ sets with an intersection Y satisfying $|Y| \leqq n-t+1$. If $|Y| < n-t+1$, we are done, thus we may suppose $|Y| = n-t+1$. If there is one among the sets A_i which does not contain Y, we are done, again. It means, as the sets are distinct, that their number m is at most 2^{t-1}. By the condition $2^{t-1} \leqq n$ this contradicts $m \geqq n+1$. The proof is complete.

It is easy to see that the family of all $(n-1)$-element subsets of X give equality in (5) if $n+t$ is even. But there are no 4 distinct subsets satisfying (3) if $n=4$, $t=1$.

While the condition of the corollary did not give stronger result than Theorem 1 gave, this is not the case here. Choose $t=1$ and take the stronger conditions $\left|\bigcap_{j=1}^{2} A_{i_j}\right| \neq 0$, $\left|\bigcap_{j=1}^{3} A_{i_j}\right| \leqq 0$. Then the $\binom{m}{2}$ intersections $A_i \cap A_j$ are all disjoint. Consequently, $\binom{m}{2} \leqq n$.

THEOREM 3. *Let $A_1, \ldots, A_m$ be a family of distinct subsets of an n element set X, and let $t>0$ be a fixed integer. Suppose that*

$$\left|\bigcap_{j=1}^{r} A_{i_j}\right| \neq r-1+t \tag{6}$$

holds for any r $(2 \leqq r \leqq m)$ *and any distinct indices* $i_1, \ldots, i_r$ $(1 \leqq i_j \leqq m)$. *Then*

$$m \leqq \sum_{\nu=1}^{t+1} \binom{n}{\nu}. \tag{7}$$

PROOF. Let us count the number of pairs (A_i, G) $(1 \leqq i \leqq m,\ G \subset A_i,\ |G| = t)$ in two different ways

$$\sum_{i=1}^{m} \binom{|A_i|}{t} = \sum_{|G|=t} |\{A_i \colon 1 \leqq i \leqq m,\ G \subset A_i\}|. \tag{8}$$

Here $|\{A_i \colon 1 \leqq i \leqq m,\ G \subset A_i,\ G \neq A_i\}| \leqq n-t$ by Theorem 1. Consequently, the right-hand side of (8) is at most $\binom{n}{t}(n-t+1)$:

$$\sum_{i=1}^{m} \binom{|A_i|}{t} \leqq \binom{n}{t}(n-t+1). \tag{9}$$

Suppose, that in contradiction to (7) $m > \sum_{\nu=1}^{t+1} \binom{n}{\nu}$. It is easy to see that $\sum_{i=1}^{m} \binom{|A_i|}{t} > \sum_{\nu=1}^{t+1} \binom{\nu}{t} \binom{n}{\nu}$ (fewer subsets with smaller sizes), and this contradicts (9). The proof is complete.

If we also assume (6) for $r=1$, then we obtain the bound $\sum_{\nu=1}^{t-1} \binom{n}{\nu} + \binom{n}{t+1}$.

(Received August 6, 1978)

MTA MATEMATIKAI KUTATÓ INTÉZETE
REÁLTANODA U. 13—15
H—1053 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 51—68

DICHTEABSCHÄTZUNGEN FÜR MEHRFACHE GITTERFÖRMIGE KUGELANORDNUNGEN IM $\mathbf{R}^m$

von

U. BOLLE

Das Ziel dieser Arbeit ist die Abschätzung der Dichte m-dimensionaler k-facher Kugelpackungen nach oben, bzw. der Dichte m-dimensionaler k-facher Kugelüberdeckungen nach unten (m stets $\geqq 2$).

0. Dabei werden unter anderem folgende Bezeichnungen benutzt:

$\mathbf{b}$, $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ Vektoren des $\mathbf{R}^m$;

$B(\mathbf{x}) := \{\mathbf{y} | \mathbf{y} \in \mathbf{R}^m \text{ und } \|\mathbf{x}-\mathbf{y}\| \leqq 1\}$;

$K(\mathbf{x}) := \{\mathbf{y} | \mathbf{y} \in \mathbf{R}^m \text{ und } \|\mathbf{x}-\mathbf{y}\| < 1\}$;

$H := \langle \mathbf{b}_1, \ldots, \mathbf{b}_r \rangle$ ein r-dimensionales Teilgitter des $\mathbf{R}^m$ mit der Basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_r\}$;
$\Delta(H)$ das Volumen eines Fundamentalparallelotops eines Gitters H;
μ_m das Lebesgue-Maß auf $\mathbf{R}^m$;

$V_m := \mu_m(B(\mathbf{0})) = \dfrac{\pi^{\frac{m}{2}}}{\Gamma\left(\frac{m}{2}+1\right)}$ das Volumen der m-dimensionalen Einheitskugel;

$c_1, c_2, \ldots, C_1, C_2, \ldots$ positive Konstanten, die nur von m (der Dimension) abhängen.

1. Sei G ein m-dimensionales Gitter des $\mathbf{R}^m$. Dann heißt $\{K(\mathbf{x}) | \mathbf{x} \in G\}$ eine *k-fache gitterförmige Packung der Einheitskugel* genau dann, wenn jedes $\mathbf{y} \in \mathbf{R}^m$ in höchstens k Kugeln $K(\mathbf{x})$ liegt. Für die abgeschlossene Einheitskugel $B(\mathbf{x})$ gilt analog: $\{B(\mathbf{x}) | \mathbf{x} \in G\}$ heißt *k-fache gitterförmige Überdeckung mit Einheitskugeln*, wenn jedes $\mathbf{y} \in \mathbf{R}^m$ in mindestens k Kugeln liegt. Wir werden im folgenden allerdings kurz von „k-Packungen“ bzw. „k-Überdeckungen“ reden, da wir uns stets auf gitterförmige Anordnungen der Einheitskugel beschränken. Die Dichte $d(G)$ einer beliebigen gitterförmigen Lagerung $\{K(\mathbf{x}) | \mathbf{x} \in G\}$ von Einheitskugeln im $\mathbf{R}^m$ ist definiert als

$$d(G) := \lim_{R \to \infty} \frac{1}{(2R)^m} \sum_{\mathbf{x} \in G} \mu_m(K(\mathbf{x}) \cap Q_R),$$

wobei

$$Q := \left\{ \mathbf{y} | \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \text{ und } |y_i| \leqq R \text{ für } 1 \leqq i \leqq m \right\}.$$

Bekanntlich existiert der Grenzwert, und es gilt

$$d(G) = \frac{V_m}{\Delta(G)}.$$

$d(G)$ ändert sich nicht, wenn man $K(\mathbf{x})$ durch $B(\mathbf{x})$ ersetzt, so daß die Dichtedefinition auch für Überdeckungen geeignet ist.

Sei $d_k^{(m)} := \sup \{d(G) | G$ m-dimensionales Gitter, das eine k-Packung liefert$\}$
$D_k^{(m)} := \inf \{d(G) | G$ m-dimensionales Gitter, das eine k-Überdeckung ergibt$\}$.
Aus dem Mahlerschen Kompaktheitslemma folgt sofort, daß es k-Packungen bzw. k-Überdeckungen mit der Dichte $d_k^{(m)}$ bzw. $D_k^{(m)}$ gibt. Der genaue Wert von $d_k^{(m)}$ bzw. $D_k^{(m)}$ ist aber nur für wenige Paare $(m; k)$ bekannt (vgl. [1]—[5]). Es stellt sich also das Problem, Abschätzungen beider Größen zu finden.

Zunächst hat man die trivialen Ungleichungen

$$d_k^{(m)} \geqq k d_1^{(m)}, \quad D_k^{(m)} \leqq k D_1^{(m)}.$$

Ferner gilt

$$\lim_{k \to \infty} \frac{d_k^{(m)}}{k} = \lim_{k \to \infty} \frac{D_k^{(m)}}{k} = 1$$

für alle Dimensionen m. Diese meines Wissens in der Literatur nirgendwo explizit bewiesene Aussage erhält man, wenn man in Gittern der Form $(b\mathbf{Z})^m$, die Anzahl der Gitterpunkte in einer Kugel mit Hilfe des Volumens der zugehörigen Gittermaschen abschätzt. Für weitere Ergebnisse sei auf [1], [2], [3] verwiesen.

Wir wollen in dieser Arbeit zwei Resultate aus [1] verallgemeinern. Dabei ergibt sich folgender

SATZ 1. a) *Es gibt Konstanten c_m, $C_m > 0$, die nur von m abhängen, so daß für $m \equiv 3,7$ (8) gilt:*

$$\frac{d_k^{(m)}}{k} < 1 - c_m k^{-\frac{m+1}{2m}}$$

$$\frac{D_k^{(m)}}{k} > 1 + C_m k^{-\frac{m+1}{2m}};$$

b) *für $m \equiv 1$ (8) gibt es $c'_m > 0$, so daß*

$$\frac{d_k^{(m)}}{k} < 1 - c'_m k^{-\frac{m+3}{2m}};$$

c) *für $m \equiv 5$ (8) gibt es ein $C'_m > 0$, so daß*

$$\frac{D_k^{(m)}}{k} > 1 + C'_m k^{-\frac{m+3}{2m}}.$$

Der Beweis des Satzes ist etwas langwierig und erfordert einige Vorbereitungen.

2. $G = \langle \mathbf{b}_1, \ldots, \mathbf{b}_m \rangle \subseteqq \mathbf{R}^m$ heißt ein Gitter in *Normaldarstellung*, wenn es folgende Eigenschaften hat:

(1) Für $1 \leqq i \leqq m-1$ ist $\langle \mathbf{b}_1, \ldots, \mathbf{b}_i \rangle$ ein i-dimensionales Teilgitter von $\langle \mathbf{b}_1, \ldots, \mathbf{b}_{i+1} \rangle$ und zwar so, daß $\Delta(\langle \mathbf{b}_1, \ldots, \mathbf{b}_i \rangle)$ minimal ist unter allen i-dimensionalen Teilgittern von $\langle \mathbf{b}_1, \ldots, \mathbf{b}_{i+1} \rangle$.

(2) Das Koordinatensystem ist so gewählt, daß für $1 \leqq i \leqq m$

$$\mathbf{b}_i = \begin{pmatrix} b_{i1} \\ \vdots \\ b_{ii} \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{mit} \quad b_{ii} > 0 \quad \text{und} \quad |b_{i+1,i}| \leqq \frac{1}{2} b_{ii}.$$

Lemma 1. *Ist $G = \langle \mathbf{b}_1, \ldots, \mathbf{b}_m \rangle$ ein Gitter in Normaldarstellung, so gilt:*

$$\Delta(G) = b_{11} \ldots b_{mm} \leqq \left(\frac{2}{\sqrt{3}}\right)^{\frac{m(m-1)}{2}} b_{mm}^m.$$

Beweis. Zunächst gilt stets:

$$b_{ii} \leqq \frac{2}{\sqrt{3}} b_{i+1,i+1} \qquad (1 \leqq i \leqq m-1).$$

Es ist klar, daß

$$\Delta(\langle \mathbf{b}_1, \ldots, \mathbf{b}_i \rangle) \leqq \Delta(\langle \mathbf{b}_1, \ldots, \mathbf{b}_{i-1}, \mathbf{b}_{i+1} \rangle),$$

da G sich in Normaldarstellung befinden sollte.

$$\Rightarrow b_{11} \ldots b_{ii} \leqq b_{11} \ldots b_{i-1,i-1} \sqrt{b_{i+1,i}^2 + b_{i+1,i+1}^2}$$

$$\Rightarrow b_{ii}^2 \leqq b_{i+1,i}^2 + b_{i+1,i+1}^2 \leqq \frac{1}{4} b_{i,i}^2 + b_{i+1,i+1}^2$$

$$\Rightarrow b_{ii} \leqq \frac{2}{\sqrt{3}} b_{i+1,i+1}.$$

Die ursprüngliche Behauptung folgt nun leicht durch Induktion: Sei für $i \geqq 1$

$$b_{11} \ldots b_{ii} \leqq \left(\frac{2}{\sqrt{3}}\right)^{\frac{i(i-1)}{2}} b_{ii}^i$$

(was für $i=1$ jedenfalls richtig ist). Dann ergibt sich:

$$b_{11} \ldots b_{i+1,i+1} \leqq \left(\frac{2}{\sqrt{3}}\right)^{\frac{i(i-1)}{2}} b_{ii}^i b_{i+1,i+1} \leqq \left(\frac{2}{\sqrt{3}}\right)^{\frac{i(i-1)}{2}} = \left(\frac{2}{\sqrt{3}}\right)^i b_{i+1,i+1}^{i+1} =$$

$$= \left(\frac{2}{\sqrt{3}}\right)^{\frac{(i+1)i}{2}} b_{i+1,i+1}^{i+1}.$$

Sei wieder $G \subseteqq \mathbf{R}^m$ ein Gitter in Normaldarstellung. Die durch G vermittelte mehrfache *Packung* heißt *vom Typ n* $(\in \mathbf{N})$ genau dann, wenn $\frac{2}{n+1} \leqq b_{mm} < \frac{2}{n}$. Die Definition des Typs läßt sich sinvoll auch noch auf $n=0$ erweitern, wenn man verlangt, daß $2 \leqq b_{mm} < \infty$ sein soll. Analog liegt eine *Überdeckung vom Typ n* vor, wenn $\frac{2}{n+1} < b_{mm} \leqq \frac{2}{n}$. Entsprechend zu definierende Überdeckungen vom Typ 0 gibt es offenbar nicht.

Der Typ ist eindeutig bestimmt, denn

$$b_{mm} = \frac{\Delta(G)}{\Delta(\langle \mathbf{b}_1, \ldots, \mathbf{b}_{m-1} \rangle)},$$

und beide Größen sind geometrische Invarianten des Gitters. Ferner definieren wir:

$$D^{(m)}(n, k) := \inf \{d(G) | G \text{ liefert eine } k\text{-Überdeckung vom Typ } n \text{ in } \mathbf{R}^m\}.$$

Es ist klar, daß nicht für jedes k-Packungen vom Typ n existieren, denn wenn n hinreichend groß ist, liegt schon eine gewisse Anzahl von Gitterpunkten in jeder Einheitskugel. Sei also

$$u(n) := \min \{k | k \in \mathbf{N}, \text{ und es gibt eine } k\text{-Packung vom Typ } n\}$$

und für $k \geqq u(n)$:

$$d^{(m)}(n, k) := \sup \{d(G) | G \text{ liefert eine } k\text{-Packung vom Typ } n\}.$$

Wir werden zunächst Abschätzungen für $d^{(m)}(n, k)$, $D^{(m)}(n, k)$ beweisen, aus denen man dann Aussagen über $d_k^{(m)}$, $D_k^{(m)}$ erhalten kann.

3. Sei im folgenden $G = \langle \mathbf{b}_1, \ldots, \mathbf{b}_m \rangle$ stets ein Gitter in Normaldarstellung im $\mathbf{R}^m$ mit $\frac{2}{n+1} \leqq b_{mm} \leqq \frac{2}{n}$. Wir können uns die Punkte von G in parallelen Hyperebenen E_i $(i \in \mathbf{Z})$ angeordnet denken, wobei $E_i := \mathbf{R}^{m-1} \times \{ib_{mm}\}$ ist. Wir betrachten nun die Schnitte einer beliebigen Kugel $K(\mathbf{x})$ (oder auch $B(\mathbf{x})$) mit diesen Hyperebenen. Das Maß eines solchen Schnittes hängt offenbar nur von der m-ten Koordinate y von $\mathbf{x}$ ab:

$$S_i(y, h) := \mu_{m-1}(E_i \cap K(\mathbf{x})) = \mu_{m-1}(E_i \cap B(\mathbf{x})) =$$

$$= \begin{cases} V_{m-1}\left(1-(ih-y)^2\right)^{\frac{m-1}{2}}, & \text{falls } |ih-y| \leqq 1 \\ 0 & \text{sonst.} \end{cases}$$

Dabei haben wir, wie von jetzt an immer, h statt b_{mm} geschrieben, um Indizes zu sparen. Sei weiter

$$S(y, h) := \sum_{i \in \mathbf{Z}} S_i(y, h).$$

S is als Funktion von y periodisch mit der Periode h; außerdem gilt

$$S\left(\frac{h}{2} - y, h\right) = S\left(\frac{h}{2} + y, h\right),$$

so daß bei festem h alle Werte von S bereits auf dem Intervall $\left[0; \frac{h}{2}\right]$ angenommen werden.

Über S gilt nun

LEMMA 2. *Sei* $G=\langle \mathbf{b}_1, \dots, \mathbf{b}_m\rangle$ *ein Gitter in Normaldarstellung,*

$$\frac{2}{n+1} \leqq h \leqq \frac{2}{n}$$

und für beliebiges $\mathbf{x}\in\mathbf{R}^m$

$$v(\mathbf{x}) := \operatorname{card}(K(\mathbf{x})\cap G)$$
$$w(\mathbf{x}) := \operatorname{card}(B(\mathbf{x})\cap G).$$

Ferner sei $P\subseteqq \mathbf{R}^{m-1}\times\{0\}$ *ein Fundamentalparallelotop von* $\langle \mathbf{b}_1, \dots, \mathbf{b}_{m-1}\rangle$. *Dann gilt:*

$$\int_P v(\mathbf{z}, y)\, d\mathbf{z} = \int_P w(\mathbf{z}, y)\, d\mathbf{z} = S(y, h)$$

für jedes $y\in\mathbf{R}$.

BEWEIS. Wir bezeichnen mit $\chi_{i_1, i_2, \dots, i_m}$ die charakteristische Funktion von $K(i_1\mathbf{b}_1+\dots+i_m\mathbf{b}_m)$, also

$$\chi_{i_1, \dots, i_m}(\mathbf{x}) = \begin{cases} 1, & \text{falls } \|\mathbf{x}-i_1\mathbf{b}_1-\dots-i_m\mathbf{b}_m\| < 1 \\ 0 & \text{sonst.}\end{cases}$$

Dann gilt offenbar

$$v(\mathbf{z}, y) = \sum_{i_1, \dots, i_m\in Z} \chi_{i_1, \dots, i_m}(\mathbf{z}, y)$$
$$(\mathbf{z}\in\mathbf{R}^{m-1},\ y\in\mathbf{R}),$$

so daß

$$\int_P v(\mathbf{z}, y)\, d\mathbf{z} = \int_P \sum_{i_1, \dots, i_m} \chi_{i_1, \dots, i_m}(\mathbf{z}, y)\, d\mathbf{z}.$$

Nun ist

$$\int_P \chi_{i_1, \dots, i_m}(\mathbf{z}, y)\, d\mathbf{z} = \mu_{m-1}\big(K(i_1\mathbf{b}_1+\dots+i_m\mathbf{b}_m)\cap P\times\{y\}\big) =$$
$$= \mu_{m-1}\big(K(i_m\mathbf{b}_m)\cap(P\times\{y\}-i_1\mathbf{b}_1-\dots-i_{m-1}\mathbf{b}_{m-1})\big),$$

wenn die Minuszeichen die Minkowski-Subtraktion bezeichnen, und insgesamt

$$\sum_{i_1, \dots, i_{m-1}} \int_P \chi_{i_1, \dots, i_m}(\mathbf{z}, y)\, d\mathbf{z} =$$
$$= \sum_{i_1, \dots, i_{m-1}} \mu_{m-1}\big(K(i_m\mathbf{b}_m)\cap(P\times\{y\}-i_1\mathbf{b}_1-\dots-i_{m-1}\mathbf{b}_{m-1})\big) =$$
$$= \mu_{m-1}\big(K(i_m\mathbf{b}_m)\cap\mathbf{R}^{m-1}\times\{y\}\big) =$$
$$= \mu_{m-1}\big(K(\mathbf{0}, y)\cap\mathbf{R}^{m-1}\times\{i_m h\}\big) = S_{i_m}(y, h),$$

also

$$\int_P v(\mathbf{z}, y)\, d\mathbf{z} = \sum_{i_m\in Z} S_{i_m}(y, h) = S(y, h).$$

Die Aussage für w folgt ebenso.

KOROLLAR. a) *G liefert eine k-Packung* $\Rightarrow S(y, h) \leqq k\mu_{m-1}(P) = k\dfrac{\Delta(G)}{h}$ *für alle* $y \in \mathbf{R}$.

b) *G führt zu einer k-Überdeckung* $\Rightarrow S(y, h) \geqq k\dfrac{\Delta(G)}{h}$ *für alle* $y \in \mathbf{R}$.

Aus dem Korollar ergibt sich leicht der für die weitere Rechnung grundlegende

SATZ 2. *Sei für* $n \in \mathbf{N}$

$$\left.\begin{aligned} a_n^{(m)} &:= \min_h \max_y hS(y, h) \\ A_n^{(m)} &:= \max_h \min_y hS(y, h) \end{aligned}\right\} \quad \begin{aligned} \frac{2}{n+1} \leqq h \leqq \frac{2}{n} \\ 0 \leqq y \leqq \frac{h}{2} \end{aligned}$$

Dann gilt:

$$\frac{d^{(m)}(n, k)}{k} \leqq \frac{V_m}{a_n^{(m)}} \quad \textit{für alle} \quad k \geqq u(n),$$

$$\frac{D^{(m)}(n, k)}{k} \geqq \frac{V_m}{A_n^{(m)}} \quad \textit{für alle} \quad k \in \mathbf{N}.$$

BEWEIS z. B. für Packungen: Sei G ein beliebiges Gitter des $\mathbf{R}^m$ in Normaldarstellung, das eine k-Packung vom Typ n liefert. Dann folgt aus dem Korollar:

$$hS(y, h) \leqq k\Delta(G)$$

für alle $y \in \mathbf{R}$, also

$$\max_y hS(y, h) \leqq k\Delta(G)$$

und damit erst recht

$$\min_h \max_y hS(y, h) = a_n^{(m)} \leqq k\Delta(G),$$

so daß

$$\frac{d(G)}{k} = \frac{V_m}{k\Delta(G)} \leqq \frac{V_m}{a_n^{(m)}}$$

für jede k-Packung vom Typ n. Der Beweis für k-Überdeckungen verläuft wieder ganz entsprechend.

Daß diese Abschätzungen nicht von vornherein trivial sind, zeigt

LEMMA 3. *Sei* $n \in \mathbf{N}$, $m \geqq 2$. *Falls $S(y, h)$ als Funktion von y für kein h konstant ist, gilt*

$$a_n^{(m)} > V_m \quad \textit{und} \quad A_n^{(m)} < V_m.$$

BEWEIS. Wir beschränken uns auf $n \equiv 0\ (2)$, der andere Fall läßt sich ganz analog behandeln. Sei also $n = 2N$, $h = \dfrac{2}{n+2x} = \dfrac{1}{N+x}$ mit $x \in \left[0; \dfrac{1}{2}\right]$. Dann haben wir

$$S(y, h) = V_{m-1} \sum_{|ih-y| \leqq 1} \left(1-(ih-y)^2\right)^{\frac{m-1}{2}} =$$

$$= V_{m-1} \sum_{i=-N+1}^{N} \left(1-(ih-y)^2\right)^{\frac{m-1}{2}} + V_{m-1} r(y, h),$$

wobei

$$r(y, h) := \begin{cases} \left(1-(Nh+y)^2\right)^{\frac{m-1}{2}}, & \text{für} \quad y < hx \\ 0 & \text{für} \quad y \geqq hx \end{cases}$$

und

$$\frac{2}{h} \int_0^{\frac{h}{2}} hS(y, h)\, dy = 2V_{m-1} \sum_{i=-N+1}^{N} \int_0^{\frac{h}{2}} \left(1-(ih-y)^2\right)^{\frac{m-1}{2}} dy + 2V_{m-1} \int_0^{\frac{h}{2}} r(y, h)\, dy =$$

$$= 2V_{m-1} \left(\sum_{i=1}^{N} \int_{\frac{2i-1}{2}}^{ih} (1-z^2)^{\frac{m-1}{2}} dz + \sum_{i=0}^{N-1} \int_{ih}^{\frac{2i+1}{2}} (1-z^2)^{\frac{m-1}{2}} dz + \int_{Nh}^{(N+x)h} (1-z^2)^{\frac{m-1}{2}} dz \right) =$$

$$= 2V_{m-1} \int_0^1 (1-z^2)^{\frac{m-1}{2}} dz = V_m.$$

Sei nun $h \in \left[\frac{2}{n+1}; \frac{2}{n}\right]$ beliebig. Wenn $S(y, h)$ als Funktion von y nicht konstant ist, so gilt $\min\limits_y hS(y, h) < V_m$, $\max\limits_y hS(y, h) > V_m$. Ferner sind $\min\limits_y hS(y, h)$ und $\max\limits_y hS(y, h)$ als Funktionen von h stetig. Daher ist $a_n^{(m)} > V_m$ bzw. $A_n^{(m)} < V_m$, wie behauptet.

Bemerkung. Daß die Voraussetzung des Lemmas wenigstens in dem für das weitere benötigten Fall ungeraden m's erfüllt ist, ergibt sich einfach daraus, daß $S(y, h)$ sich auf $\left[0; \frac{h}{2}\right]$ aus zwei Polynomen vom Grade $m-1$ in y zusammengesetzt, deren höchster Koeffizient von h nicht abhängt.

4. Es bleibt also noch, $S(y, h)$ genauer zu berechnen. Dazu benötigen wir eine Reihe von Sätzen aus der Analysis die der Bequemlichkeit halber im diesen Abschnitt zusammengestellt werden sollen.

1. Sei f: $[a; b] \to \mathbf{R}$ stetig und $\lim\limits_{x \to a_+} f(x) = \lim\limits_{x \to b_-} f(x) = 0$; ferner konvergiere

$$\sum_{r=-\infty}^{+\infty} \int_a^b f(x) \cos(2\pi r x)\, dx.$$

Dann gilt

$$\sum_{\substack{a<j<b \\ j \text{ ganz}}} f(j) = \sum_{r=-\infty}^{+\infty} \int_a^b f(x) \cos(2\pi r x)\, dx$$

([6], S. 113).

2. Für $M \in \mathbf{N}_0 = \mathbf{N} \cup \{0\}$ und $z > 0$ gilt:

$$\int_0^1 (1-t^2)^M \cos(zt)\, dt \frac{2^M M!}{z^{M+1}} \cos\left(z - \frac{M+1}{2}\pi\right) -$$

$$-\frac{M 2^{M-2}(M+1)!}{z^{M+2}} \sin\left(z - \frac{M+1}{2}\pi\right) + O\left(\frac{1}{z^{M+3}}\right)$$

([8], S. 366, 368).

3. Für $\operatorname{Re}(s) < 0$ und $0 < a \leqq 1$ gilt:

$$\sum_{j=1}^{\infty} \frac{\sin\left(2\pi j a + \frac{\pi s}{2}\right)}{j^{1-s}} = \frac{(2\pi)^{1-s}}{2\Gamma(1-s)} \zeta(s, a),$$

wobei $\zeta(s, a)$ für die verallgemeinerte ζ-Funktion steht ([8], S. 268).

4. Für $M \in \mathbf{N}_0$ und $0 < a \leqq 1$ hat man:

$$\zeta(-M, a) = -\frac{B_{M+1}(a)}{M+1}$$

wo B_{M+1} das Bernoulli-Polynom $(M+1)$-ten Grades bezeichnet.

Wir benutzen die Definition der Bernoulli-Polynome nach [7]. Die angegebene Aussage findet man etwa in [8], S. 267, allerdings ist dabei zu beachten, daß die dort als „Bernoulli-Polynome" bezeichneten Φ_n mit unserem B_n durch die Gleichung

$$\Phi_n(x) = B_n(x) - B_n(0)$$

$(n \geqq 1)$ zusammenhängen.

5. Die Bernoulli-Polynome haben folgende Eigenschaften:

5.1 $B'_{n+1}(x) = (n+1) B_n(x)$;

5.2 $B_n\left(\frac{1}{2} - x\right) = (-1)^n B_n\left(\frac{1}{2} + x\right)$;

5.3 Für $n \geqq 3$, $n \equiv 1\,(2)$:

$$B_n(0) = B_n\left(\frac{1}{2}\right) = B_n(1) = 0,$$

und B_n besitzt auf $[0; 1]$ keine weiteren Nullstellen, ferner

$$\operatorname{sgn} B_n\left(\frac{1}{4}\right) = (-1)^{\frac{n+1}{2}};$$

5.4 Für $n > 2$, $n \equiv 0\,(2)$:

$$B_n(0) = B_n(1) = -\frac{2^{n-1}}{2^{n-1}-1} B_n\left(\frac{1}{2}\right) = b_n$$

(b_n steht für n-te Bernoulli-Zahl); B_n besitzt genau eine Nullstelle x_n in $\left[0; \frac{1}{2}\right]$;

$\operatorname{sgn} B_n(0) = (-1)^{\frac{n}{2}+1}$;

B_n besitzt relative Extrema bei 0, $\frac{1}{2}$, 1 und keine weiteren auf [0, 1].

5.5 $$B_{2r}\left(\frac{1}{4}\right) = -\frac{1}{4^r}\left(1-\frac{1}{2^{2r-1}}\right) B_{2r}(0) \sim \sqrt{8\pi r}\left(\frac{r}{2\pi e}\right)^{2r}.$$

Für die Aussagen 5.1—5.5 sei auf [7] und [8] verwiesen. Weiter benötigen wir noch eine Eigenschaft der nach 5.4 eindeutig bestimmten Nullstelle x_{2r} von B_{2r} auf $\left[0, \frac{1}{2}\right]$, die man durch explizite Abschätzung aus [5] erhält.

5.6 *Für $r \geqq 1$ gilt*

$$x_{2r} < x_{2r+2}.$$

BEWEIS. Setzen wir mit LEHMER [7]

$$\vartheta := 2\pi\left(\frac{1}{4} - x_{2r}\right).$$

Dann entnimmt man aus der genannten Arbeit, daß

$$\sin\vartheta = 2^{-2r}\cos 2\vartheta + 3^{-2r}\sin 3\vartheta - 4^{-2r}\cos 4\vartheta - 5^{-2r}\sin 5\vartheta + \ldots$$

und $0 < \vartheta < 2^{-2r}$ ist.

Wir wollen zunächst nachrechnen, daß

$$|\sin\vartheta - 2^{-2r}| < 3\cdot 4^{-2r}$$

ist für $r \geqq 3$:

$$|\sin\vartheta - 2^{-2r}\cos 2\vartheta| < 3\cdot 6^{-2r} + 4^{-2r} + \int_4^\infty x^{-2r}\,dx =$$

$$= 4^{-2r}\left[3\left(\frac{3}{2}\right)^{-2r} + 1 + \frac{4}{2r-1}\right] < 2{,}5\cdot 4^{-2r} \quad \text{für} \quad r \geqq 3.$$

Weiter ist

$$|\cos 2\vartheta - 1| \leqq 2\vartheta^2(1 + (2\vartheta)^2 + (2\vartheta)^4 + \ldots) <$$

$$< 2\cdot 4^{-2r}\frac{1}{1-4^{-2r+1}} < 2{,}5\cdot 4^{-2r} \quad \text{für} \quad r \geqq 3$$

und damit

$$|\sin\vartheta-2^{-2r}| \leqq 2{,}5\cdot 4^{-2r}+2{,}5\cdot 8^{-2r} < 3\cdot 4^{-2r}.$$

Schließlich erhält man mit

$$|\sin\vartheta-\vartheta| \leqq \frac{\vartheta^3}{6}(1+\vartheta^2+\vartheta^4+\ldots) = \frac{\vartheta^3}{6}\frac{1}{1-\vartheta^2} < 0{,}2\cdot 8^{-2r}$$

$$|\vartheta-2^{-2r}| < 3\cdot 4^{-2r}+0{,}2\cdot 8^{-2r} < 4^{-2r+1} \quad \text{für} \quad r \geqq 3.$$

Das ergibt

$$x_{2r} = \frac{1}{4}-\frac{2^{-2r}}{2\pi}+c_r 4^{-2r},$$

wobei $|c_r| \leqq 1$ für $r \geqq 3$, und

$$x_{2r+2}-x_{2r} = \frac{3}{8\pi}2^{-2r-2}+4^{-2r-2}(c_{r+1}-16c_r) \geqq$$

$$\geqq 2^{-2r-2}\left(\frac{3}{8\pi}-17\cdot 2^{-2r-2}\right) > 0 \quad \text{für} \quad r \geqq 3.$$

$x_2 < x_4 < x_6$ kann man der Tabelle am Schluß von [7] entnehmen.

5. Um

$$S(y,h) = V_{m-1}\sum_{|ih-y|\leqq 1}\left(1-(ih-y)^2\right)^{\frac{m-1}{2}}$$

weiterzubehandeln, beschränken wir uns von nun an auf den Fall $m=2M+1$ $(M\in\mathbf{N})$. Wir benutzen zunächst 4.1 mit $a=-\frac{1-y}{h}$, $b=\frac{1+y}{h}$ und $f(x):=\left(1-(xh-y)^2\right)^M$. Dann ergibt sich

$$S(y,h) = V_{m-1}\sum_{r=-\infty}^{\infty}\int_a^b (1-(zh-y)^2)^M\cos(2\pi rz)\,dz =$$

$$= \frac{V_{m-1}}{h}\sum_{r=-\infty}^{+\infty}\cos\frac{2\pi ry}{h}\int_{-1}^{+1}(1-u^2)^M\cos\frac{2\pi ru}{h}\,du,$$

da die Reihe nach 4.2 konvergiert. Nach leichter Umformung erhält man

$$hS(y,h) = V_m+4V_{m-1}\sum_{r=1}^{\infty}\cos\frac{2\pi ry}{h}\int_0^1(1-u^2)^M\cos\frac{2\pi ru}{h}\,du$$

und weiter unter Verwendung von 4.2, 4.3, und 4.4:

$$\begin{aligned}
hS(y,h) = {} & V_m+\frac{2V_{m-1}M!}{\pi^{M+1}}h^{M+1}\sum_{r=1}^{\infty}\frac{1}{r^{M+1}}\cos\frac{2\pi ry}{h}\times \\
& \times\cos\left(\frac{2\pi r}{h}-\frac{M+1}{2}\pi\right)-\frac{M(M+1)!}{2\pi^{M+2}}V_{m-1}h^{M+2}\sum_{r=1}\frac{1}{r^{M+2}}\times \\
& \times\cos\frac{2\pi ry}{h}\cdot\sin\left(\frac{2\pi r}{h}-\frac{M+1}{2}\pi\right)+O(h^{M+3}) = \\
& = V_m+2^M V_{m-1}h^{M+1}(\zeta(-M,a_1)+\zeta(-M,a_2))- \\
& -2^{M-1}MV_{m-1}h^{M+2}(\zeta(-M-1,a_1)+\zeta(-M-1,a_2))+O(h^{M+3}) = \\
& = V_m-\frac{2^M V_{m-1}}{M+1}h^{M+1}(B_{M+1}(a_1)+B_{m+1}(a_2))+ \\
& +\frac{2^{M-1}MV_{m-1}}{M+2}h^{M+2}(B_{M+2}(a_1)+B_{M+2}(a_2))+O(h^{M+3}),
\end{aligned}$$

wenn wir $a_1, a_2\in]0;1]$ so wählen, daß

$$\frac{1+y}{h}\equiv a_1 \bmod 1 \quad \text{und} \quad \frac{1-y}{h}\equiv a_2 \bmod 1 \quad \text{ist.}$$

Der Kürze halber schreiben wir $T_i(y,h)$ statt $B_i(a_1)+B_i(a_2)$ $(i=M+1, M+2)$. Nun gilt folgendes

LEMMA 4. *Sei $m=2M+1$, $M\in\mathbf{N}$, $n\in\mathbf{N}$. Dann haben wir*

$$\left.\begin{array}{l}
\text{(1) } \textit{für } M\equiv 1,3\ (4) \\
\max\limits_h \min\limits_y T_{M+1}(y,h)\leqq -c_1<0 \\
\min\limits_h \max\limits_y T_{M+1}(y,h)\geqq c_2>0; \\
\text{(2) } \textit{für } M\equiv 0\ (4) \textit{ und alle hinreichend kleinen } h \\
\max\limits_h \min\limits_y \Big(T_{M+1}(y,h)- \\
\quad -\dfrac{1}{2}\dfrac{M(M+1)}{M+2}hT_{M+2}(y,h)\Big)\leqq -c_3h; \\
\text{(3) } \textit{für } M\equiv 2\ (4) \textit{ und alle hinreichend kleinen } h \\
\min\limits_h \max\limits_y \Big(T_{M+1}(y,h)- \\
\quad -\dfrac{1}{2}\dfrac{M(M+1)}{M+2}hT_{M+2}(y,h)\Big)\geqq c_4h
\end{array}\right\}
\begin{array}{l}
0\leqq y\leqq \dfrac{h}{2} \\[2ex]
\dfrac{2}{n+1}\leqq h\leqq \dfrac{2}{n}.
\end{array}$$

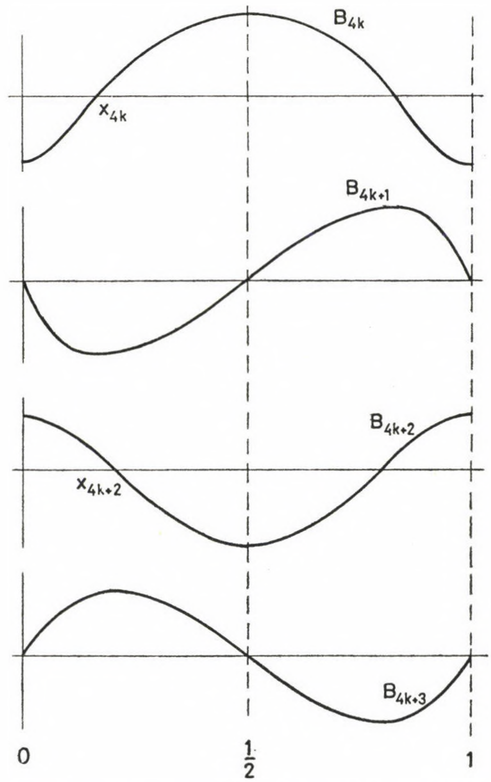
Die c_i sind dabei positive Zahlen, die wohl von m, nicht aber von n abhängen.

BEWEIS. *Zu* (1). Sei $M\equiv 3(4)$, also $M+1\equiv 0(4)$. Das zugehörige Bernoulli-Polynom findet man in der Abbildung schematisch dargestellt. Für die Nullstellen gilt nach 5.2, 5.4

$$x_{M+1}<\frac{1}{4},\quad 1-x_{M+1}>\frac{3}{4},$$

und sogar

$$\left|B_{M+1}\left(\frac{1}{4}\right)\right|\sim\sqrt{4\pi M}\left(\frac{M+1}{4e}\right)^{M+1}\to\infty \quad \text{für} \quad M\to\infty.$$



Graphen der Bernoulli-Polynome (schematisch)

Jedes $h\in\left[\frac{2}{n+1};\,\frac{2}{n}\right]$ läßt sich in der Form $h=\frac{2}{n+2x}$ mit einem geeigneten $x\in\left[0,\,\frac{1}{2}\right]$ schreiben.

1. Betrachten wir zunächst den Fall geraden n's: $n=2N$ ($N\in\mathbf{N}$). Dann erhal-

ten wir

$$\frac{1+y}{h} = N+x+\frac{y}{h}, \quad \text{also} \quad a_1 = \begin{cases} 1 & \text{falls } x = y = 0 \\ x+\frac{y}{h} & \text{sonst} \end{cases}$$

$$\left(0 \leqq x \leqq \frac{1}{2} \text{ und } 0 \leqq y \leqq \frac{h}{2}\right).$$

$$\frac{1-y}{h} = N+x-\frac{y}{h} \quad \text{und daher} \quad a_2 = \begin{cases} x-\frac{y}{h} & \text{falls } \quad x > \frac{y}{h} \\ x-\frac{y}{h}+1 & \text{sonst.} \end{cases}$$

Wählen wir für beliebiges $h \in \left[\frac{2}{n+1}; \frac{2}{n}\right]$ $y_1 := xh \left(\leqq \frac{1}{2}h\right)$, so erhalten wir

$$\max_h \min_y T_{M+1}(y, h) \leqq \max_h T_{M+1}(y_1, h) \leqq$$

$$\leqq B_{M+1}(0) + B_{M+1}\left(\frac{1}{2}\right) = \frac{1}{2^M} B_{M+1}(0) =: -c_1 < 0.$$

Um $\min_h \max_y T_{M+1}(y, h)$ abzuschätzen, wählen wir

$$y_2 := \begin{cases} \frac{h}{2} & \text{für } \quad 0 \leqq x \leqq \frac{1}{4} \\ 0 & \text{für } \quad \frac{1}{4} < x \leqq \frac{1}{2}. \end{cases}$$

Dann gilt offenbar

$$\min_h \max_y T_{M+1}(y, h) \geqq \min_h T_{M+1}(y_2, h) \geqq$$

$$\geqq 2B_{M+1}\left(\frac{1}{4}\right) =: c_2 > 0 \ (5.4, 5.5).$$

2. Sei jetzt $n = 2N+1$ ($N \in \mathbf{N}_0$). Dann haben wir

$$\frac{1+y}{h} = N+x+\frac{1}{2}+\frac{y}{h}, \quad \frac{1-y}{h} = N+x+\frac{1}{2}-\frac{y}{h}$$

und damit

$$a_1 = \begin{cases} x+\frac{1}{2}-\frac{y}{h} & \text{für } \quad x+\frac{y}{h} \leqq \frac{1}{2} \\ x+\frac{y}{h}-\frac{1}{2} & \text{sonst;} \end{cases}$$

$$a_2 = \begin{cases} 1, & \text{falls } \quad x = 0, \ y = \frac{h}{2} \\ x+\frac{1}{2}-\frac{y}{h} & \text{sonst.} \end{cases}$$

Wählen wir $y_1 := \left(\frac{1}{2} - x\right) h$, so erhalten wir

$$\max_h \min_y T_{M+1}(y, h) \leqq B_{M+1}(1) + B_{M+1}\left(\frac{1}{2}\right) = -c_1 < 0.$$

Für y_2 wählen wir

$$y_2 := \begin{cases} 0 & \text{für} \quad 0 \leqq x \leqq \frac{1}{4} \\ \frac{h}{2} & \text{sonst.} \end{cases}$$

Dann wird

$$\min_h \max_y T_{M+1}(y, h) \geqq \min_h T_{M+1}(y_2, h) \geqq 2B_{M+1}\left(\frac{1}{4}\right) =: c_2 > 0.$$

Damit ist der Fall $M = 3(4)$ erledigt. Die Abschätzung für $M \equiv 1(4)$ verläuft ganz analog. Wir beschränken uns daher darauf, die y_i und das Ergebnis anzugeben.

3. $n = 2N$.

$$y_1 := \begin{cases} \frac{h}{2} & \text{für} \quad 0 \leqq x \leqq \frac{1}{4} \\ 0 & \text{sonst.} \end{cases}$$

$$\max_h \min_y T_{M+1}(y, h) \leqq 2B_{M+1}\left(\frac{1}{4}\right) < 0,$$

$$y_2 := xh,$$

$$\min_h \max_y T_{M+1}(y, h) \geqq B_{M+1}(0) + B_{M+1}\left(\frac{1}{2}\right) > 0.$$

4. $n = 2N + 1$.

$$y_1 := \begin{cases} 0 & \text{für} \quad 0 \leqq x \leqq \frac{1}{4} \\ \frac{1}{2} h & \text{sonst.} \end{cases}$$

$$\max_h \min_y T_{M+1}(y, h) \leqq 2B_{M+1}\left(\frac{1}{4}\right) < 0,$$

$$y_2 := \left(\frac{1}{2} - x\right) h,$$

$$\min_h \max_y T_{M+1}(y, h) \geqq B_{M+1}(0) + B_{M+1}\left(\frac{1}{2}\right) > 0.$$

Zu (2). $M \equiv 0(4)$.

(5.6) Sei $$\delta := \frac{1}{4}(x_{M+2} - x_M) > 0.$$

1. $n=2N$. Wählen wir

$$y_1 := \begin{cases} 0 & \text{für} \quad 0 \leqq x < \frac{1}{2}-\delta \\ (x-x_M)h & \text{für} \quad \frac{1}{2}-\delta \leqq x \leqq \frac{1}{2}, \end{cases}$$

so erhalten wir

$$\text{für} \quad 0 \leqq x \leqq x_M\colon \; T_{M+1}(y_1, h) \leqq 0$$
$$T_{M+2}(y_1, h) \geqq 2B_{M+2}(x_M) > 0,$$

$$\text{für} \quad x_M < x < \frac{1}{2}-\delta\colon \; T_{M+1}(y_1, h) < 2B_{M+1}\left(\frac{1}{2}-\delta\right) < 0,$$

$$\text{für} \quad \frac{1}{2}-\delta \leqq x \leqq \frac{1}{2}\colon \; T_{M+1}(y_1, h) = B_{M+1}(a_1)+B_{M+1}(x_M) \leqq 0$$
$$T_{M+2}(y_1, h) = B_{M+2}(a_1)+B_{M+2}(x_M) > B_{M+2}(x_M),$$

weil

$$a_1 = 2x-x_M \geqq 1-2\delta-x_M =$$
$$= 1-\frac{1}{2}(x_M+x_{M+2}) > 1-x_{M+2}.$$

Insgesamt ergibt sich

$$\max_h \min_y \left(T_{M+1}(y, h) - \frac{1}{2}\frac{M(M+1)}{M+2} hT_{M+2}(y, h)\right) \leqq -c_3 h.$$

2. $n=2N+1$. Wählen wir

$$y_1 := \begin{cases} \frac{h}{2} & \text{für} \quad 0 \leqq x < \frac{1}{2}-\delta \\ \left(\left(\frac{1}{2}-x+x_M\right)\right) h & \text{für} \quad \frac{1}{2}-\delta \leqq x \leqq \frac{1}{2}, \end{cases}$$

so bekommen wir ganz entsprechend wie oben

$$\text{für} \quad 0 \leqq x \leqq x_M\colon \; T_{M+1}(y_1, h) = 2B_{M+1}(x) \leqq 0$$
$$T_{M+2}(y_1, h) \geqq 2B_{M+2}(x_M) > 0,$$

$$\text{für} \quad x_M < x < \frac{1}{2}-\delta\colon \; T_{M+1}(y_1, h) < 2B_{M+1}\left(\frac{1}{2}-\delta\right) < 0$$

$$\text{für} \quad \frac{1}{2}-\delta \leqq x \leqq 1\colon \; T_{M+1}(y_1, h) = B_{M+1}(x_M)+B_{M+1}(a_2) \leqq 0$$
$$T_{M+2}(y_1, h) = B_{M+2}(x_M)+B_{M+2}(a_2) > 0,$$

$$\text{weil} \quad x_{M+2} > x_M \quad \text{und} \quad a_2 = 2x-x_M > 1-x_{M+2}.$$

Zusammengefaßt ergibt sich wieder die Behauptung (2).

Zu (3). Der Beweis ist fast wörtlich der gleiche wie zu (2); Wir geben daher nur die jeweils zu wählenden y_i an.

$$\delta := \frac{1}{4}(x_{M+2}-x_M).$$

1. $n=2N$:

$$y_1 := \begin{cases} 0 & \text{für } 0 \leqq x < \frac{1}{2}-\delta \\ (x-x_M)h & \text{für } \frac{1}{2}-\delta \leqq x \leqq \frac{1}{2}. \end{cases}$$

2. $n=2N+1$:

$$y_2 := \begin{cases} \frac{h}{2} & \text{für } 0 \leqq x < \frac{1}{2}-\delta \\ \left(\frac{1}{2}+x_M-x\right)h & \text{für } \frac{1}{2}-\delta \leqq x \leqq \frac{1}{2}. \end{cases}$$

Aus Lemma 3 ergibt sich ohne Schwierigkeiten für die Abschätzung von $d^{(m)}(n,k)$, $D^{(m)}(n,k)$

SATZ 3. *Sei $m \in \mathbf{N}$ und $m \geqq 3$. Dann gelten*

(1) *für* $m \equiv 3, 7 (8)$:

$$\frac{d^{(m)}(n,k)}{k} \leqq 1-\frac{c'_m}{n^{\frac{m+1}{2}}}, \quad \frac{D^{(m)}(n,k)}{k} \geqq 1+\frac{C'_m}{n^{\frac{m+1}{2}}};$$

(2) *für* $m \equiv 1 (8)$:

$$\frac{d^{(m)}(n,k)}{k} \leqq 1-\frac{c'_m}{n^{\frac{m+3}{2}}};$$

(3) *für* $m \equiv 5 (8)$:

$$\frac{D^{(m)}(n,k)}{k} \geqq 1+\frac{C'_m}{n^{\frac{m+3}{2}}}.$$

Dabei sind c'_m, C'_m positive Zahlen, die nicht von n abhängen.

BEWEIS. Wir beweisen nur die Aussage (1) für Packungen; alles übrige folgt dann ganz entsprechend. Sei also $m \equiv 3 (8)$ oder $m \equiv 7 (8)$. Dann liefert Lemma 4

$$a_n^{(m)} = \min_h \max_y hS(y,h) = V_m - \frac{2^M V_{m-1}}{M+1} \times$$

$$\times \max_h \min_y h^{M+1} T_{M+1}(y,h) + O\left(\frac{1}{n^{M+2}}\right) \geqq$$

$$\geqq V_m + \frac{c''_1}{n^{M+1}} + O\left(\frac{1}{n^{M+2}}\right) \geqq V_m + \frac{c'''_1}{n^{\frac{m+1}{2}}},$$

da $\frac{2}{n+1} \leqq h < \frac{2}{n}$ ist. Mit Satz 2 haben wir

$$\frac{d^{(m)}(n, k)}{k} \leqq \frac{V_m}{a_n^{(m)}} \leqq 1 - \frac{c'_m}{n^{\frac{m+1}{2}}}, \quad \text{wie behauptet.}$$

Daß wir dabei $c'_m > 0$ wählen können, ergibt sich für alle hinreichend großen n aus Lemma 4, für „kleine" n aus Lemma 3.

6. Es bleibt schließlich noch Satz 1 zu beweisen, wobei wir uns wieder, um Wiederholungen zu vermeiden, auf die Fälle $m=3$, $7(8)$ beschränken. Sei zunächst G ein Gitter in Normaldarstellung, das eine optimale k-Packung im $\mathbf{R}^m$ liefert, und bezeichnen wir den Typ der Packung mit n_k, d. h.

$$\frac{2}{n_k+1} \leqq h = b_{mm} < \frac{2}{n_k}.$$

Nach Lemma 1 gilt

$$\Delta(G) \leqq \left(\frac{2}{\sqrt{3}}\right)^{\frac{m(m-1)}{2}} h^m < \left(\frac{2}{\sqrt{3}}\right)^{\frac{m(m-1)}{2}} \left(\frac{2}{n_k}\right)^m.$$

Andererseits ist sicher

$$k \geqq d_k^{(m)} = \frac{V_m}{\Delta(G)} > \left(\frac{\sqrt{3}}{2}\right)^{\frac{m(m-1)}{2}} \left(\frac{n_k}{2}\right)^m V_m,$$

also $n_k \leqq \alpha_m k^{\frac{1}{m}}$ mit einem nur von m abhängigen α_m. Dann ergibt Satz 3:

$$\frac{d_k^{(m)}}{k} = \frac{d^{(m)}(n_k, k)}{k} \leqq 1 - \frac{c'_m}{n_k^{\frac{m+1}{2}}} \leqq 1 - \frac{c_m}{k^{\frac{m+1}{2m}}}.$$

Um den Beweis für k-Überdeckungen zu führen, haben wir nur die Abschätzungen für n_k unwesentlich zu modifizieren, alles übrige bleibt sinngemäß gleich.

Zu einer Schranke für n_k kommt man auf folgende Weise:

$$kD_1^{(m)} \geqq D_k^{(m)} = D^{(m)}(n_k, k) = \frac{V_m}{\Delta(G)} > \left(\frac{\sqrt{3}}{2}\right)^{\frac{m(m-1)}{2}} \left(\frac{n_k}{2}\right)^m V_m,$$

woraus sich $n_k \leqq \beta_m k^{\frac{1}{m}}$ ergibt mit einem nur von m abhängigen β_m, das allerdings noch das auch nicht genau bekannte $D_1^{(m)}$ enthält.

LITERATURVERZEICHNIS

[1] BOLLE, U.: Mehrfache Kreisanordnungen in der euklidischen Ebene, Dissertation, Dortmund, 1976.

[2] FEJES TÓTH, G.: Multiple packing and covering of the plane with circles, *Acta Math. Acad. Sci. Hungar.* **27** (1976), 135—140.

[3] FEJES TÓTH, G.: Multiple packing and covering of spheres, *Acta Math. Acad. Sci. Hungar.* **34** (1979), 165—176.

[4] FEW, L.: Multiple packing of spheres: A survey, *Proc. Colloquium on Convexity* (Copenhagen, 1965), Københavns Univ. Math. Inst., Copenhagen, 1967, 88—93.

[5] FEW, L. and KANAGASABAPATHY, P.: The double packing of spheres, *J. London Math. Soc.* **44** (1969), 141—146.

[6] LANDAU, E.: *Ausgewählte Abhandlungen zur Gitterpunktlehre*, Herausgegeben von Arnold Walfisz, VEB Deutscher Verlag der Wissenschaften, Berlin, 1962, 292 pp.

[7] LEHMER, D. H.: On the maxima and minima of Bernoulli polynomials, *Amer. Math. Monthly* **47** (1940), 533—538.

[8] WHITTAKER, E. T. and WATSON, G. N.: *A course of modern analysis. An introduction to the general theory of infinite process and of analytic functions; with an account of the principal transcendental functions,* Fourth edition, Reprinted, Cambridge University Press, New York, 1962, VII+608 pp.

(Eingegangen am 13. April 1979)

GYMNASIUM LÜCHOW
D—3130 LÜCHOW
WEST GERMANY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 69—76

GENERALIZED REGULAR NEAR-RINGS

by

A. K. GOYAL and S. C. CHOUDHARY

Introduction

The concepts of π-regularity, semi π-regularity, right and left N-regularity have been defined in near-rings. Corresponding concepts in ring theory are given in [1] and [10]. Section 1 deals with the properties, relations among these and characterizations of these near-rings. Conditions are given under which such near-rings become regular. It is shown that every finite near-ring with identity is π-regular, semi-regular, right and left N-regular. In Section 2, we have dealt with the radical properties of these near-rings. It is proved that the quasi-radical $D(R)$ and the ideal radical $J_0(R)$ ([6], [14]) of π-regular near-ring R are nil, and if R is a right N-regular with identity, then the radical subgroup of R [4] is also nil. Let $Q(R)$ denote the sum of all quasi-regular ideals of the unitary near-ring R. If R is a ring with 1, then $Q(R)= =J(R)$, the Jacobson radical of R. It has been shown that for a near-ring with identity, $Q(R)\subseteqq J_0(R)\subseteqq D(R)\subseteqq J(R)$ and strict inequality may hold in this chain. Also contrary to the ring case, the radical $J(R)$ may not be a nil ideal for these near-rings and therefore, the conditions are given under which $J(R)$ becomes nil.

Preliminaries

Throughout R will denote a left near-ring satisfying $0\cdot x=x\cdot 0=0$ for all x in R.

A subgroup H of R is called an R-subgroup if $HR=\{hr|h\in H,\ r\in R\}\subseteqq H$. A normal subgroup B of R is said to be right ideal of R if $(r_1+b)r_2-r_1r_2\in B$ for all $b\in B$ and $r_1, r_2\in R$. Every right ideal is also an R-subgroup. The converse holds in rings but not in near-rings. An ideal I of R is a right ideal which also satisfies $RI\subseteqq I$.

We denote by $r(a)=\{x\in R|ax=0\}$ to be the right annihilator of an element $a\in R$. It is easy to see that $r(a)$ is a right ideal of R. If R contains no non-zero nilpotent elements, then it can be seen that $xy=0$ implies $yx=0$ and $xry=0$ for all $x, y, r\in R$ [5 Lemma 1].

A right ideal B of R is said to be strictly small if for every R-subgroup A of R, $R=B+A$ implies $R=A$.

A near-ring R is called regular [1] if for each $a\in R$ there exists $x\in R$ such that $a=axa$.

An element $a\in R$ is said to be right quasi-regular if a is in the right ideal generated

AMS (MOS) subject classifications (1970). Primary 16A56; Secondary 16A30.

Key words and phrases. π-regular, semi-π-regular, right and left N-regular, radical subgroup, quasi-radical, the radical JCR, local near rings.

by the set $\{x-ax|x\in R\}$. Thus a is right quasi-regular if the right ideal generated by the set $\{x-ax|x\in R\}$ coincides with R.

Corresponding to the Jacobson radical in ring case the (right) ideals $J_0(R)$, $D(R)$, $J_1(R)$, $J_2(R)$ have been obtained in near-rings by BETSCH [6]. If R contains an identity then $J_1(R)=J_2(R)$ and it is denoted by $J(R)$. In general, we have $J_0(R)\subseteq D(R)\subseteq J_1(R)\subseteq J_2(R)$. $D(R)$ is a right ideal of R and $D(R)$ is an ideal iff $J_0(R)=D(R)$. If R is a d.g. near-ring with identity, then $J_0(R)$ is the ideal radical, $D(R)$ is the quasi-radical and $J_1(R)=J_2(R)$ is the radical discussed by LAXTON [12]. For the definitions of these radicals we also refer to RAMAKOTAIAH [14] where 'quasi-regular' means 'right quasi-regular'. It is known [14] that $J_0(R)$ is the sum of all the right quasi-regular ideals of R and $D(R)$ is the sum of all the right quasi-regular right ideals of R.

Throughout n will denote a positive integer.

§ 1

DEFINITION 1.1. A near-ring R is said to be π-regular (semi π-regular), if for every $a\in R$, there exist $x\in R$ and an integer n such that $a^n=a^nxa^n$ ($a^n=a^nxa$). Such an element a is called π-regular (semi π-regular).

DEFINITION 1.2. A near-ring R is said to be right N-regular (left N-regular) if for every $a\in R$, there exist $x\in R$ and an integer n such that $a^n=a^{n+1}x$ ($a^n=xa^{n+1}$). Clearly, every regular near-ring is π-regular but not conversely and every π-regular near-ring is semi π-regular. Also every right N-regular near-ring is semi π-regular (as $a^n=a^{n+1}x$ implies $a^{n+1}=a^n\cdot a=a^{n+1}xa$).

EXAMPLE 1.3. Let $R=\{0, a, b, c\}$ with addition and multiplication be defined as follows:

$+$	0	a	b	c
0	0	a	b	c
a	a	0	c	b
b	b	c	0	a
c	c	b	a	0

$\cdot$	0	a	b	c
0	0	0	0	0
a	0	0	a	a
b	0	0	b	b
c	0	a	b	c

It can be seen that $\{R, +, \cdot\}$ is π-regular, right N-regular and left N-regular near-ring but R is not regular, as the element a is not a regular element.

REMARKS. In a π-regular near-ring, a^nx and xa^n are idempotents. We note that in a regular near-ring ax and xa are non-zero idempotents. This property has often been used in proving the theorems in regular near-rings [1]. But in a π-regular near-ring, the idempotents a^nx and xa^n may be zero; as in Example 1.3 we have $a^2=a^2ba^2$ and the idempotents a^2b and ba^2 both are zero. Also in Example 1.3 $\{0, a\}$ is a nilpotent R-subgroup of R. This shows that contrary to the case of regular (near) ring, a π-regular near-ring may have non-zero nilpotent R-subgroups.

THEOREM 1.4. *R is a π-regular near-ring if and only if for all $a\in R$, $a^n\in a^nR$ and the principal R-subgroup generated by a^n is the principal R-subgroup generated by an idempotent. The proof is same as that for regular near-rings* [3].

THEOREM 1.5. *The direct products and the homomorphic images of π-regular (semi π-regular, right N-regular, left N-regular) near-rings are also π-regular (semi π-regular, right N-regular, left N-regular).*

The proof is straightforward.

PROPOSITION 1.6. *The center of a π-regular near-ring is also π-regular and in fact right N-regular as well as left N-regular.*

PROOF. Let R be π-regular and $a\in Z(R)$, the center of R. Then $a^n=a^nxa^n$ for some $x\in R$ and some n. This implies $a^n=a^nxa^nxa^n$. It is sufficient to show that $xa^nx\in Z(R)$. If $y\in R$, then since $a^n\in Z(R)$, we have $(a^nx)y=a^n(xy)=(xy)a^n=$ $=(xy)a^nxa^n=a^nxyxa^n=a^nxa^nyx=a^nyx=y(a^nx)$ and so $a^nx\in Z(R)$. Now $(xa^nx)=y=$ $=xa^nxy=xya^nx=a^nxyx=y(a^nx)x=y(xa^nx)$ and therefore, $xa^nx\in Z(R)$. It is easy to see that $Z(R)$ is in fact right N-regular as well as left N-regular.

THEOREM 1.7. *If R is a right and left N-regular near-ring, then it is π-regular.*

PROOF. Let $a\in R$. Then $a^n=a^{n+1}x$ and $a^m=ya^{m+1}$ for some $x, y\in R$ and some positive integers n, m. Therefore, $a^n=a\cdot a^nx=aa^{n+1}xx=a^{n+2}x^2$. In this way, we get $a^n=a^{n+m}x^m$. Similarly, from $a^m=ya^{m+1}$, we get $a^m=y^na^{m+n}$. Now $a^{m+n}=a^n\cdot a^m=$ $=a^{m+n}x^m\cdot y^na^{m+n}=a^{m+n}za^{m+n}$, where $z=x^my^n\in R$ and so R is π-regular.

THEOREM 1.8. *If R is a semi π-regular near-ring with no non-zero nilpotent elements, then R is regular.*

PROOF. If R is semi π-regular, then for $a\in R$, we have $a^n=a^nxa$ for some $x\in R$ and some n. This implies that $a(a^{n-1}-a^{n-1}xa)=0$, and so $(a^{n-1}-a^{n-1}xa)a=0$ since R contains no non-zero nilpotent elements. This would imply $(a^{n-1}-a^{n-1}xa)^2=0$ and therefore $a^{n-1}=a^{n-1}xa$. By downward induction, we obtain $a=axa$. Hence R is regular.

COROLLARY 1.9. If R is a π-regular (or right N-regular) near-ring with no non-zero nilpotent elements, then R is regular.

THEOREM 1.10. *Let R be a π-regular near-ring with no non-zero nilpotent elements. Then R is regular as well as left N-regular.*

PROOF. R is regular by the above corollary. Let $a\in R$. Then $a^n=a^nxa^n$ for some $x\in R$ and some n. This implies $a(a^{n-1}-a^{n-1}xa^n)=0$, and so $(a^{n-1}-a^{n-1}xa^n)a=0$ since R contains no non-zero nilpotent elements. From this we obtain

$$(a^{n-1}-a^{n-1}xa^n)^2=0$$

and so $a^{n-1}=a^{n-1}xa^n$. By downward induction, we have $a=axa^n$, which gives $a^m=ya^{m+1}$, where $y=axa^{n-2}\in R$ and so R is left N-regular.

THEOREM 1.11. *Let R be a near-ring with identity and with no non-zero nilpotent elements. Then the following are equivalent:*

(i) *R is π-regular;*
(ii) *R is right N-regular;*
(iii) *R is semi π-regular.*

PROOF. It is obvious that (i) implies (iii) and (ii) implies (iii). It follows by Theorem 1.8 that (iii) implies (i), and (i) implies (ii) is only to be shown. If (i) holds and $a\in R$, then $a^n=a^n xa^n$ for some $x\in R$ and some n. Clearly, $a^n x$ is an idempotent. Since R contains no non-zero nilpotent elements and R is with identity, by BELL [5, Lemma 2], every idempotent in R is central. This implies $a^n=a^n\cdot a^n x=a^{n+1}y$, where $a^{n-1}x=y\in R$. Hence R is right N-regular.

THEOREM 1.12. *If R is a π-regular near-ring such that for each non-zero $a\in R$, there exists a unique $x\in R$ and an integer n such that $a^n=a^n xa^n$, then R is a near-field.*

PROOF. Let $ab=0$ and $a\neq 0$. Then $a^n b=0$ for any n. This implies $a^n(z+b)a^n=a^n za^n$. By the given hypothesis, we have $z+b=z$, which means $b=0$. So R has no ($\neq 0$) zero-divisors. Since R is π-regular we have $a^n xa^n=a^n$ for some $x\in R$ and some n and $za^n x=e$ is an idempotent. As R has no ($\neq 0$) zero-divisors, $e\neq 0$ and e is a left identity $(\text{as } ex=y\Rightarrow eex=ey\Rightarrow ex=ey\Rightarrow e(x-y)=0\Rightarrow x=y)$. Moreover R will have no ($\neq 0$) nilpotent elements. Therefore, by Corollary 1.9, R is regular. The result now follows from Corollary 3.8 of HEATHERLY [11], which states that a regular near-ring with no ($\neq 0$) zero-divisors and a non-zero right distributive element is a near-field.

DEFINITION 1.13. A near-ring R is said to be an S-near-ring if $a\in aR$ for each $a\in R$. Clearly, if R contains a (right) identity, then R is an S-near-ring. But there exist near-rings (e.g., near-rings 1, 2 of 2.2 given by CLAY [8]) which are S-near-rings without (right) identity.

We now give some characterizations of these near-rings under ascending and descending chain conditions.

THEOREM 1.14. *Let R be an S-near-ring. Then R is right N-regular if and only if for all $a\in R$, the descending chain of principal R-subgroups $aR\supseteqq a^2R\supseteqq a^3R\supseteqq\ldots\supseteqq a^nR\supseteqq a^{n+1}R\ldots$ terminates after a finite number of steps.*

PROOF. Since R is an S-near-ring, the principal R-subgroup generated by a is of the form aR. Let $aR\supseteqq a^2R\supseteqq a^3R\supseteqq\ldots\supseteqq a^nR\supseteqq a^{n+1}R\supseteqq\ldots$ be a descending chain of the principal R-subgroups of R. If R is right N-regular and $a\in R$, then $a^n=a^{n+1}x$ for some $x\in R$ and some n. Now $a^nR=a^{n+1}xR\subseteqq a^{n+1}R=a^n\cdot aR\subseteqq a^nR$ and so $a^nR=a^{n+1}R$ which means that the given chain stops. Conversely, if $a^nR=a^{n+k}R$ (for every $k\geqq 0$), then $a^nR=a^{n+1}R$ and so $a^n=a^{n+1}x$ for some $x\in R$ because R is an S-near-ring ($a^n\in a^nR$). Hence R is right N-regular.

COROLLARY 1.15. If R is an S-near-ring satisfying the d.c.c. on R-subgroups, then R is right N-regular.

THEOREM 1.16. *Let R be a right N-regular near-ring with identity. Then R is left N-regular if, and only if, every ascending chain of right annihilators $r(a)\subseteqq r(a^2)\subseteqq r(a^3)\ldots\subseteqq r(a^n)\subseteqq r(a^{n+1})\subseteqq\ldots$ terminates after a finite number of steps.*

PROOF. Let R be a right N-regular near-ring with identity. If R is left-N-regular, then for $a\in R$, $a^n=xa^{n+1}$ for some $x\in R$ and some n. Let $r(a)\subseteqq r(a^2)\subseteqq r(a^3)\subseteqq\ldots\subseteqq r(a^n)\subseteqq r(a^{n+1})\subseteqq\ldots$ be an ascending chain of right annihilators. If $y\in r(a^{n+1})$, then $a^{n+1}y=0$ which gives $xa^{n+1}y=0$ and so $a^ny=0$, i.e., $y\in r(a^n)$ and therefore $r(a^{n+1})\subseteqq r(a^n)$. Hence $r(a^n)=r(a^{n+1})$, which means that the given chain stops.

Conversely, suppose $r(a^p)=r(a^{p+1})=\dots$. Since R is right N-regular, $a^m=a^{m+1}x$ for some $x\in R$ and some m. Now m and p may be taken to be the same, say equal to n, i.e., $r(a^n)=r(a^{n+1})=r(a^{n+k})$, $k\geqq 0$ and $a^n=a^{n+1}z$ for some $z\in R$. Therefore,

$$(1) \qquad a^n z^k = a^{n+1} z^{k+1} \quad \text{for all} \quad k \geqq 0.$$

Now $a^{n+1}=a^n\cdot a=a^{n+1}za$ implies $a^{n+1}(a-za)=0$, i.e., $(a1-za)\subseteqq r(a^{n+1})=r(a^n)$, so $a^n(1-za)=0$ and we have $a^n=a^nza$. Again, $a^{n+1}=a^n\cdot a=a^nza\cdot a=a^nza^2=a^{n+1}z^2a^2$ from (1). This implies $a^{n+1}(1-z^2a^2)=0$ and so $(1-z^2a^2)\in r(a^{n+1})=r(a^n)$. Therefore, $a^n(1-z^2a^2)=0$, i.e., $a^n=a^nz^2a^2$. Now assume that $a^n=a^nz^ja^j$ for some integer j, then $a^{n+1}=a^n\cdot a=a^nz^ja^j\cdot a=a^{n+1}z^{j+1}a^{j+1}$ from (1). This implies $a^{n+1}(1-z^{j+1}a^{j+1})=0$ and so $(1-z^{j+1}a^{j+1})\in r(a^{n+1})=r(a^n)$. Therefore, $a^n(1-z^{j+1}a^{j+1})=0$, i.e., $a^n= =a^nz^{j+1}a^{j+1}$. Thus $a^n=a^nz^{k+1}a^{k+1}$ for all $k\geqq 0$, so $a^n=a^nz^{n+1}a^{n+1}=xa^{n+1}$, where $a^nz^{n+1}=x\in R$. R is therefore left N-regular.

This theorem with Theorem 1.7 immediately gives the following

COROLLARY 1.17. If R is a right N-regular near-ring with identity satisfying the a.c.c. on right ideals, then R is π-regular.

From 1.15, 1.16 and 1.17, we immediately have the following

THEOREM 1.18. *Every finite near-ring with identity is π-regular, semi π-regular, right N-regular and left N-regular.*

THEOREM 1.19. *Let R be a local near-ring with identity. Then R is π-regular if and only if it is right N-regular.*

PROOF. In view of MAXSON [13, Theorem 4.2], if R is a local near-ring with identity, then the only idempotents of R are 0 and 1. Let R be π-regular and $a\in R$. Then $a^n=a^nxa^n$ for some $x\in R$ and some n. Therefore, a^nx is an idempotent and so $a^nx=0$ or $a^nx=1$. If $a^nx=0$, then clearly $a^n=a^nxa^n=0=a^{n+1}\cdot 0$; and if $a^nx=1$, then $a^n=a^n\cdot 1=a^n\cdot a^nx=a^{2n}x=a^{n+1}a^{n-1}x=a^{n+1}y$, where $a^{n-1}x=y\in R$. Hence in either case R is right N-regular. Conversely, let R be right N-regular. Since R is local with 1, from MAXSON [13, Theorem 2.8], R contains a unique maximal R-subgroup. This means that R satisfies the a.c.c. on R-subgroups and hence on right ideals. The result now immediately follows from Corollary 1.17.

§ 2

THEOREM 2.1. *If R is a π-regular near-ring, then the quasi-radical $D(R)$ is a nil right ideal.*

PROOF. Let R be π-regular and $a\in D(R)$. Then $a^m=a^mxa^m$ for some $x\in R$ and some integer m. Clearly, $a^mx=e$ (say) is an idempotent and $e\in D(R)$ since $D(R)$ is a right ideal of R. If $e\neq 0$, then e is a non-zero right quasi-regular element because each element of $D(R)$ is right quasi-regular (see [14, Theorem 2.2]). But by RAMAKOTAIAH [14, Theorem 4.2], a non-zero idempotent cannot be a right quasi-regular element. Hence $e=0$, and so $a^m=ea^m=0$, i.e., a is a nilpotent element, which proves the result.

Since $D(R)$ contains all the nil right ideals of R [14, Corollary 2.5] and $U(R)\subseteqq \subseteqq J_0(R)\subseteqq D(R)$, where $U(R)$ is the upper nil-radical of R, which is the sum of all the nil ideals of R [14, Corollary 2 of Theorem 4], this theorem immediately gives the following corollaries:

COROLLARY 2.2. For a π-regular near-ring, $D(R)$ is the sum of all the nil right ideals of R.

COROLLARY 2.3. For a π-regular near-ring, $J_0(R)$ is a nil ideal and $J_0(R)=$ $=U(R)$, the sum of all the nil ideals of R.

THEOREM 2.4. *Let R be a π-regular near-ring having all the nilpotent elements in the center of R. Then $D(R)=J_0(R)=U(R)=$the sum of all the nil ideals of R.*

PROOF. It suffices to show that the elements of $D(R)$ are closed from the left by the elements of R. If $a\in D(R)$, then a is nilpotent by the Theorem 2.1. Therefore, a lies in the center and so $ax=xa$ for all $x\in R$. Hence $xa\in D(R)$ for all $x\in R$.

DEFINITION 2.5 ([4]). An element $x\in R$ is called quasi-regular if there is an element $x'\in R$ such that $(1-x)x'=1$. An ideal (right ideal, R-subgroup) is called quasi-regular if its every element is quasi-regular.

If R happens to be a ring with identity, then it can be seen that the concepts of 'right quasi-regularity' and 'quasi-regularity' become equivalent. But in near-rings with identity, we note that quasi-regularity implies right quasi-regularity, whereas the converse need not hold. For this we consider the following

EXAMPLE 2.6. Let $R=\{0, a, b, c\}$ with addition and multiplication be defined as follows:

$+$	0	a	b	c
0	0	a	b	b
a	a	0	c	b
b	b	c	0	a
c	c	b	a	0

$\cdot$	0	a	b	c
0	0	0	0	0
a	0	a	b	c
b	0	b	0	0
c	0	c	b	c

Then $\{R, +, \cdot\}$ is a unitary near-ring and the element b is right quasi-regular but not quasi-regular. The only quasi-regular element in R is '0'. Also, if R is a ring with 1, then every nilpotent element is quasi-regular. But in near-rings with 1, this may not be true; as in this Example, the element b is nilpotent but not quasi-regular.

DEFINITION 2.7 ([4]). A radical subgroup of R is defined as the intersection of all the maximal R-subgroups of R.

We shall use the following theorem (mentioned here as a Lemma) of BEIDLEMAN.

LEMMA 2.8 ([4, Theorem 2.2]). *The radical subgroup A of a near-ring R is a quasi-regular R-subgroup that contains all the quasi-regular right ideals of R.*

We shall denote by $Q(R)$ the sum of all the quasiregular ideals.

THEOREM 2.9. *$Q(R)$ is a quasi-regular ideal.*

PROOF. If B is any quasi-regular ideal of R, then by the above Lemma, $B \subseteq A$, the radical subgroup of R. Hence $Q(R) \subseteq A$. Since A is quasi-regular by 2.8, $Q(R)$ is also quasi-regular.

THEOREM 2.10. *For a near-ring R with identity*

$$Q(R) \subseteq J_0(R) \subseteq D(R) \subseteq J(R).$$

PROOF. It is known that $J_0(R) \subseteq D(R) \subseteq J(R)$ ([6], [14], [12]) and so it remains to see that $Q(R) \subseteq J_0(R)$. This follows from Theorem 2.9 and the fact that each quasi-regular ideal is right quasi-regular ideal and $J_0(R)$ contains all the right quasi-regular ideals [14, Theorem 2.3].

NOTE. If R happens to be a ring with 1, then $Q(R)=J_0(R)=D(R)=J(R)$, the Jacobson radical of R. Also in a ring the upper nil-radical is always contained in the Jacobson radical [9]. But in near-rings the situations may be different, since in Example 2.6, $(0)=Q(R) \subsetneqq U(R)=\{0, b\}$, where $U(R)$ is the upper nil-radical of the near-ring R. The example 2.6 also shows that $Q(R) \subsetneqq J_0(R)$ $(=\{0, b\})$. Moreover, the examples where $J_0(R) \subsetneqq D(R) \subsetneqq J(R)$ are given in ([6] and [12]). Therefore, we may have the situation where the strict inequality may hold in the above chain.

This Theorem with Theorem 2.1, gives the following

COROLLARY 2.11. If R is a π-regular near-ring with 1, then $Q(R)$ is a nil ideal and so $U(R)=Q(R)=J_0(R)$.

THEOREM 2.12. *If R is a right N-regular near-ring with identity, then the radical subgroup of R is nil.*

PROOF. If A is the radical subgroup of the right N-regular near-ring R, then A is quasi-regular R-subgroup by Lemma 2.8. Let $a \in A$, then $a^n=a^{n+1}x$ for some $x \in R$ and some n, since R is right N-regular. As $a \in A$, we have $ax \in A$ and ax is a quasi-regular element. Therefore, $(1-ax)x'=1$ for some $x' \in R$. This implies that $a^n(1-ax)x'=a^n$, i.e., $a^n=(a^n-a^{n+1}x)x'=0$ and hence A is nil.

COROLLARY 2.13. If R is a right N-regular near-ring with identity, then $Q(R)$ is a nil ideal.

NOTE. If R is π-regular or semi π-regular or right or left N-regular ring with identity, then the Jacobson radical $J(R)$ is nil (see [1], Remark p. 39). But in near-rings $J(R)$ may not be a nil ideal. For this, consider the near-ring T of LAXTON [12, Section 4], for which $J(T)$ is non-nilpotent ideal. That near-ring T is a finite (d.g.) near-ring with identity. Therefore by Theorem 1.19, T is π-regular, semi π-regular, right N-regular and left N-regular. Moreover for finite near-rings, the concepts of nilpotence and nil coincide (see [3] and [7]). Hence $J(T)$ is non nil. Also the finite near-ring R_3 with 1 of BETSCH [6, 4.3] where $J_0(R_3) \subsetneqq D(R_3) \subsetneqq J(R_3)$, shows that $J(R_3)$ may not be a nil ideal. $J(R)$ becomes nil under some conditions given below:

THEOREM 2.14. *If R is a right N-regular local near-ring with identity and $R \neq J(R)$, then $J(R)$ is a nil ideal.*

PROOF. Since R is a local near-ring, by Lemma 2.9 and Theorem 2.10 of MAXSON [13], $J(R)$ will be a quasi-regular ideal and so $J(R) \subseteqq Q(R)$. This with Theorem 2.10 gives $J(R) = Q(R)$ which is nil by Corollary 2.13.

THEOREM 2.15. *If R is a right N-regular near-ring with identity in which $J(R)$ is strictly small, then $J(R)$ is a nil ideal.*

PROOF. In view of BEIDLEMAN [2, Cor. 1], if $J(R)$ is strictly small then $J(R)$ is the radical subgroup of R. The result now follows from Theorem 2.12.

REFERENCES

[1] AZUMAYA, G.: Strongly π-regular rings, *J. Fac. Sci. Hokkaido Univ. Ser. I,* **13** (1954), 34—39.
[2] BEIDLEMAN, J. C.: A radical for near-ring modules, *Michigan Math. J.* **12** (1965), 377—383.
[3] BEIDLEMAN, J. C.: A note on regular near-rings, *J. Indian Math. Soc.* **33** (1969), 207—210.
[4] BEIDLEMAN, J. C.: Quasi-regularity in near-rings, *Math. Z.* **89** (1965), 224—229.
[5] BELL, H. E.: Near-rings in which every element is a power of itself, *Bull. Austral. Math. Soc.* **2** (1970), 363—368.
[6] BETSCH, G.: *Struktursätze für Fastringe,* Doctoral Dissertation, Universität Tübingen, 1963.
[7] CHOUDHARY, S. C. and TEWARI, K.: (NB) property in near-rings, *Rivista Math. Univ. Parma* (to appear).
[8] CLAY, J. R.: The near-rings on groups of low order, *Math. Z.* **104** (1968), 364—371.
[9] DIVINSKY, N. J.: *Rings and Radicals,* Mathematical Expositions no. 14, University of Toronto Press, 1965.
[10] DRAZIN, M. P.: Rings with central idempotent or nilpotent elements, *Proc. Edinburgh Math. Soc.* (2) **9** (1958), 157—165.
[11] HEATHERLY, H. E.: Regular near-rings, *J. Indian Math. Soc.* **38** (1974), 345—354.
[12] LAXTON, R. R.: Prime ideals and ideal radical of a d.g. near-ring, *Math. Z.* **83** (1964), 8—17.
[13] MAXSON, C. J.: On local near-rings, *Math. Z.* **106** (1968), 197—205.
[14] RAMAKOTAIAH, D.: Radical for near-rings, *Math. Z.* **97** (1967), 45—56.
[15] VANDERWALT, A. P. J.: Prime ideals and nil radicals in near-rings, *Arch. Math.* **15** (1964), 408—414.

(Received May 21, 1979)

COLLEGE OF TECHNICAL AND AGRICULTURAL ENGINEERING
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF UDAIPUR
IND—313 001 UDAIPUR
INDIA

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 77—82

A TOPOLOGICAL SEMIGROUP OF QUOTIENTS

by

JOHN K. LUEDEMAN

Abstract

A monoid S has a maximal semigroup of right quotients $Q(S)$ in the sense of McMorris. For a locally compact topological monoid S, a topological semigroup of quotients $C(S)$ is constructed, and every topological right S-system is shown to be continuously embeddable in $C(S)$. A compact connected Hausdorff semigroup cannot be properly embedded as an open subsemigroup in any semigroup of quotients. A necessary and sufficient condition for S to be openly embedded in its classical semigroup of quotients is also given.

In his thesis, R. L. Johnson [2] described a topological ring of quotients $C(R)$ of a topological ring R in which R is an open subring. The crux of his work was determining the additive subgroup $C(R)$ of the maximal ring of quotients $Q(R)$ of R in which multiplication is continuous, because $Q(R)$ is already a topological group when one takes the neighborhood system of 0 in R as a basis for the neighborhood system in $Q(R)$.

F. R. McMorris [4] constructed a semigroup of quotients $Q(S)$ for a semigroup S and proved that it possessed properties similar to those of $Q(R)$ for rings. In this paper we propose a candidate $C(S)$ for a maximal topological semigroup of quotients in which a locally compact monoid S is openly embedded. Unlike the theory of topological rings, the topology of S is not necessarily determined by the neighborhood system of one of its elements. Consequently we must first find a method of extending the topology of S to $Q(S)$, or at least a subset thereof.

We require the following definitions from [4].

Definition. A right S-system $T \supseteqq Y$ is a *dense S-extension* of the right S-system Y if for $t_1 \neq t_2$, $t_3 \in T$ there is some $s \in S$ with $t_1 s \neq t_2 s$ and $t_3 s \in Y$.

A semigroup T is a *semigroup of right quotients* of S if T is a dense S-extension of S.

The *maximal semigroup of right quotients* $Q(S)$ *of* S is constructed as follows: Let $\mathcal{D} = \{D : D$ is a dense right ideal of $S\}$, and let $T = \bigcup_{D \in \mathcal{D}} \operatorname{Hom}(D, S)$ where $\operatorname{Hom}(D, S)$ is the collection of all right S-homomorphisms from D to S. Define a right congruence Θ on T by $f \Theta g$ iff there is a dense right ideal D of S on which f and g agree. Then

$$Q(S) = T/\Theta.$$

AMS (MOS) subject classifications (1980). Primary 20M50; Secondary 20M99.

Key words and phrases. Topological semigroup of quotients, semigroup of quotients, open embedding.

Note that if $q \in Q(S)$, then $q^{-1}S = \{s \in S \mid qs \in S\}$ is a dense right ideal of S, and if $q \in Q(S)$ and $s \in q^{-1}S$ and q is represented by f, then $f(s) = qs$.

We refer the reader to McMorris [4] for a discussion of semigroups of quotients in the algebraic case.

§ 1. The construction of $C(S)$

Let S be a locally compact topological monoid. A topological semigroup $T \supseteq S$ is a topological semigroup of right quotients of S if T is a dense S-extension of S. Consequently, multiplication in T must be bicontinuous, and the evaluation mappings $q: q^{-1}S \to S$ given by $q(t) = qt$ must be continuous for all $q \in T$. As is well-known [3], this requires that the topology of T contain the compact open topology of T and so of S. Also, if we wish S to be open in T, $q^{-1}S$ is necessarily an open dense right ideal of S.

With the foregoing discussion as rationale, we topologize $Q(S)$ by taking as a subbasis the collection $\{M(K, U): K \text{ compact in } S,\ U \text{ open in } S\}$ where $M(K, U) = \{q \in Q(S): qK \subset U\}$. Note that this implies that $K \subset q^{-1}S$ and that $S = M(\{1\}, S)$ is open in $Q(S)$.

Let $C(S) = \{q \in Q(S): q^{-1}S \text{ is open in } S \text{ and } q \text{ is continuous on } q^{-1}S\}$.

Proposition 1.1. *$C(S)$ is a subsemigroup of $Q(S)$ containing S.*

Proof. Since S is a topological semigroup, for $t \in S$, $t^{-1}S = S$ is open in S and t is continuous on S.

Moreover, if $r, q \in C(S)$, then $(rq)^{-1}S = q^{-1}(r^{-1}S) \cap q^{-1}S$ which is an open dense right ideal of S. Also, since $q^{-1}(r^{-1}S)$ is mapped into $r^{-1}S$ by q, and q is continuous on $q^{-1}S$ and r is continuous on $r^{-1}S$, the composition rq is continuous on $q^{-1}(r^{-1}S) \cap q^{-1}S$. Thus $C(S)$ is closed under the multiplication of $Q(S)$ and so is a subsemigroup.

Proposition 1.2. *$C(S)$ is a topological semigroup in the subspace topology of $Q(S)$.*

Proof. We need only show that multiplication is bicontinuous in $C(S)$. To this end let $f, g \in C(S)$ and $fg \in M(K, U)$. Then K is compact in S, $K \subset g^{-1}S$ and $g(K) \subset U$ is compact. Hence for $k \in K$, there is an open neighborhood W_k of k with $\overline{W}_k$ compact and $g(k) \in W_k \subset \overline{W}_k \subset f^{-1}(U)$ since both f and g are continuous. Since $\{W_k: k \in K\}$ is an open cover of K, there are $k_1, \dots, k_n \in K$ with $W = \bigcup_1^n W_{k_i} \supset K$. Then $g \in M(K, W)$, $f \in M(\overline{W}, U)$ and $M(\overline{W}, U)M(K, W) \subset M(K, U)$.

Corollary 1.3. *S is a topological subsemigroup of $C(S)$.*

Proof. S is open in $Q(S)$ and so in $C(S)$ and since S is locally compact, it has the subspace topology of $Q(S)$.

We call $C(S)$ the *maximal topological semigroup of quotients* of S. The name is justified by the next result.

THEOREM 1.4. *Let T be a topological semigroup of quotients openly containing S. Then there is a continuous monomorphism α: $T \to C(S)$ whose restriction to S is the identity.*

PROOF. For $t \in T$, $t^{-1}S$ is an open dense right ideal of S. Thus φ: $t^{-1}S \to S$ given by $\varphi(x)=tx$ determines a unique element $q \in Q(S)$ [4]. Define α: $T \to Q(S)$ by $\alpha(t)=q$.

α is one-to-one for if $\alpha(t)=\alpha(g)$, then $tx=gx$ for all x in some dense right ideal of S. Thus [4] $ts=gs$ for all $s \in S$ and so $t=t1=g1=g$.

α is continuous for if $t \in T$, $\alpha(t)=q$ and $q \in M(K, U)$ then K is compact in S, U is open in S and for $k \in K$, $q(k)=qk=tk \in U$. Then $t \in M^*(K, U)=\{t \in T \mid tK \subset U\}$. Since T is a topological semigroup, the topology of T contains the compact open topology of S and so $M^*(K, U)$ is open in T and $\alpha(M^*(K, U)) \subseteq M(K, U)$.

$\alpha(T) \subset C(S)$ for if $\alpha(t)=q$, $q^{-1}S=t^{-1}S$, a dense open right ideal of S, and since the mapping $s \to ts$ is continuous on $t^{-1}S$, q is continuous on $q^{-1}S$ and $q \in C(S)$.

Finally, $\alpha(s)=s$ for $s^{-1}S=S$ and if $\alpha(s)=q$, then $q^{-1}S=S$ and so $q(1)= =q \cdot 1=q \in S$ and $q=q \cdot 1=s \cdot 1=s$.

In the above theorem we would like α to be a homeomorphism, however this is seldom the case. Since T is a topological semigroup, the multiplication (composition of mappings t: $t^{-1}S \to S$) is continuous and so the topology of T contains the compact open topology of S. However, the compact open topology of S does not necessarily make the multiplication in T (or $Q(S)$) continuous. In general, the topology of T is finer than the compact open topology of T of S. Thus $C(S)$ is algebraically the maximal topological semigroup in which S can be openly embedded and topologically $C(S)$ has the coarsest topology which will do the job.

A brief examination of the proof of the last theorem shows that the semigroup structure of T was little used. T need only be a topological S-system — i.e., the mapping $T \times S \to T$ is bicontinuous. In this case if S is dense in T we may continuously embed T in $C(S)$. This is useful in considering $C(C(S))$. Since $C(S) \subset Q(S)$, we may give $Q(Q(S))$ the compact open topology determined by the compact and the open subsets of $C(S)$. Forming $C(C(S))$ as those $t \in Q(Q(S))$ for which $t^{-1}C(S)$ is a dense and open S-subsystem of $C(S)$ and t: $t^{-1}C(S) \to C(S)$ is continuous, then as S-systems

$$S \subset C(S) \subset C(C(S)) \subset Q(Q(S)) = Q(S)$$

algebraically. Moreover, $C(S)$ is open in $C(C(S))$ and so S is open in $C(C(S))$. Consequently, there is a continuous embedding $\mathscr{B}$: $C(C(S)) \to C(S)$ which is a right S-homomorphism whose restriction to S is the identity. Thus $C(C(S))=C(S)$ algebraically but the topology of $C(C(S))$ is generally finer than the topology of $C(S)$. Moreover, $C(C(S))$ is not in general a topological semigroup.

Returning to $C(S)$, we have the following alternate description of $C(S)$. JOHNSON [2] required S to be a commutative ring for this description of $C(S)$ to be valid.

PROPOSITION 1.5. $C(S)=\{q \in Q(S): q|_S$ *is continuous*$\}$.

PROOF. Let T denote the set on the right-hand side of the above equation. If $q \in T$, then $q \in Q(S)$ so $q^{-1}S$ is a dense right ideal of S. Since $q|_S$ is continuous, and S is open in $Q(S)$, $q^{-1}S$ is open in S and $Q(S)$. Thus, q: $q^{-1}S \to S$ is continuous and $q \in C(S)$.

Conversely, since $C(S)$ is a topological semigroup openly containing S, if $q \in C(S)$, then left multiplication by q is continuous so $q|_S$ is continuous.

§ 2. Products

Let $\{S_i : i \in I\}$ be a collection of topological semigroups each S_i having a zero 0_i and an identity e_i. Let ΠS_i denote the cartesian product of these semigroups. We require the following proposition due to HINKLE [1].

PROPOSITION 2.1. *$Q(\Pi S_i)$ is semigroup isomorphic to $\Pi Q(S_i)$.*

Sketch of proof. For $q \in Q(\Pi S_i)$, let $q_i \in Q(S_i)$ be defined by $q_i(s_i) = \pi_i q(s_i)$ where π_i: $\Pi S_j \to S_i$ is the i^{th} projection mapping. Then φ: $Q(\Pi S_i) \to \Pi Q(S_i)$ given by $\varphi(q) = \Pi q_i$ is the required isomorphism.

In order for ΠS_i to be locally compact, we require that all but a finite number of the S_i be compact while the remaining S_i are locally compact. Let $\mathcal{T}_1$ be the topology of $Q(\Pi S_i)$ and $\mathcal{T}_2$ be the product topology on $\Pi Q(S_i)$ where $Q(\Pi S_i)$ and $Q(S_i)$ are given the compact open topology of Section 1.

THEOREM 2.2. $\mathcal{T}_1 \supseteq \mathcal{T}_2$.

PROOF. We will show that the isomorphism φ defined above is continuous, thus when $Q(\Pi S_i)$ and $\Pi Q(S_i)$ are identified, $\mathcal{T}_1 \supseteq \mathcal{T}_2$. To this end let $\Pi q_i \in \Pi Q(S_i)$ and U be a basic open neighborhood of Πq_i. Then $U = \prod_{k \in K} Q(S_k) \times \prod_{j \in J} U_j$ where J is finite, $J \cap K = \emptyset$, $J \cup K = I$, and U_j is an open neighborhood of q_j for all $j \in J$. Let $\varphi(q) = \Pi q_i$, then by the definition of $\mathcal{T}_1$ and without loss of generality, $U_j = M(K_j, V_j)$ for $j \in J$ where K_j is compact and V_j is open in S_j. Therefore $V = \prod_{k \in K} S_k \times \prod_{j \in J} V_j$ is open and $K = \prod_{k \in K} \{0_k\} \times \prod_{j \in J} K_j$ is compact in ΠS_i, and $q \in M(K, V)$. Moreover, if $y \in M(K, V)$ then for $j \in J$, $y_j(k_j) \in V_j$ for $k_j \in K_j$ and $j \in J$. Thus $\varphi(y) \in U$ and φ is continuous.

COROLLARY 2.3. *If I is finite, then φ is a homeomorphism.*

PROOF. Let $x \in Q(\Pi S_i)$ and $U = M(K, V) = M(K, \Pi V_i)$ be a subbasic open neighborhood of x. Then $\varphi(x) = \Pi x_i$ and if $K_i = \pi_i(K)$, then $x_i(K_i) \subset V_i$ so that $x_i \in M(K_i, V_i)$. Since $\Pi M(K_i, V_i)$ is open in $\Pi Q(S_i)$, and $\varphi(x) \in \Pi M(K_i, V_i) = \varphi(M(K, V))$ for all x, then φ is an open mapping.

We next look at $C(\Pi S_i)$ and $\Pi C(S_i)$. In this discussion we identify $Q(\Pi S_i)$ and $\Pi Q(S_i)$ algebraically.

THEOREM 2.4. *As semigroups, $C(\Pi S_i) \subset \Pi C(S_i)$.*

PROOF. Let $q = \Pi q_i \in C(\Pi S_i)$, then $q^{-1}(\Pi S_i)$ is a dense open right ideal of ΠS_i. Since $\Pi 0_i \in q^{-1}(\Pi S_i)$, there is a basic open neighborhood $\prod_{k \in K} S_k \times \prod_{j \in J} U_j \subset q^{-1}(\Pi S_i)$ of $\Pi 0_i$, where J is finite, $J \cap K = \emptyset$ and $J \cup K = I$. Then for $k \in K$, $q_k^{-1} S_k = S_k$ — i.e., $q_k \in S_k$. Now, let $x_j \in q_j^{-1} S_j$, $j \in J$. Then there is $x \in q^{-1}(\Pi S_i)$ with $\pi_j x = x_j$. Since $q^{-1}(\Pi S_i)$ is open, there is a basic open set $x \in U \subseteq q^{-1}(\Pi S_i)$

so that $x_j \in U_j = \pi_j(U) \subset q_j^{-1} S_j$ is open. Since this is true for all $x_j \in q_j^{-1} S_j$, $q_j^{-1} S_j$ is an open dense right ideal of S_j. Since q is continuous on $q^{-1}(\Pi S_i)$ and the j^{th} projection π_j is continuous, we have that $q_j = \pi_j q$ is continuous on $q_j^{-1} S_j$ and so $q_j \in C(S_j)$ and $q = \Pi q_i \in \pi C(S_i)$.

The following result shows the structure of $C(\Pi S_i)$ in $Q(\Pi S_i)$.

THEOREM 2.5. *$C(\Pi S_i) = \{q = \Pi q_i \in Q(\Pi S_i)$: $q_j \in S_j$ for all but a finite number of $j \in I$ and $q_j \in C(S_j)$ for all $j \in I\}$.*

PROOF. Let T denote the right-hand set in the statement of the theorem. In the proof of the previous theorem we showed that $C(\Pi S_i) \subseteqq T$.

Conversely, let $t = \Pi t_i \in T$. Then $t^{-1}(\Pi S_i) = \Pi t_i^{-1} S_i$. Since $t_i^{-1} S_i = S_i$ for all but a finite number of i and the remaining $t_j^{-1} S_j$ are dense open right ideals of S_j, then $t^{-1}(\Pi S_i) = \Pi t_i^{-1} S_i$ is a dense open right ideal of ΠS_i. Moreover, since $t = \Pi t_i$ and each t_i is continuous on $t_i^{-1} S_i$, then t is continuous on $t^{-1}(\Pi S_i)$ so that $t \in C(\Pi S_i)$.

COROLLARY 2.6. *$C(\Pi S_i) = \Pi C(S_i)$ as semigroups iff $C(S_i) = S_i$ for all but a finite number of $i \in I$. Moreover, if $C(\Pi S_i) = \Pi C(S_i)$ as topological semigroups, $C(S_i) = S_i$ for all but a finite number of $i \in I$.*

§ 3. The position of $C(S)$

If S is a discrete topological semigroup then only finite sets are compact and S is locally compact. The compact open topology on $Q(S)$ makes each q: $q^{-1} S \to S$ continuous and so $C(S) = Q(S)$.

PROPOSITION 3.1. *If S is a discrete topological semigroup, then $Q(S) = C(S)$.*

On the other extreme, suppose that S has no proper open dense right ideals, then for $q \in C(S)$, $q^{-1} S = S$ and so $q = q(1) \in S$.

PROPOSITION 3.2. *If S has no proper open right ideals then $C(S) = S$.*

LEMMA 3.3. *If S is Hausdorff then $C(S)$ is Hausdorff.*

PROOF. Let $q_1 \neq q_2$ be elements of $C(S)$. Then for some $s \in q_1^{-1} S \cap q_2^{-1} S$, $q_1 s \neq q_2 s$. Since S is Hausdorff there are disjoint open sets V_1 and V_2 with $q_1 s \in V_1$ and $q_2 s \in V_2$. Thus $q_1 \in M(\{s\}, V_1)$, $q_2 \in M(\{s\}, V_2)$ and $M(\{s\}, V_1) \cap M(\{s\}, V_2) = \emptyset$.

THEOREM 3.4. *If S is a compact connected Hausdorff topological semigroup then $C(S) = S$.*

PROOF. Let $q \in C(S)$, then $q^{-1} S$ is open. Moreover, $q^{-1} S$ is closed since S is compact in $Q(S)$ and so closed. Thus $q^{-1} S = S$ and so $q \in S$.

Note that by Theorem 1.4, a connected compact Hausdorff topological semigroup cannot be openly embedded as a proper (algebraically) dense subsemigroup of any topological semigroup.

The following example shows that $C(S)$ may lie strictly between S and $Q(S)$.

EXAMPLE. Let $T = \{0\} \cup \{1/2^n$: $n = 0, 1, 2, \ldots\}$ be the multiplicative subsemigroup of the real numbers with the subspace topology. Then T is compact, Hausdorff and $\{0\}$ is the only limit point. It is easily seen that $Q(T) = \{0\} \cup \{2^n$: n an integer$\}$.

Let S be the product of a countable number of copies of T with the product topology. Then $q\in Q(S)\setminus C(S)$ if q is defined by $\pi_i(q)=2$ for all i. Moreover, $y\in C(S)\setminus S$ if y is defined by $\pi_1(y)=2$ and $\pi_i(y)=0$ for $i\neq 1$ by Theorem 2.5. Thus $S\subsetneqq C(S)\subsetneqq Q(S)$.

§ 4. The classical semigroup of quotients

In this section let S be a locally compact monoid, C be the set of cancellable elements of S, and let S have the common right multiple property with respect to C — i.e., if for $a\in S$, $b\in C$ there is $x\in S$, $y\in C$ satisfying $bx=ay$. In this case we may form the classical semigroup of right quotients $Q(S, C)$ of S with respect to C. The elements of $Q(S, C)$ may be represented as ac^{-1} where $a\in S$ and $c\in C$. The following result gives necessary and sufficient conditions for S to be openly embeddable in $Q(S, C)$.

THEOREM 4.1 [5]. *A necessary and sufficient condition that S be openly embeddable in $Q(S, C)$ is that for all $c\in C$, left multiplication by c is an open mapping.*

PROOF. $\Rightarrow$: Let S be openly embeddable in $Q(S, C)$ and $c\in C$. Then $c^{-1}\in Q(S, C)$ and so if V is open in S, $(c^{-1})^{-1}V=cV$ is open in S. Thus left multiplication by c is an open mapping.

$\Leftarrow$: Conversely, let left multiplication by c be an open mapping for all $c\in C$. Since $Q(S, C)\subseteqq Q(S)$ we give $Q(S, C)$ the subspace topology from $Q(S)$. We will show that $Q(S, C)\subseteqq C(S)$. To this end let $ab^{-1}\in Q(S, C)$. Then $(ab^{-1})^{-1}S= =\{s\in S\colon ab^{-1}s\in S\}$ is a dense right ideal of S. Choose $x\in(ab^{-1})^{-1}S$ so that $ab^{-1}x\in S$. Let V be an open neighborhood of $ab^{-1}x$. Then $x\in b(a^{-1}V)$ which is open in S since $b\in C$ and $a^{-1}V$ is open in S.

Thus $ab^{-1}[b(a^{-1}V)]\subset V\subset S$ and so $b(a^{-1}V)\subset(ab^{-1})^{-1}S$. Since this is true for all $x\in(ab^{-1})^{-1}S$, $(ab^{-1})^{-1}S$ is open in S. This also shows that ab^{-1} is continuous on $(ab^{-1})^{-1}S$ and so $ab^{-1}\in C(S)$. Consequently, both algebraically and topologically $S\subset Q(S, C)\subset C(S)$ and so S is openly embeddable in $Q(S, C)$.

REFERENCES

[1] HINKLE, C. V.: Generalized semigroups of quotients, *Trans. Amer. Math. Soc.* **183** (1973), 87—118.

[2] JOHNSON, R. L.: Rings of quotients of topological rings, *Math. Ann.* **179** (1969), 203—211.

[3] KELLEY, J. L.: *General Topology,* Van Nostrand, Princeton, New Jersey, 1955.

[4] MCMORRIS, F. R.: On quotient semigroups, *Journal of Mathematical Sciences* **7** (1972), 48—56.

[5] WEINERT, H. J.: Semigroups of Right Quotients of Topological Semigroups, *Trans. Amer. Math. Soc.* **147** (1970), 333—348.

(Received July 23, 1979)

DEPARTMENT OF MATHEMATICAL SCIENCES
O-104 MARTIN HALL
CLEMSON UNIVERSITY
CLEMSON, SC 29631
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 83—86

A COMMUTATIVITY THEOREM FOR RINGS WITH CONSTRAINTS INVOLVING NILPOTENT ELEMENTS

by

HAZAR ABU-KHUZAM and ADIL YAQUB

A well-known theorem of JACOBSON [4; p. 217] asserts that a ring R with the property that for each x in R there exists an integer $n(x)>1$ such that $x^{n(x)}=x$ is necessarily commutative. Our objective is to prove the following theorem the case $N=\{0\}$ of which yields Jacobson's Theorem.

THEOREM 1. *Suppose that R is a ring with center Z and N is the set of nilpotent elements of R. Suppose that* (i) *N is commutative;* (ii) *for all $a\in N$ and $b\in R$, $ab-ba$ commutes with b;* (iii) *for all $b\in R$, we have $b\in Z$ or $b^{n(b)}-b\in N$ for some integer $n(b)>1$. Then R is commutative (and conversely).*

Theorem 1 also generalizes some of the results in [5], as will be shown below. Moreover, we give examples which demonstrate that Theorem 1 need not be true if any of the hypotheses is deleted.

PROOF of Theorem 1. The proof will be broken into several claims.

CLAIM 1. *The idempotents of R are in the center of R.*

PROOF. Suppose that $e^2=e\in R$, $x\in R$. Since $ex-exe\in N$, hypothesis (ii) implies that $[(ex-exe)e-e(ex-exe)]$ commutes with e, and hence $ex=exe$. Similarly, $xe=exe$.

CLAIM 2. *N is an ideal in R.*

PROOF. The proof was essentially given in [3]. However, for the sake of self-containedness, we give it here also.

Let $a\in N$, $b\in R$. Let $a^h=0$. If $ab\in Z$, then $(ab)^h=a^h b^h=0$, and hence $ab\in N$. Similarly, $ba\in N$. So suppose $ab\notin Z$. Then, by hypothesis (iii), $(ab)^n-ab\in N$ for some integer $n=n(ab)>1$. Hence we have $[(ab)^n-ab]^m=0$ for some positive integer m, and thus there exists a polynomial f with integer coefficients such that $(ab)^m=(ab)^{m+1}f(ab)$. Let $c=f(ab)$, $d=c^m$, $e=(ab)^m d$. Then, as is readily verified,

$$(ab)^m=(ab)^m e \quad \text{and} \quad e^2=e. \tag{1}$$

(Compare with the proof of Lemma 1.3.2 in [2].) By Claim 1, e is in the center of R, and hence

$$e=e^2=e(ab)^m d=aeb(ab)^{m-1}d=\ldots=a^h e\{b(ab)^{m-1}d\}^h=0,$$

AMS (MOS) subject classifications (1970). Primary 16A70.

since $a^h=0$. Thus, $e=0$ and hence by (1), $(ab)^m=0$. Therefore, ab is nilpotent. Similarly, ba is nilpotent. We have thus shown that ab and ba are nilpotent, for all $a\in N$ and $b\in R$. Combining this and hypothesis (i), we conclude that N is an ideal in R.

CLAIM 3. *Any homomorphic image of R satisfies the conditions* (i), (ii), (iii).

PROOF. Suppose f: $R\to R^*$ is a homomorphism of R onto R^*. Let N^* be the set of nilpotent elements of R^*, and Z^* the center of R^*. We claim that

$$N^*\subset f(N)\cup Z^*. \tag{2}$$

The argument is similar to that given in [3]. Thus, suppose that $d^*\in N^*$ with $(d^*)^k=0$, and suppose $d^*\notin Z^*$. Choose $d\in R$ such that $f(d)=d^*$. Then $d\notin Z$, and hence by hypothesis (iii), $d-d^n\in N$ for some integer $n=n(d)>1$. Let $d'=d^{n-2}$. Then $d-d^2d'\in N$. Since N is an ideal in R (by Claim 2), we conclude that

$$d-d^{k+1}(d')^k=(d-d^2d')+dd'(d-d^2d')+\ldots+(dd')^{k-1}(d-d^2d')$$

is in N. Recall that $f(d)=d^*$ and $(d^*)^k=0$, and hence $d^*\in f(N)$. This proves (2).
In view of (2), we see that N^* is commutative, and hence condition (i) holds for R^*. Now, suppose that $a^*\in N^*$, $b^*\in R^*$. If $a^*\in Z^*$ then, clearly, $a^*b^*-b^*a^*$ commutes with b^*. So suppose that $a^*\notin Z^*$. Then, by (2), $a^*\in f(N)$ and hence $a^*=f(a)$ for some a in N. Let $f(b)=b^*$, $b\in R$. By hypothesis (ii), $ab-ba$ commutes with b, and hence $a^*b^*-b^*a^*$ commutes with b^*, which proves that condition (ii) holds for R^*. That condition (iii) holds for R^* is obvious.
We are now in a position to complete the proof of Theorem 1. In view of Claim 3, we may (and shall) assume that R is subdirectly irreducible. First, we will show the following:

$$\begin{array}{c}\text{If}\quad x\in R\quad\text{and}\quad x\notin Z,\quad\text{then}\quad x\in N\quad\text{or}\\ R\text{ has an identity }1\text{ and }x\text{ has an inverse in }R.\end{array} \tag{3}$$

To prove (3), let $x\in R$ and $x\notin Z$. Then, by hypothesis (iii), $x-x^n\in N$ for some integer $n>1$. Hence, as we have seen in the proof of Claim 2, $x^m=x^{m+1}x'$ for some positive integer m and some $x'\in\langle x\rangle$. Then, as is readily verified (compare with Lemma 1.3.2 in [2]),

$$x^m=x^me,\quad e^2=e,\quad\text{where}\quad e=x^m(x')^m. \tag{4}$$

By Claim 1, e is a *central* idempotent in the subdirectly irreducible ring R, and hence, using the Peirce-decomposition (see [1; Lemma 9]),

$$e=0\quad\text{or } e \text{ is the identity, 1, of } R. \tag{5}$$

Note that if $1\notin R$, then $e=0$ and hence by (4), $x\in N$, which proves (3). So suppose $1\in R$. Combining (4) and (5), we obtain (3).
Now, if $N\subseteq Z$, then Theorem 1 follows at once from hypothesis (iii) and HERSTEIN's Theorem [4; § X.2, Theorem 2]. So suppose $N\nsubseteq Z$. Then,

$$ab\neq ba\quad\text{for some}\quad a\in N,\quad\text{and some}\quad b\in R;\quad a,b\textit{ fixed}. \tag{6}$$

Hence, $b\notin Z$ and $b\notin N$ (since, by hypothesis (i), N is commutative). Therefore, by (3), $1\in R$ and

$$b^{-1}\text{ exists in }R. \tag{7}$$

Let $\bar{b}=b+N\in R/N$. Since $b\notin Z$, hypothesis (iii) implies that, for some integer $n>1$, $b-b^n\in N$, and hence

(8) $$(\bar{b})^n = \bar{b}, \qquad (n > 1).$$

Thus, by (7) and (8), $(\bar{b})^{n-1}=\bar{1}$. Moreover, note that not both $2b$ and $3b$ commute with a, since $ab\neq ba$. Assume, without loss of generality, that $(2b)a\neq a(2b)$. Then by a similar argument as above, we have $(\overline{2b})^{m-1}=\bar{1}$ for some positive integer $m>1$, and hence $(\overline{2b})^{(m-1)(n-1)}=\bar{1}$. But $(\bar{b})^{(m-1)(n-1)}=\bar{1}$, and so $(2^{(m-1)(n-1)}-1)\bar{1}=0$, where $m>1$, and $n>1$. It follows that the characteristic of R/N is not zero, and therefore, since R is subdirectly irreducible,

(9) $$\text{characteristic of } R = p^k \qquad (p \text{ prime, } k \geqq 1).$$

Hence

(10) $$\text{characteristic of } R/N = p^\alpha \qquad (p \text{ prime, } \alpha \geqq 1).$$

Let $F=\langle\bar{b}\rangle\subseteqq R/N$. In view of (7)—(10), we see that F is a finite commutative ring with identity and with no nonzero nilpotent elements and hence, as is well-known,

$$F = \sum_{i=1}^{t}{}^{\bullet} F_i \qquad \text{(direct sum)},$$

$$F_i = GF(p^{k_i}); \qquad i = 1, \ldots, t, \qquad (p \text{ as in (9)}).$$

Let $\lambda=k_1k_2\ldots k_t$. Then, as is readily verified,

(11) $$(\bar{b})^{p^{\lambda k}} = \bar{b}, \quad \text{and hence} \quad b^{p^{\lambda k}}-b\in N.$$

Since $a\in N$, $b^{p^{\lambda k}}-b\in N$, and N is commutative, we have

(12) $$a \text{ commutes with } b^{p^{\lambda k}}-b.$$

An easy induction (see the proof of Lemma 4 in § X.2 of [4]) shows that if $x, y\in R$, and x commutes with $xy-yx$, then for each positive integer k, we have

$$x^k y - yx^k = kx^{k-1}(xy-yx).$$

In our case, by hypothesis (ii), b commutes with $ab-ba$, and hence

(13) $$ab^{p^{\lambda k}}-b^{p^{\lambda k}}a = p^{\lambda k}b^{p^{\lambda k}-1}(ab-ba) = 0, \quad \text{by (9)}.$$

Combining (12) and (13), we conclude that $ab=ba$, which contradicts (6). This final contradiction proves Theorem 1.

As stated above, the case $N=\{0\}$ of Theorem 1 yields Jacobson's Theorem (quoted in the introduction).

As a further corollary of Theorem 1, we have the following theorem proved in [5].

COROLLARY. *Suppose R is a ring with center Z, and N is the set of nilpotent elements of R. Suppose that* (i) *N is commutative;* (ii) *for all x, y in R, $xy-yx\in Z$;* (iii) *for every x in R, there exists a positive integer $n=n(x)$ and a polynomial $f(\lambda)=f_x(\lambda)$ with integer coefficients such that $x^n=x^{n+1}f(x)$ (both n and $f(\lambda)$ depend on x). Then R is commutative.*

PROOF. A careful examination of the proofs of Claim 1 and Claim 2 above shows that N is an ideal in R. Moreover, since $x^n = x^{n+1}f(x)$, $x^{n-1}(x-x^2f(x))=0$, and hence

$$\begin{aligned}(x-x^2f(x))^n &= (x-x^2f(x))^{n-1}(x-x^2f(x)) \\ &= h(x)x^{n-1}(x-x^2f(x)) = 0.\end{aligned}$$

Thus, $x-x^2f(x)\in N$. Hence, by Corollary 3.5 of STEWART [6], $x-x^{k+1}\in N$ for some $k=k(x)\geqq 1$. The corollary now follows at once from Theorem 1.

We conclude with the following examples which show that Theorem 1 need not be true if we delete any of the hypotheses.

EXAMPLE 1. Let

$$R = \left\{ \left. \begin{pmatrix} a & b & c \\ 0 & a & d \\ 0 & 0 & a \end{pmatrix} \right| a, b, c, d \in GF(3) \right\}.$$

In this example, hypothesis (i) in Theorem 1 fails and, moreover, R is not commutative.

EXAMPLE 2. Let

$$R = \left\{ \left. \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right| 0, 1 \in GF(2) \right\}.$$

Here, hypothesis (ii) in Theorem 1 fails and again R is not commutative.

EXAMPLE 3. The ring of quaternions shows that hypothesis (iii) in Theorem 1 cannot be dropped.

In conclusion, we would like to express our indebtedness and gratitude to the referee Mr. E. KISS for his valuable comments and helpful suggestions.

REFERENCES

[1] HERSTEIN, I. N.: The structure of a certain class of rings, *Amer. J. Math.* **75** (1953), 864—871.

[2] HERSTEIN, I. N.: *Noncommutative rings,* Carus Math. Monographs, No. 15, Math. Assoc. America, 1968.

[3] IKEHATA, S. and TOMINAGA, H.: A commutativity theorem, *Math. Japonica* **24** (1979), 29—30.

[4] JACOBSON, N.: *Structure of rings,* Amer. Math. Soc. Colloq. Publications, 37, 1964.

[5] PUTCHA, M. S. and YAQUB, ADIL: Rings with constraints on nilpotent elements and commutators, *Studia Sci. Math. Hungar.* **12** (1977), 193—197.

[6] STEWART, P. N.: Semi-simple radical classes, *Pacific J. Math.* **32** (1970), 249—259.

(Received August 15, 1979)

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF PETROLEUM AND MINERALS
DHAHRAN
SAUDI ARABIA

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
SANTA BARBARA, CA 93106
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 87—94

APPROXIMATION BY A CLASS OF LINEAR OPERATORS INVOLVING A LOWER TRIANGULAR MATRIX

by

R. N. MOHAPATRA and B. N. SAHNEY

Abstract

The order and class of saturation of linear operators defined by considering Cesàro, Hölder, Nörlund and Riesz means of Fourier series of continuous, 2π periodic functions have been considered by various authors. The object of this paper is to generalize these results by considering a linear operator based on lower-semi-matrix transformation of the Fourier series of a continuous function.

1. Notations and definition

The Fourier series of a continuous, 2π periodic function $f(x)$ is given by

$$\frac{1}{2}a_0+\sum_{k=1}^{\infty}(a_k\cos kx+b_k\sin kx)\equiv\sum_{k=0}^{\infty}A_k(x). \tag{1.1}$$

Let $D=(d_{nk})$ be a lower-triangular infinite matrix. We shall write $\bar{D}=(\bar{d}_{nk})$ where $\bar{d}_{nk}=\sum_{r=k}^{n}d_{nr}$.

Let us write

$$D_n(f;x)=\sum_{r=0}^{n}\bar{d}_{nr}A_r(x). \tag{1.2}$$

Clearly, $D=D_n(f)$ is a linear operator on C^*, the set of continuous functions with period 2π. All norms considered will be supnorms.

We shall say that the class of operators D_n is saturated with order $\varphi(n)$ and belongs to a saturation class $\mathscr{K}$ if there exists a non-increasing function $\varphi(n)$ and a class of functions $\mathscr{K}$ with following properties:

$$\|f(x)-D_n(f;x)\|=o(\varphi(n))\Rightarrow f(x)\quad\text{is constant,} \tag{1.3}$$

$$\|f(x)-D_n(f;x)\|=O(\varphi(n))\Rightarrow f(x)\in\mathscr{K} \tag{1.4}$$

and

$$f(x)\in\mathscr{K}\Rightarrow\|f(x)-D_n(f;x)\|=O(\varphi(n)). \tag{1.5}$$

Throughout the paper $\tilde{f}$, the conjugate function of f, is taken as

$$\tilde{f}(x)=\frac{1}{2\pi}\int_0^{\pi}\{f(x+t)-f(x-t)\}\cot\frac{1}{2}t\,dt$$

AMS (MOS) subject classifications (1980). Primary 41A40.
Key words and phrases. Operators, saturation, summability, matrices.

if the integral converges absolutely for all x and if the integral

$$\int_0^{\pi} |f(x+t)-f(x-t)| \cot \frac{1}{2} t \, dt$$

is an integrable function of x.

2. Introduction

The problem of considering the saturation class of operators involving $(C, 1)$-means of Fourier series has been considered by ZAMANSKI [14] see also ALEXITS [1]. SUNOUCHI and WATARI [13] have considered an analogous problem by taking the (C, α)- mean, Abel means and Riesz means of Fourier series. Some of these results have been extended by other authors (see [12], [9], [10], [11]). Many authors have also considered the general problem of saturation (see [2], [3], [7], [8], [9]). Recently GOEL, HOLLAND et. al. [6] have considered the saturation class of an operator based on the Nörlund mean of a Fourier series. The object of this paper is to obtain the saturation class of the operators $D_n(f; x)$ and unify several of these results.

We prove the following

THEOREM. *Let the matrix* $D=(d_{nk})$ *satisfy*

$$d_{n0} > 0, \; d_{nk} \geqq 0 \quad \text{and} \quad \sum_{k=0}^{n} d_{nk} = 1. \tag{2.1}$$

Then the following hold:

(a) $\|f-D_n(f)\| = o(d_{n0}) \Rightarrow f$ *is a constant.*

(b) *If in addition to* (2.1) $d_{nk}/d_{n0} \to \gamma_k$ *($n \to \infty$, k fixed),*

$$\eta_k = \sum_{i=0}^{k-1} \gamma_i, \quad k/\eta_k = O(1) \quad \text{and} \quad \sum_{k=1}^{\infty} (k+1)|\Delta^2(k/\eta_k)| < \infty, \tag{2.2}$$

then

$$\|f-D_n(f)\| = O(d_{n0}) \Rightarrow \tilde{f} \in \operatorname{Lip} 1.$$

(c) *If in addition to* (2.1)

$$\sum_{k=0}^{n} |d_{nk}-d_{n,k+1}| = O(d_{n0}), \tag{2.3}$$

then

$$\tilde{f} \in \operatorname{Lip} 1 \Rightarrow \|f-D_n(f)\| = O(d_{n0}).$$

3. Five lemmas

We shall need the following lemmas:

LEMMA 1 ([4], Lemma 2). *If* $\alpha \geqq 0$, $p \geqq 0$, $E_n = O(n^{-p})$ *and* $\sum A_n^{\alpha} |\Delta^{\alpha} E_n| < \infty$ *then*

(i) $\sum A_n^{\lambda+p} |\Delta^{\lambda+1} E_n| < \infty$ *for* $-1 \leqq \lambda \leqq \alpha$,

(ii) $A_n^{\lambda+p} \Delta^{\lambda} E_n$ *is of bounded variation for* $0 \leqq \lambda \leqq \alpha$ *and tends to zero as* $n \to \infty$ *except when* $p=0$ *and* $\lambda = 0$.

LEMMA 2. *When* (2.2) *holds, we have*

$$\Delta(n/\eta_n) = O(1/n) \tag{3.1}$$

and

$$\sum_{n=m}^{\infty} (n+1) |\Delta^2(1/\eta_n)| = O(1/m). \tag{3.1'}$$

PROOF. On setting $E_n = n/\eta_n$ in Lemma 1 we have the estimate of (3.1). The estimate (3.1′) follows from

$$\Delta^2 \left[\frac{1}{\eta_n}\right] = \frac{2n/\eta_n}{n(n+1)(n+2)} + \frac{2\Delta(n/\eta_n)}{(n+1)(n+2)} + \frac{\Delta^2(n/\eta_n)}{(n+2)}$$

on appropriate manipulation.

LEMMA 3. *Let the matrix* (d_{nk}) *satisfy* (2.1). *Then*

$$\int_0^{\pi} \left| \sum_{k=0}^{n} (d_{nk} - d_{n,k+1}) \int_t^{\pi} \frac{\sin (k+1)u}{u^2}\, du \right| dt = O\left(\sum_{k=0}^{n} |d_{nk} - d_{n,k+1}| \right). \tag{3.2}$$

PROOF. Since

$$\int_t^{\pi} \frac{\sin (k+1)u}{u^2}\, du = (k+1) \int_{(k+1)t}^{(k+1)\pi} \frac{\sin \theta}{\theta^2}\, d\theta = \begin{cases} O\left[(k+1) \log \dfrac{1}{(k+1)t}\right] & ((k+1)t < 1), \\ O\big(1/(k+1)t^2\big) \quad ((k+1)t > 1), \end{cases}$$

we find that the expression the left side of (3.2) is

$$O\left[\sum_{k=0}^{n} |d_{nk} - d_{n,k+1}| (k+1) \int_0^{1/(k+1)} \log \frac{1}{(k+1)t}\, dt \right] +$$

$$+ O\left[\sum_{k=0}^{n} |d_{nk} - d_{n,k+1}| (k+1)^{-1} \int_{1/(k+1)}^{\pi} t^{-2}\, dt \right] =$$

$$= O\left[\sum_{k=0}^{n} |d_{nk} - d_{n,k+1}| \right].$$

LEMMA 4. *Let* $\sigma_m(x)=\sum_{k=0}^{m}\left\{1-\frac{k}{m+1}\right\}A_k(x)$ *and* $\{\eta_k\}$ *satisfy* (2.2). *Then*

$$\|T_m(x)\| \equiv \left\|\sum_{k=1}^{m}\eta_k\left[1-\frac{k}{m+1}\right]A_k(x)\right\| = O(1) \tag{3.3}$$

implies

$$\|\sigma_m(x)-\varphi(x)\| = O(1/m)$$

where

$$\varphi(x) = \frac{1}{2}a_0+\sum_{k=1}^{\infty}(k+1)T_k(x)\Delta^2(1/\eta_k). \tag{3.4}$$

PROOF. From the definitions of $\sigma_n(x)$ and $T_m(x)$ we have

$$\eta_k A_k(x) = (k+1)T_k(x)-2kT_{k-1}(x)+(k-1)T_{k-2}(x)$$

and

$$\sigma_m(x) = \frac{1}{2}a_0+\frac{1}{m+1}\sum_{k=1}^{m}\frac{m+1-k}{k}\{(k+1)T_k(x)-2kT_{k-1}(x)+(k-1)T_{k-2}(x)\} =$$

$$= \frac{1}{2}a_0+\frac{1}{m+1}\left\{\sum_{k=1}^{m-1}(k+1)T_k(x)\left[\frac{m+1-k}{\eta_k}-\frac{2(m-k)}{\eta_{k+1}}+\frac{m-1-k}{\eta_{k+2}}\right]+\right.$$

$$\left.+\frac{(m+1)T_m(x)}{\eta_m}\right\} =$$

$$= \frac{1}{2}a_0+\sum_{k=1}^{m-1}(k+1)T_k(x)\Delta^2\left[\frac{1}{\eta_k}\right]-\frac{1}{m+1}\sum_{k=1}^{m-1}(k+1)T_k(x)\Delta^2\left[\frac{k}{\eta_k}\right]+\frac{T_m(x)}{\eta_m} =$$

$$= \frac{1}{2}a_0+\sum_{k=1}^{\infty}(k+1)T_k(x)\Delta^2\left[\frac{1}{\eta_k}\right]+O\left[\frac{1}{m}\right],$$

by (3.3) and (2.2).

LEMMA 5 ([1], also see [13], [15], and [11], Theorem 10). *If* $\sigma_n(f;x)=\sigma_n(x)$ *then*

$$\|\sigma_n(f)-f\|=O\left[\frac{1}{n}\right]\Leftrightarrow\tilde{f}\in\mathrm{Lip}\,1.$$

4. Proof of the Theorem

(a) From (1.2) and (1.1)

$$\frac{1}{\pi}\int_{-\pi}^{\pi}D_n(f;x)\cos kx\,dx = \frac{1}{\pi}\sum_{r=0}^{n}\bar{d}_{nr}\int_{-\pi}^{\pi}A_r(x)\cos kx\,dx = \bar{d}_{nk}a_k.$$

Thus

$$a_k-\bar{d}_{nk}a_k = \frac{1}{\pi}\int_{-\pi}^{\pi}\{f(x)-D_n(f;x)\}\cos kx\,dx$$

implies that $|a_k-\bar{d}_{nk}a_k|=O(\|f-D_n(f)\|)=o(d_{n0})$, by the hypothesis. Since, by (2.1), $(1-\bar{d}_{nk})/d_{n0}\geqq 1$, $a_k=0$ for $k\geqq 1$. Similarly $b_k=0$ $(k\geqq 1)$. Hence $f(x)=\frac{1}{2}a_0$ for all x. This proves (a).

(b) Since

$$(1-\bar{d}_{nk})A_k(x)=\frac{1}{\pi}\int_{-\pi}^{\pi}(f(x+t)-D_n(f;x+t))\cos kt\,dt, \tag{4.1}$$

on multiplying both sides by $(1-k/(m+1))$ and summing from 0 to m, we can conclude that

$$\left\|\sum_{k=0}^{m}(1-\bar{d}_{nk})\left[1-\frac{k}{m+1}\right]A_k(x)\right\|=O(d_{n0}) \tag{4.2}$$

uniformly in $n\geqq m>0$. Dividing (4.2) by d_{n0} and replacing 1 by $\sum_{k=0}^{n}d_{nk}$, we make $n\to\infty$. Then we obtain

$$\left\|\sum_{k=1}^{n}\eta_k\left[1-\frac{k}{m+1}\right]A_k(x)\right\|=O(1). \tag{4.3}$$

From Lemma 4 we can infer that $\|\sigma_m(x)-\varphi(x)\|=O\left[\frac{1}{m}\right]$ ($\varphi(x)$ is defined in Lemma 4). But since $\sigma_m(x)\to f(x)$ everywhere by Fejér's theorem, it follows that $f(x)=\varphi(x)$. Now the proof is completed by appealing to Lemma 5.

(c) Let $\tilde{S}_n(\tilde{f}(x))$ be the n-th partial sum of the conjugate series associated with $\tilde{f}(x)$ (see ZYGMUND [15]) we have

$$\tilde{S}_n(\tilde{f}(x))=\frac{1}{2\pi}\int_{0}^{\pi}\{\tilde{f}(x-t)-\tilde{f}(x+t)\}\frac{\cos t/2-\cos\left[n+\frac{1}{2}\right]t}{\sin t/2}\,dt. \tag{4.4}$$

Hence writing $D_n(\tilde{S}_n(\tilde{f}(x)))=\sum_{k=0}^{n}d_{nk}\tilde{S}_k(\tilde{f}(x))$,

$$D_n(\tilde{S}_n(\tilde{f}(x)))=\sum_{k=0}^{n}d_{nk}\frac{1}{2\pi}\int_{0}^{\pi}\{\tilde{f}(x+t)-\tilde{f}(x-t)\}\cot t/2\,dt$$

$$-\sum_{k=0}^{n}d_{nk}\frac{1}{2\pi}\int_{0}^{\pi}\{\tilde{f}(x+t)-\tilde{f}(x-t)\}\frac{\cos\left[k+\frac{1}{2}\right]t}{\sin t/2}\,dt.$$

Since $\tilde{f}\in\mathrm{Lip}\,1\Rightarrow -f+\frac{1}{2}a_0$ is identical to $\tilde{\tilde{f}}$ we have

$$f(x)-D_n(f;x)=\frac{1}{2\pi}\int_{0}^{\pi}\{\tilde{f}(x+t)-\tilde{f}(x-t)\}K_n(t)\,dt, \tag{4.5}$$

where

$$K_n(t) = \sum_{k=0}^{n} d_{nk} \frac{\cos\left[k+\frac{1}{2}\right]t}{\sin t/2}. \tag{4.6}$$

Applying Abel's transformation to the sum on the right of (4.6), we get

$$K_n(t) = \sum_{k=0}^{n} (d_{nk}-d_{n,k+1}) \frac{\sin(k+1)t}{\sin^2 t/2} =$$

$$= \left\{\frac{2}{t^2}+O(1)\right\} \sum_{k=0}^{n} (d_{nk}-d_{n,k+1}) \sin(k+1)t =$$

$$= 2\sum_{k=0}^{n} (d_{nk}-d_{n,k+1}) \frac{\sin(k+1)t}{t^2} + O(d_{n0}),$$

by (2.3). Consequently, $f(x)-D_n(f; x)=I_1+O(d_{n0})$, where

$$I_1 = \frac{1}{\pi}\int_0^{\pi} \{\tilde{f}(x+t)-\tilde{f}(x-t)\} \sum_{k=0}^{n} (d_{nk}-d_{n,k+1}) \frac{\sin(k+1)t}{t^2}\, dt. \tag{4.7}$$

To complete the proof we need to show $I_1=O(d_{n0})$.

Set

$$N_n(t) = \int_t^{\pi} \sum_{k=0}^{n} (d_{nk}-d_{n,k+1}) \frac{\sin(k+1)t}{t^2}\, dt.$$

Hence, by integration by parts

$$I_1 = -\frac{1}{\pi}\int_0^{\pi} \{\tilde{f}^1(x+t)+\tilde{f}^1(x-t)\} N_n(t)\, dt = O\left[\int_0^{\pi} |N_n(t)|\, dt\right]$$

since $\tilde{f}^1=O(1)$ whenever $\tilde{f}\in \operatorname{Lip} 1$. By appealing to Lemma 3 and (2.3) $I_1=O(d_{n0})$.

This completes the proof of the theorem.

REMARK 1. If the matrix $D=(d_{nk})$ is such that instead of the restrictions imposed on the entries of D, we have (2.1) and

$$d_{nk}/d_{n0} \to 1 \quad \text{as} \quad n\to\infty \quad \text{for each fixed } k \tag{4.8}$$

then also the conclusion (b) of the theorem is true as indicated below.

Dividing both sides of (4.2) by d_{n0} and making $n\to\infty$ we have

$$\left\|\sum_{k=0}^{m}\left[1-\frac{k}{m+1}\right](-kA_k(x))\right\| = O(1).$$

One can easily follow the argument in [11, p. 100 and p. 101] and complete the proof.

However, (4.8) may not be satisfied by many standard summability matrices; for example, if

$$d_{n,k} = \begin{cases} \dfrac{p_k}{p_n}, & k = 0, 1, \dots, n, \\ 0, & k > n, \end{cases}$$

when $p_n \geqq 0$ for all n and $p_0 \neq 0$, we see that

$$\frac{d_{nk}}{d_{n0}} = \frac{p_k}{p_0}.$$

Hence when $n \to \infty$, d_{nk}/d_{n0} need not converge to 1, as $n \to \infty$ (for example $p_k = 1/(k+1)$, $k=0, 1, \dots$). Thus (4.8) is more restrictive than (2.2).

Remark 2. Although the condition (2.3) is satisfied by many infinite matrices, it fails to hold when (d_{nk}) is the matrix associated with the Harmonic Summability, since in this case

$$d_{nk} = \begin{cases} \dfrac{1}{(n-k+1)\log n}, & k = 0, 1, \dots, n; \\ 0, & k > n. \end{cases}$$

Hence our result cannot be used to obtain the saturation class of an operator associated with Harmonic Summability.

Remark 3. By specializing the matrix $D=(d_{nk})$ to be the Cesàro, Nörlund or discrete Riesz matrices (see [7] for definitions and other references) we can deduce the results proved in [6], [8], [13], [14] and a result concerning an operator associated with the discrete Riesz method.

For some general results involving an infinite matrix that is not necessarily lower triangular see [10].

REFERENCES

[1] Alexits, G.: Sur l'ordre de grandeur de l'approximation d'une fonction par les moyennes de sa série de Fourier, *Mat. Fiz. Lapok* **48** (1941), 410—422.

[2] Aljančič, S.: Classes de saturation des procédés de sommation de Hölder et de Riesz, *C. R. Acad. Sci. Paris* **246** (1958), 2567—2569.

[3] Aljančič, S.: Approximation of continuous functions by typical means of their Fourier series, *Proc. Amer. Math. Soc.* **12** (1961), 681—688.

[4] Bosanquet, L. S.: Note on Bohr—Hardy Theorem, *J. London Math. Soc.* **17** (1942), 166—173.

[5] Favard, J.: Sur la saturation des procédés de sommation, *J. de Mathemat.* **36** (1957), 359—372.

[6] Goel, D. S., Holland, A. S. B., Nasim, C. and Sahney, B. N.: Best approximation by a saturation class of polynomial operators, *Pacific J. Math.* **55** (1974), 149—155.

[7] Hardy, G. H.: *Divergent Series*, Oxford, 1949.

[8] Harsiladze, F.: Classes of saturation for certain methods of summability, *Dokl. Akad. Nauk SSSR* **122** (1958), 353—355 (in Russian).

[9] Ikeno, K. and Suzuki, Y.: Some remarks of saturation problem in the local approximation, *Tôhoku Math. J.* **20** (1968), 214—233.

[10] Kuttner, B., Mohapatra, R. N. and Sahney, B. N.: Saturation results for a class of linear operators (to appear).

[11] Lorentz, G. G.: *Approximation of Functions,* Holt, Rinehart and Winston, 1966.

[12] Sunouchi, G.: On the class of saturation in the theory of approximation I, II, III, *Tôhoku Math. J.* **12** (1960), 339—344 and *ibid.* **13** (1961), 112—118 and 320—328.

[13] Sunouchi, G. and Watari, C.: On determination of the class of saturation in the theory of approximation of functions, *Proc. Japan Acad.* **34** (1958), 477—481.

[14] Zamanski, M.: Classes de saturation de certaines procédés d'approximation des séries de Fourier des fonctions continues, *Ann. École Normale Sup.* **66** (1949), 19—93.

[15] Zygmund, A.: *Trigonometric Series,* I and II, Cambridge, 1959.

(Received August 16, 1979)

AMERICAN UNIVERSITY OF BEIRUT
BEIRUT
LEBANON

DEPARTMENT OF MATHEMATICS AND STATISTICS
THE UNIVERSITY OF CALGARY
CALGARY, ALBERTA
T2N 1N4
CANADA

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 95—98

ON FAITHFUL IRREDUCIBLE REPRESENTATIONS OF FINITE GROUPS

by

P. P. PÁLFY

1. Introduction

The problem that which groups have a faithful irreducible representation goes back to W. BURNSIDE [1, p. 476—478]. In course of years several authors were interested in the question, e.g. K. SHODA [8], L. WEISNER [9], R. KOCHENDÖRFFER [4], W. GASCHÜTZ [2]. The purpose of this note is to give a formula expressing the square-sum of degrees of faithful irreducible representations of a finite group G over a splitting field F, such that $\operatorname{char} F \nmid |G|$. In this expression there occur some structure constants of the socle of G. The proof hardly makes use of representation theory, its main tool is the Moebius function of the lattice of normal subgroups of G. As a corollary we obtain a condition for G to have a faithful irreducible representation, which is in fact a special case of the Kochendörffer's one [4, Satz 5] for fields satisfying the above conditions.

2. The socle

First we summarize some properties of the socle of G. Let $\mathcal{S}$ be the set of all minimal normal subgroups of G, and $\mathcal{S}_1, \ldots, \mathcal{S}_r, \mathcal{S}_{r+1}, \ldots, \mathcal{S}_{r+s}$ be the G-isomorphism classes of $\mathcal{S}$, where $\mathcal{S}_i$ consists of abelian or non-abelian subgroups for $1 \leqq i \leqq r$, resp., $r+1 \leqq i \leqq r+s$. Set $S=\langle M \mid M \in \mathcal{S}\rangle$ and $S_i=\langle M \mid M \in \mathcal{S}_i\rangle$ for $1 \leqq i \leqq r+s$. At last choose $A_i \in \mathcal{S}_i$ for $1 \leqq i \leqq r$. The following facts are more or less well-known.

LEMMA 1. *Let N be a normal subgroup of G.*

(i) *If $N \leqq S$, then $N=(N \cap S_1) \times (N \cap S_2) \times \ldots \times (N \cap S_{r+s})$.*

(ii) *If $N \leqq S_i$ for an i, $1 \leqq i \leqq r$, then N is the direct product of some minimal normal subgroups of G G-isomorphic to A_i. Let us denote by t_i the number of such direct factors of S_i. Then $N \mapsto \operatorname{Hom}_G(A_i, N)$ is a lattice isomorphism between $\{N \mid N \lhd G, N \leqq S_i\}$ and the subspace-lattice of the t_i-dimensional vector space $\operatorname{Hom}_G(A_i, S_i)$ over $\operatorname{End}_G(A_i)$.*

(iii) *If $N \leqq S_i$ for an i, $r+1 \leqq i \leqq r+s$, then either $N=S_i$ or $N=1$.*

Concerning (i), (iii) and the first assertion of (ii) we refer to the paper of R. REMAK [6], particularly to his Theorems 3a, 5 and 1, respectively. Therefore we sketch only the proof of the second assertion of (ii). One can see immediately that $\varphi_1, \ldots, \varphi_k \in \operatorname{Hom}_G(A_i, S_i)$ are linearly independent over $\operatorname{End}_G(A_i)$ iff the subgroup sum $\sum_{j=1}^{k} \operatorname{Im} \varphi_j$ is direct and φ_j's are not zero. Hence we have

$$\dim_{\operatorname{End}_G(A_i)} \operatorname{Hom}_G(A_i, S_i)=t_i.$$

Moreover, let W be a subspace of $\mathrm{Hom}_G(A_i, S_i)$ and $\varphi_1, \ldots, \varphi_k$ be a basis of W. Then it follows also from the preceding remark that $\sum_{\varphi\in W} \mathrm{Im}\,\varphi = \sum_{j=1}^{k} \mathrm{Im}\,\varphi_j$, where the sum on the right-hand side is direct. Applying this, one can easily check that $W \mapsto \sum_{\varphi\in W} \mathrm{Im}\,\varphi$ is the inverse of the mapping $N \mapsto \mathrm{Hom}_G(A_i, N)$. Since these mappings are inclusion preserving, they are lattice isomorphisms, as we claimed.

3. The Moebius function

The Moebius function of the lattice of normal subgroups of G is defined by

$$\sum_{\substack{N \lhd G\\ N\leqq K}} \mu(N) = \begin{cases} 1 & \text{if } \; K=1\\ 0 & \text{if } \; K\neq 1,\, K\lhd G.\end{cases} \tag{1}$$

Making use of some results from the general theory of Moebius functions (see G.-C. ROTA [7]) we can determine μ as the following lemma shows.

LEMMA 2. *Let N be a normal subgroup of G.*

(i) *If $N\not\leqq S$, then $\mu(N)=0$.*

(ii) *If $N\leqq S$, then $\mu(N)=\prod_{i=1}^{r+s} \mu(N\cap S_i)$.*

(iii) *If $N\leqq S_i$ for an i, $1\leqq i\leqq r$, and N is the direct product of k minimal normal subgroups of G, then*

$$\mu(N) = (-1)^k |\mathrm{End}_G(A_i)|^{\binom{k}{2}}.$$

(iv) *If $N=S_i$ for an i, $r+1\leqq i\leqq r+s$, then $\mu(N)=-1$.*

PROOF. (i) follows from a general result of P. HALL [3, the dual of Theorem 2.3, see p. 140]. The assertion (ii) is a simple consequence of our Lemma 1 (i) and [7, Proposition 3.5]. Now (iii) follows from Lemma 1 (ii) and [7, Example 2, p. 351]. At last, for $r+1\leqq i\leqq r+s$ by Lemma 1 (iii) S_i is an atom in the lattice of normal subgroups of G, thus obviously $\mu(S_i)=-1$, so (iv) holds, too.

4. Results

THEOREM. *Let G be a finite group, F be a splitting field for G with* $\mathrm{char}\,F\nmid |G|$. *Let us denote by* $\mathrm{Irr}(G)$ *the set of pairwise non-equivalent irreducible representations of G over F. Moreover let r, s, S_i, A_i, t_i denote the same as in Lemma* 1 *and set $q_i= |\mathrm{End}_G(A_i)|$ for $1\leqq i\leqq r$. Then we have*

$$\sum_{\substack{\varphi\in \mathrm{Irr}(G)\\ \varphi \text{ faithful}}} (\deg\varphi)^2 = |G| \prod_{i=1}^{r} \prod_{k=0}^{t_i-1} \left(1-\frac{q_i^k}{|A_i|}\right) \times \prod_{i=r+1}^{r+s} \left(1-\frac{1}{|S_i|}\right). \tag{2}$$

PROOF. First of all let us remark that under our assumptions on F

$$\sum_{\substack{\varphi\in \mathrm{Irr}(G)\\ \mathrm{Ker}\,\varphi\geqq N}} (\deg\varphi)^2 = |G/N| \tag{3}$$

holds for any $N\lhd G$. Now consecutive applications of (1), (3) and Lemmas 1, 2 yield:

$$\sum_{\substack{\varphi\in \mathrm{Irr}(G)\\ \mathrm{Ker}\,\varphi=1}} (\deg\varphi)^2 = \sum_{K\lhd G}\Big(\sum_{\substack{N\lhd G\\ N\leqq K}}\mu(N)\Big)\sum_{\substack{\varphi\in \mathrm{Irr}(G)\\ \mathrm{Ker}\,\varphi=K}} (\deg\varphi)^2 =$$

$$= \sum_{N\lhd G}\mu(N)\sum_{\substack{K\lhd G\\ N\leqq K}}\sum_{\substack{\varphi\in \mathrm{Irr}(G)\\ \mathrm{Ker}\,\varphi=K}} (\deg\varphi)^2 = \sum_{N\lhd G}\mu(N)\sum_{\substack{\varphi\in \mathrm{Irr}(G)\\ \mathrm{Ker}\,\varphi\geqq N}} (\deg\varphi)^2 =$$

$$= \sum_{N\lhd G}\mu(N)|G/N| = |G|\sum_{\substack{N\lhd G\\ N\leqq S}}\frac{\mu(N)}{|N|} =$$

$$= |G|\sum_{\substack{N_1\lhd G\\ N_1\leqq S_1}}\cdots\sum_{\substack{N_{r+s}\lhd G\\ N_{r+s}\leqq S_{r+s}}}\prod_{i=1}^{r+s}\frac{\mu(N_i)}{|N_i|} = |G|\prod_{i=1}^{r+s}\Big(\sum_{\substack{N_i\lhd G\\ N_i\leqq S_i}}\frac{\mu(N_i)}{|N_i|}\Big) =$$

$$= |G|\prod_{i=1}^{r}\Big(\sum_{j=0}^{t_i}\begin{bmatrix} t_i\\ j\end{bmatrix}_{q_i}\frac{(-1)^j q_i^{\binom{j}{2}}}{|A_i|^j}\Big)\times\prod_{i=r+1}^{r+s}\Big(1-\frac{1}{|S_i|}\Big),$$

where in the latter expression the Gaussian coefficient

$$\begin{bmatrix} t_i\\ j\end{bmatrix}_{q_i} = \frac{(q_i^{t_i}-1)(q_i^{t_i}-q_i)\dots(q_i^{t_i}-q_i^{j-1})}{(q_i^j-1)(q_i^j-q_i)\dots(q_i^j-q_i^{j-1})}$$

stands for the number of normal subgroups N satisfying $N\leqq S_i$ and $|N|=|A_i|^j$ (cf. Lemma 1 (ii)). In order to prove (2) it remains only to check that

$$\prod_{k=0}^{t-1}\Big(1-\frac{q^k}{a}\Big) = \sum_{j=0}^{t}\begin{bmatrix} t\\ j\end{bmatrix}_q\frac{(-1)^j q^{\binom{j}{2}}}{a^j} \tag{4}$$

holds for arbitrary natural numbers $t\geqq 0$, $a>0$, $q>1$. Proceeding by induction on t, we get

$$\prod_{k=0}^{t}\Big(1-\frac{q^k}{a}\Big) = \Big(1-\frac{q^t}{a}\Big)\sum_{j=0}^{t}\begin{bmatrix} t\\ j\end{bmatrix}_q\frac{(-1)^j q^{\binom{j}{2}}}{a^j} =$$

$$= \sum_{j=0}^{t+1}\Big(\begin{bmatrix} t\\ j\end{bmatrix}_q(-1)^j q^{\binom{j}{2}} - q^t\begin{bmatrix} t\\ j-1\end{bmatrix}_q(-1)^{j-1}q^{\binom{j-1}{2}}\Big)\frac{1}{a^j} =$$

$$= \sum_{j=0}^{t+1}\Big(\begin{bmatrix} t\\ j\end{bmatrix}_q + q^{t-j+1}\begin{bmatrix} t\\ j-1\end{bmatrix}_q\Big)\frac{(-1)^j q^{\binom{j}{2}}}{a^j}.$$

Moreover, we have

$$\begin{bmatrix} t \\ j \end{bmatrix}_q + q^{t-j+1} \begin{bmatrix} t \\ j-1 \end{bmatrix}_q = \begin{bmatrix} t+1 \\ j \end{bmatrix}_q$$

(see, e.g., [5, (3.1), p. 104]), which completes the proof.

Examining when the two sides of (2) be positive we obtain a condition for G to have a faithful irreducible representation. (The notations are the same as above.)

COROLLARY. G possesses a faithful irreducible representation over F iff for all i, $1 \leqq i \leqq r$ $q_i^{t_i} \leqq |A_i|$, i.e., $t_i \leqq \dim_{\mathrm{End}_G(A_i)} A_i$ holds.

REFERENCES

[1] BURNSIDE, W.: *Theory of groups of finite order,* 2nd ed., Cambridge, 1911.
[2] GASCHÜTZ, W.: Endliche Gruppen mit treuen absolut irreduziblen Darstellungen, *Math. Nachr.* **12** (1954), 253—255.
[3] HALL, P.: The Eulerian functions of a group, *Quart. J. Math. Oxford* **7** (1936), 134—151.
[4] KOCHENDÖRFFER, R.: Über treue irreduzible Darstellungen endlicher Gruppen, *Math. Nachr.* **1** (1948), 25—39.
[5] PÓLYA, G. and ALEXANDERSON, G. L.: Gaussian binomial coefficients, *Elemente Math.* **26** (1971), 102—109.
[6] REMAK, R.: Über minimale invariante Untergruppen in der Theorie der endlichen Gruppen, *J. Reine Angew. Math.* **162** (1930), 1—16.
[7] ROTA, G.-C.: On the foundations of combinatorial theory I. Theory of Möbius functions, *Z. Wahrscheinlichkeitstheorie* **2** (1964), 340—368.
[8] SHODA, K.: Bemerkungen über vollständig reduzible Gruppen, *J. Fac. Sci. Univ. Tokyo* **2** (1931), 203—209.
[9] WEISNER, L.: Condition that a finite group be multiply isomorphic with each of its irreducible representations, *Amer. J. Math.* **61** (1939), 709—712.

(Received September 26, 1979)

MTA MATEMATIKAI KUTATÓ INTÉZETE
REÁLTANODA U. 13—15
H—1053 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 99—104

AN OPTIMAL STOPPING RULE FOR s_n/n RELATED TO MARTINGALES

by

GÁBOR HANÁK

Introduction

Let $y_1, y_2, \dots$ be real valued random variables on the probability space $(\Omega, \mathcal{A}, \mathsf{P})$, $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots \subset \mathcal{A}$ be a sequence of sub-σ-algebras such that y_n is $\mathcal{F}_n$-measurable for $n=1, 2, \dots$. Then $\{y_n, \mathcal{F}_n\}_{n=1}^{\infty}$ is called an adapted sequence. Let $x_n = f_n(y_1, \dots, y_n)$ where f_n is Borel-measurable for $n=1, 2, \dots$. Assume that x_n is finitely integrable for $n=1, 2, \dots$. Denote $\bar{C} = \{t\colon t$ is a stopping rule with respect to $\{\mathcal{F}_n\}_{n=1}^{\infty}\}$, $C = \{t \in \bar{C}\colon \mathsf{P}(t<\infty)=1\}$.

We are interested in the following question: Is there a $t \in \bar{C}$ ($t \in C$) such that $\mathsf{E}x_t \geqq \mathsf{E}x_r$ for each $r \in \bar{C}$ ($r \in C$). (Define $x_r I(r=\infty) = \limsup_{n\to\infty} x_n I(r=\infty)$ if $r \in \bar{C}$.) If there exists such a t then it is called an optimal stopping rule in $\bar{C}$ (in C) with respect to the sequence $\{x_n, \mathcal{F}_n\}_{n=1}^{\infty}$.

The case when $y_1, y_2, \dots$ are iid, $\mathcal{F}_n = \sigma\{y_1, \dots, y_n\}$ and $f_n(y_1, \dots, y_n) = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{s_n}{n}$ for $n=1, 2, \dots$ has been investigated in several papers. The problem was originally formulated by CHOW and ROBBINS [1] who considered the sequence $\{y_n\}_{n=1}^{\infty}$ as the result of coin tossing and proved the existence of an optimal stopping rule. This result was generalized by DVORETZKY [3] who assumed y_n to be iid nondegenerate r.v.'s with mean zero and finite variance and proved the existence of a minimal optimal stopping rule in C and gave a description of the structure of that stopping rule. This latter result was sharpened by SHEPP [8]. The structure of all optimal stopping rule in $\bar{C}$ (in C) was investigated by KLASS [4]. SIEGMUND, SIMONS and FEDER [9] found an optimal stopping rule in the following case: $\mathsf{E}y_1 = 0$, $\mathsf{E}|y_1|^{\max(2, b)} < \infty$ and $f_n(y_1, \dots, y_n) = n^{-a}|s_n|^b$ or $= c_n|s_n|^b$ $n=1, 2, \dots$ where $2a > b > 0$ and $\{c_n\}_{n=1}^{\infty}$ is a certain real sequence. The case $f_n(y_1, \dots, y_n) = c_n s_n$ was also studied by TEICHER and WOLFOWITZ [10] and KLASS [5].

RUIZ-MONCAYO [7] dealt not only with the iid case but also with the case of a stationary Markov chain with countable state space.

Now let $\{y_n, \mathcal{F}_n\}_{n=1}^{\infty}$ be a square-integrable martingale difference sequence, $s_n = \sum_{i=1}^{n} y_i$, $A_n = \sum_{i=1}^{n} \mathsf{E}(y_i^2 | \mathcal{F}_{i-1})$ $n=1, 2, \dots$ ($\mathcal{F}_0 = \{\emptyset, \Omega\}$) and assume for sake of simplicity that $\mathsf{E}y_1 = 0$ and $\mathsf{E}y_1^2 \neq 0$. Note that $\{A_n\}_{n=1}^{\infty}$ is the increasing process associated with the submartingale $\{s_n^2, \mathcal{F}_n\}_{n=1}^{\infty}$ by the Doob-decomposition (see e.g. NEVEU [6], p. 148).

We investigate the sequence $x_n = \dfrac{e(s_n)}{f(A_n)}$ where e and f are certain Borel-measurable real functions. In the iid case the denominator n stands for the time gone.

So it is reasonable to take a function of A_n instead of a function of n as the denominator of x_n in the martingale case because, essentially, $\{A_n\}_{n=1}^{\infty}$ measures the progress of time of the process.

Let $g_n = \operatorname{ess\,sup}\{\mathsf{E}(x_t|\mathcal{F}_n)\colon t\in C,\ t\geqq n\}$ $n=1,2,\dots$, $s=\inf\{n\geqq 1\colon x_n=g_n\}$ and $V=\sup\{\mathsf{E}x_t\colon t\in C\}$. g_n can be interpreted as the greatest expected value of the stopped martingale if the past is given and the martingale is not stopped before n. Thus it is natural to expect that s is optimal in $\bar{C}$ (resp. in C). Assuming certain conditions Theorem 1 (resp. Theorem 2) proves this expectation.

Results

Let $e(x)=x^2$ $(x\in\mathbf{R})$ and let $f\colon \mathbf{R}^+\to\mathbf{R}^+$ be a function with the following properties:

$$\text{(i)}\qquad f' \text{ exists and } f'\geqq 0,$$

$$\text{(ii)}\qquad \int_0^{\infty}\frac{1}{1+f(t)}\,dt<\infty.$$

(E.g., $f(x)=x^{\alpha}$ or $f(x)=(x+1)(\log(x+1))^{\alpha}$ with $\alpha>1$ in both cases will do.) Set

$$F(x)=\int_x^{\infty}\frac{1}{f(t)}\,dt.$$

Conditions (i) and (ii) imply that $F(x)$ is well-defined for each positive x.

The first fundamental result is the following

LEMMA. *Let $\{z_n,\mathcal{F}_n\}_{n=1}^{\infty}$ be an adapted sequence in $L_2(\Omega,\mathcal{A},\mathsf{P})$, let $\{s_n,\mathcal{F}_n\}_{n=1}^{\infty}$ be the martingale introduced above and let $f\colon\mathbf{R}^+\to\mathbf{R}^+$ be a function fulfilling conditions* (i) *and* (ii). *Then for every $n\in\mathbf{N}$ and every $t\in\bar{C}$ for which $t\geqq n$ a.s. we have*

$$\text{(1)}\qquad \mathsf{E}\left(\frac{(z_n+s_t-s_n)^2}{f(A_t)}\middle|\mathcal{F}_n\right)\leqq\frac{z_n^2}{f(A_n)}+F(A_n)\quad a.s.$$

PROOF. Let

$$h_k=\frac{(z_n+s_k-s_n)^2}{f(A_k)}+F(A_k),\qquad k=n,n+1,\dots.$$

First prove that $\{h_k,\mathcal{F}_k\}_{k=n}^{\infty}$ is a supermartingale. For each $k>n$ we have

$$\mathsf{E}(h_k|\mathcal{F}_{k-1}):=\frac{(z_n+s_{k-1}-s_n)^2+\mathsf{E}(y_k^2|\mathcal{F}_{k-1})}{f(A_{k-1}+\mathsf{E}(y_k^2|\mathcal{F}_{k-1}))}+F\big(A_{k-1}+\mathsf{E}(y_k^2|\mathcal{F}_{k-1})\big)\stackrel{\text{def}}{=}$$

$$\stackrel{\text{def}}{=}\frac{c+u}{f(d+u)}+F(d+u)\stackrel{\text{def}}{=}G(u)\quad\text{a.s.,}$$

where

$$c=(z_n+s_{k-1}-s_n)^2,$$

$$d=A_{k-1}$$

and

$$u = \mathsf{E}(y_k^2|\mathcal{F}_{k-1}).$$

Then by (i) G: $\mathbf{R}^+ \cup \{0\} \to \mathbf{R}^+$, G' exists, $G' \leqq 0$ and $G(0) \geqq G(u)$ $\forall u \geqq 0$ so we have the supermartingale property, i.e.,

$$\mathsf{E}(h_k|\mathcal{F}_n) \leqq h_n \quad \text{a.s.} \quad \forall k \geqq n. \tag{2}$$

Since $\{h_k, \mathcal{F}_k\}_{k=n}^{\infty}$ is a nonnegative supermartingale and $t \geqq n$, by a well-known result (see e.g. CHOW et al. [2] p. 78) we can change k to t in (2) and the inequality remains true. □

The Lemma enables us to estimate g_n. Assume $\mathsf{P}(A_n \to \infty) = 1$, in this case the Lemma has a few interesting consequences.

Set $\{z_n\}_{n=1}^{\infty} = \{s_n\}_{n=1}^{\infty}$, $n=1$ then $h_k = \frac{s_k^2}{f(A_k)} + F(A_k)$, $k=1, 2, \dots$, is a nonnegative supermartingale. Since $F(A_k)$ converges to zero a.s. $\frac{s_k^2}{f(A_k)}$ converges a.s. Furthermore $F(A_k)$ converges to zero in L_1, too, because it is dominated by a positive constant in the following way: $F(A_k) \leqq F(A_1) = F(\mathsf{E}s_1^2)$. Consequently, if we show $\mathsf{E}\frac{s_k^2}{f(A_k)} \to 0$ as $k \to \infty$ we obtain that h_k and $\frac{s_k^2}{f(A_k)}$ converge to zero a.s. and in L_1, too. Since

$$\mathsf{E}\frac{s_k^2}{f(A_k)} \leqq 2\mathsf{E}\frac{(s_k - s_n)^2}{f(A_k)} + 2\mathsf{E}\frac{s_n^2}{f(A_k)} \leqq 2\mathsf{E}(F(A_n)) + 2\mathsf{E}\frac{s_n^2}{f(A_k)} \tag{3}$$

(the second inequality is a consequence of the Lemma with $\{z_n\}_{n=1}^{\infty} = \{0\}_{n=1}^{\infty}$ and $t \equiv k$) it is sufficient to show that $\mathsf{E}\frac{s_n^2}{f(A_k)} \to 0$ as $k \to \infty$. This last convergence is obviously true by Lebesgue's theorem since $\frac{s_n^2}{f(A_k)} \to 0$ a.s. as $k \to \infty$ and $\frac{s_n^2}{f(A_k)} \leqq \frac{s_n^2}{f(A_1)} = \frac{s_n^2}{f(\mathsf{E}s_1^2)} \in L_1$. Since $g_n \leqq \frac{s_n^2}{f(A_n)} + F(A_n)$ a.s. for each n by the Lemma we have the following

COROLLARY of the Lemma. Assume the conditions of the Lemma and in addition $\mathsf{P}(A_n \to \infty) = 1$. Then g_n and $\frac{s_n^2}{f(A_n)}$ converge to zero a.s. and in L_1, too as $n \to \infty$.

THEOREM 1. *Let $\{y_n, \mathcal{F}_n\}_{n=1}^{\infty}$ be a square-integrable martingale difference sequence, $s_n = \sum_{i=1}^{n} y_i$, $A_n = \sum_{i=1}^{n} \mathsf{E}(y_i^2|\mathcal{F}_{i-1})$, $n=1, 2, \dots$, $(\mathcal{F}_0 = \{\emptyset, \Omega\})$, f: $\mathbf{R}^+ \to \mathbf{R}^+$ be a function fulfilling conditions* (i) *and* (ii) *and $x_n = \frac{s_n^2}{f(A_n)}$. Assume that $\mathsf{E}y_1 = 0$, $\mathsf{E}y_1^2 \neq 0$ and $\mathsf{P}(A_n \to \infty) = 1$. Then s is an optimal stopping rule in $\bar{C}$ with respect to the sequence $\{x_n, \mathcal{F}_n\}_{n=1}^{\infty}$.*

PROOF. It is known (cf. [2], Theorem 4.10, p. 82) that in the case $V<\infty$ s is optimal in $\bar{C}$ if and only if

$$\int_{\{s<\infty\}} x_s^+ < \infty \tag{4}$$

and

$$\{I(s>n)\, g_n^+\}_{n=1}^\infty \tag{5}$$

is uniformly integrable. The Lemma implies (4) and $V<\infty$ simultaneously setting $\{z_n\}_{n=1}^\infty=\{s_n\}_{n=1}^\infty$, $n=1$ and taking expectation in (1), while (5) is a simple consequence of $\mathsf{E}g_n\to 0$. □

We need a simple preliminary proposition before stating Theorem 2.

PROPOSITION. *Let $\{x_n, \mathcal{F}_n\}_{n=1}^\infty$ be an adapted sequence and assume that s is optimal in $\bar{C}$ with respect to $\{x_n, \mathcal{F}_n\}_{n=1}^\infty$. Assume moreover that there exists a sequence $\{B_n\}_{n=1}^\infty\subset\mathcal{A}$ such that*

$$B_n\in\mathcal{F}_n \quad \forall n\in\mathbf{N}, \tag{6}$$

$$I(s=\infty)\leqq I\left(\bigcup_{n=1}^\infty B_n\right) \quad \text{a.s.} \tag{7}$$

and

$$\mathsf{E}(x_s|\mathcal{F}_n) < x_n \quad \text{a.s. on } B_n \quad \forall n\in\mathbf{N}. \tag{8}$$

Then s is optimal in C with respect to the sequence $\{x_n, \mathcal{F}_n\}_{n=1}^\infty$.

PROOF. It is obviously enough to show that $\mathsf{P}(s=\infty)=0$. Suppose the contrary and let $t=\inf\{n\geqq 1\colon B_n\}$, $r=t\wedge s$. Then $r\in\bar{C}$ and

$$\mathsf{E}x_s=\int_{\{t<s\}} x_s+\int_{\{t\geqq s\}} x_s=\sum_{n=1}^\infty\int_{\{t=n<s\}} x_s+\int_{\{t\geqq s\}} x_s<$$

$$<\sum_{n=1}^\infty\int_{\{t=n<s\}} x_n+\int_{\{t\geqq s\}} x_s=\int_{\{t<s\}} x_t+\int_{\{t\geqq s\}} x_s=\mathsf{E}x_r$$

which is a contradiction. □

THEOREM 2. *Assume the conditions of Theorem 1 and furthermore*

$$\mathsf{P}\left(\limsup_{n\to\infty}\frac{s_n^2}{2f(A_n)\cdot F(A_n)}>1+\delta\right)=1 \tag{9}$$

for some $\delta>0$. Then s is an optimal stopping rule in C with respect to the sequence $\{x_n, \mathcal{F}_n\}_{n=1}^\infty$.

PROOF. According to Theorem 1, s is optimal in $\bar{C}$ with respect to $\{x_n, \mathcal{F}_n\}_{n=1}^\infty$. Let

$$B_n=\left\{s>n,\ \mathsf{P}(s<\infty|\mathcal{F}_n)<\frac{\delta}{2(2+\boldsymbol{\delta})},\ \frac{s_n^2}{f(A_n)F(A_n)}>2+\delta\right\},\quad n=1,2,\ldots \tag{10}$$

then (6) is obviously true. Condition (9) and Lévy's convergence theorem imply (7) thus by the Proposition we only have to prove (8). Now, on B_n we have

$$\mathsf{E}\left(\frac{s_s^2}{f(A_s)}\middle|\mathcal{F}_n\right) \leqq 2\mathsf{E}\left(\frac{s_n^2}{f(A_s)}\middle|\mathcal{F}_n\right)+2\cdot\mathsf{E}\left(\frac{(s_s-s_n)^2}{f(A_s)}\middle|\mathcal{F}_n\right) \leqq$$

$$\leqq 2\frac{s_n^2}{f(A_n)}\mathsf{P}(s<\infty|\mathcal{F}_n)+2F(A_n) < 2\frac{s_n^2}{f(A_n)}\frac{\delta}{2(2+\delta)}+2\frac{s_n^2}{f(A_n)}\frac{1}{2+\delta} = \frac{s_n^2}{f(A_n)}$$

a.s.

(The second inequality is an application of the Lemma with $\{z_n\}_{n=1}^{\infty}=\{0\}_{n=1}^{\infty}$.) □

Corollary of Theorem 2. Assume the conditions of Theorem 1 and furthermore

$$\mathsf{P}\left(\limsup_{n\to\infty}\frac{s_n}{\sqrt{2A_n\log\log A_n}}=1\right)=1 \tag{11}$$

(Law of the Iterated Logarithm) and that there exists a function g: $\mathbf{R}^+\to\mathbf{R}^+$ such that

(i) g' exists and $g'\geqq -1$,

(12) (ii) $\limsup_{t\to\infty}\dfrac{g(t)}{t\log\log t}<1$,

(iii) $f(x)=cg(x)\exp\{J(x)\}$

for some $c>0$, where $J'=1/g$. Then s is an optimal stopping rule in C with respect to the sequence $\{x_n, \mathcal{F}_n\}_{n=1}^{\infty}$.

Proof. Conditions (11), (12) and $\mathsf{P}(A_n\to\infty)=1$ clearly imply (9). □

Remark 1. If we define f by (12iii) then (12i) and (12ii) imply (i) and (ii), respectively.

Example. Let $c>0$ be arbitrary and $g(x)=x/c$. Conditions (12i) and (12ii) are obviously true, thus, by Theorem 2, the Law of the Iterated Logarithm and $\mathsf{P}(A_n\to\infty)=1$ imply that s is optimal in C with respect to the sequence $\left\{\dfrac{s_n^2}{A_n^{1+c}}, \mathcal{F}_n\right\}_{n=1}^{\infty}$.

Remark 2. The corresponding results can be obtained for the sequences $\left\{\dfrac{s_n}{\sqrt{f(A_n)}}, \mathcal{F}_n\right\}_{n=1}^{\infty}$ and $\left\{\dfrac{(s_n^+)^2}{f(A_n)}, \mathcal{F}_n\right\}_{n=1}^{\infty}$ with a slight modification of the proofs.

Remark 3. Let $p\in(0, 1)$ be arbitrary and $R(t)=t^p$ $(t\geqq 0)$, then R is a monotone increasing and concave function. The following result is well-known (see e.g. [2], p. 90): let $\{u_n, \mathcal{F}_n\}_{n=1}^{\infty}$ be an arbitrary adapted sequence for which $V<\infty$. Denote the corresponding objects for the sequence $\{R(u_n), \mathcal{F}_n\}_{n=1}^{\infty}$ by writing a subscript R. If s is optimal in $\bar{C}$ (in C) with respect to $\{u_n, \mathcal{F}_n\}_{n=1}^{\infty}$ then $V_R<\infty$ and s_R is also optimal in $\bar{C}_R$ (in C_R) with respect to $\{R(u_n), \mathcal{F}_n\}_{n=1}^{\infty}$. In this way the optimality of

s in $\bar{C}$ (in C) with respect to $\left\{\frac{s_n^2}{A_n^{1+c}}, \mathscr{F}_n\right\}_{n=1}^{\infty}$ ($c>0$) implies the optimality of $s_{a,b}$ in $\bar{C}_{a,b}$ (in $C_{a,b}$) with respect to $\left\{\frac{|s_n|^b}{A_n^a}, \mathscr{F}_n\right\}_{n=1}^{\infty}$ ($2a>b>0$, $b\leqq 2$).

In the iid case ($\mathsf{E}y_1^2=\sigma^2<\infty$) $A_n=\sigma^2 n$ and the Law of the Iterated Logarithm is true. Thus Theorem 1 (Theorem 2) tells us that s is always optimal in $\bar{C}$ (in C) with respect to the sequence $\left\{\frac{s_n^2}{f(\sigma^2 n)}, \mathscr{F}_n\right\}_{n=1}^{\infty}$ whenever (i) and (ii) ((12)) are (is) valid. For example, in the iid case, s is optimal in C with respect to the sequence $\left\{\frac{|s_n|^b}{n^a}, \mathscr{F}_n\right\}_{n=1}^{\infty}$ on the basis of the above example and Remark 3 ($2a>b>0$, $b\leqq 2$).

Hence the existence parts of the results mentioned in the Introduction follow from our results. However, Theorems 1 and 2 provide only this existence, namely, they state that s is an optimal stopping rule but they tell us nothing about the structure of s. In the martingale case this structure is not known yet.

REFERENCES

[1] CHOW, Y. S. and ROBBINS, H.: On optimal stopping rules for s_n/n, *Illinois J. Math.* **9** (1965), 444—454.

[2] CHOW, Y. S., ROBBINS, H. and SIEGMUND, D.: *Great Expectations: The Theory of Optimal Stopping*, Houghton Mifflin, Boston, 1971.

[3] DVORETZKY, A.: Existence and properties of certain optimal stopping rules, *Proc. Fifth Berkeley Symp. Math. Statist. Prob.*, Vol. 1, Univ. of California Press, 1967, 441—452.

[4] KLASS, M. J.: Properties of optimal extended-valued stopping rules for s_n/n, *Ann. Prob.* **1** (1973), 719—757.

[5] KLASS, M. J.: On stopping rules and expected supremum of s_n/a_n and $|s_n|/a_n$, *Ann. Prob.* **2** (1974), 889—905.

[6] NEVEU, J.: *Discrete-Parameter Martingales*, North-Holland, Amsterdam, 1975.

[7] RUIZ-MONCAYO, A.: Optimal stopping for functions of Markov chains, *Ann. Math. Statist.* **39** (1968), 1905—1912.

[8] SHEPP, L. A.: Explicit solutions to some problems of optimal stopping, *Ann. Math. Statist.* **40** (1969), 993—1010.

[9] SIEGMUND, D., SIMONS, G. and FEDER, P.: Existence of optimal stopping rules for rewards related to S_n/n, *Ann. Math. Statist.* **39** (1968), 1228—1235.

[10] TEICHER, H. and WOLFOWITZ, J.: Existence of optimal stopping rules for linear and quadratic rewards, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **5** (1966), 361—368.

(Received October 3, 1979)

MTA MATEMATIKAI KUTATÓ INTÉZETE
REÁLTANODA U. 13—15
H—1053 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 105—123

POISSON FORMULA AND ESTIMATIONS FOR THE LENGTH SPECTRUM OF COMPACT HYPERBOLIC SPACE FORMS

by

PAUL GÜNTHER

Introduction

Let V be an n-dimensional compact Riemannian manifold of constant negative curvature $-\gamma^2$ $(\gamma>0)$. Further let Ω be the set of nontrivial free homotopy classes of V. In every class $\theta\in\Omega$ there is one and only one closed geodesic line c_θ; let $l(\theta)$, $\nu(\theta)$ be its length and its multiplicity, respectively. The parallel displacement along c_θ induces an isometry of the tangent space in $c_\theta(0)=c_\theta(l(\theta))$; we denote its eigenvalues by $\beta_1, \beta_2, \ldots, \beta_{n-1}, 1$; we have $|\beta_i|=1$; with $\alpha_i=\mathfrak{Re}\,\beta_i$ we define

$$\sigma(\theta) = \frac{1}{\nu(\theta)} \prod_{i=1}^{n-1} (\cosh \gamma l(\theta) - \alpha_i)^{-1/2}. \tag{0.1}$$

The quantities $l(\theta)$, $\nu(\theta)$, $\sigma(\theta)$ determine the length spectrum of V introduced by H. Huber [12], [13]. (See also [21], [1], [4].)

Let ω be a harmonic 1-form on V. For every closed curve c the value of $\int_c \omega$ depends only on the free homotopy class of c and is a linear combination of the periods of ω with integer coefficients. We define for $\theta\in\Omega$:

$$f(\theta) = \exp 2\pi i \int_{c_\theta} \omega. \tag{0.2}$$

If A is the contravariant vector field belonging to ω a formal self-adjoint second order differential operator L on V is well defined by

$$L[u] = \triangle u + 4\pi i A(u) - 4\pi^2 \|A\|^2 u. \tag{0.3}$$

There is a complete orthonormal system of eigenfunctions $\{\varphi_i\}_{i\geqq 0}$, $L[\varphi_i]+\mu_i\varphi_i=0$ with nonnegative eigenvalues μ_i depending on ω. The connection between this eigenvalue spectrum Spec $\{V, \omega\}$ and the length spectrum of V is described by a Poisson formula, which we shall prove in the first part of this paper.

AMS (MOS) subject classifications (1970). Primary 53C35, 58G99, 50C05; Secondary 35Q05.

Key words and phrases. Euler—Poisson—Darboux equation for a hyperbolic space, hyperbolic space forms, eigenvalues of the Laplace—Beltrami operator, spectral geometry, closed geodesics, length spectrum, Poisson formula.

THEOREM A. *In the sense of distributions over* $\mathbf{R}$ *we have*

$$(0.4)\qquad \begin{aligned} 2D_n &:= 2\sum_{i=0}^{\infty}\cos\left(\sqrt{\mu_i-[\gamma(n-1)/2]^2}\,t\right) = \\ &= \operatorname{vol} V\cdot T_n+2^{\frac{1-n}{2}}\sum_{\theta\in\Omega} f(\theta)\,l(\theta)\,\sigma(\theta)\{\delta_{l(\theta)}+\delta_{-l(\theta)}\}. \end{aligned}$$

The distribution $T_n\in\mathscr{D}'(\mathbf{R})$ *is defined by*

$$(0.5)\qquad \langle T_n,\varphi\rangle = (-2\pi)^{\frac{1-n}{2}}\Lambda^{\left(\frac{n-1}{2}\right)}\{\varphi(t)+\varphi(-t)\}_{|t=0}$$

if n *is odd and by*

$$(0.6)\qquad \langle T_n,\varphi\rangle = (-2\pi)^{\frac{1-n}{2}}\int_{-\infty}^{\infty}\cosh\gamma t/2\cdot\Lambda^{\left(\frac{n}{2}\right)}\{\varphi(t)+\varphi(-t)\}\,dt$$

if n *is even; in both cases* $\varphi\in\mathscr{D}(\mathbf{R})$ *and* $\Lambda:=(\gamma/\sinh\gamma t)\,d/dt$.

The distribution D_n is the trace of the fundamental solution for the hyperbolic differential equation over $\mathbf{R}\times V$:

$$(0.7)\qquad \frac{\partial^2 u}{\partial t^2}-L[u]-\gamma\left(\frac{n-1}{2}\right)^2 u = 0.$$

According to a general theorem of J. J. DUISTERMAAT and V. W. GUILLEMIN [5]

$$(0.8)\qquad \operatorname{supp}\operatorname{sing} D_n = \{\pm l(\theta)\mid\theta\in\Omega\}\cup\{0\}.$$

It seems remarkable that for n odd D_n vanishes outside its singular support. This expresses the fact that in this case equation (0.7) obeys Huygens principle (in the strict sense).

From Theorem A we derive a more complicated formula analogous to (0.4) for the distribution

$$(0.9)\qquad E_n := \sum_{i=0}^{\infty}\cos\left(\sqrt{\mu_i}\,t\right)$$

(Proposition 3.3).

Poisson formulas (or equivalent Jacobi formulas) for the distribution E_n in the case $\omega=0$ were given by several authors with different degree of expliciteness and for different types of manifolds with negative curvature. (See the papers [2], [3], [4], [5], [14], [15], [17], [20].) By our knowledge the case $\omega\neq 0$ was not yet treated.

In the second part of this paper we give estimations for the length spectrum with the "weights" $f(\theta)$. Setting

$$(0.10)\qquad P(t) := \sum_{\theta\in\Omega,\ \nu(\theta)=1} f(\theta)\,\eta_{l(\theta)}(t),$$

$$(0.11)\qquad N(t) := \sum_{\theta\in\Omega} f(\theta)\,\eta_{l(\theta)}(t),$$

(η_x is the Heaviside function with jump point x) we can prove

THEOREM B. *If $F(t)$ equals one of the functions $P(t)$ or $N(t)$, then for $t\to\infty$:*

$$(0.12)\qquad F(t) = \sum_{\frac{(n-1)^2}{2n}\leqq\delta_i\leqq\frac{n-1}{2}} E\left(t, \gamma\left[\frac{n-1}{2}+\delta_i\right]\right)+O\left(E\left(t, \frac{\gamma(n-1)(2n-1)}{2n}\right)\right).$$

Here is

$$(0.13)\qquad \gamma\delta_i = \left(\gamma^2\left[\frac{n-1}{2}\right]^2-\mu_i\right)^{1/2}, \quad i = 0, 1, 2, \ldots, l;$$

$\mu_0, \mu_1, \ldots, \mu_l$ *are those of the eigenvalues for which* $0\leqq\mu_i\leqq\gamma^2(n-1)^2/4$. *Finally*

$$(0.14)\qquad E(t, \alpha) = \int_1^t \frac{e^{\alpha\tau}}{\tau}\,d\tau, \quad \alpha > 0.$$

This theorem was given by H. HUBER [13] in the case $n=2$ and without weights, i.e., for $\omega=0$. Weaker results also for $\omega=0$ were proved in the papers [1], [21], [4].

In order to prove Theorems A and B we use the theory of the Euler—Poisson—Darboux equation with parameter λ:

$$(0.15)\qquad \frac{\partial^2 u}{\partial t^2}+\lambda\gamma\coth\gamma t\frac{\partial u}{\partial t}-L[u]+\frac{\gamma^2}{4}\{\lambda^2-(n-1)^2\}u = 0.$$

If $\mathfrak{Re}\,\lambda$ is large enough the solution of the singular Cauchy problem for (0.15) possesses a simple integral representation with kernel $A(t, x, y, \lambda)$, $x, y\in V$. A "correspondence principle"[1] allows to perform the transition from λ to $\lambda\pm 2k$, $k\in\mathbf{Z}$ in a clear manner. The investigation of the aforesaid kernel is in the first instance combined with the Selberg trace formula, thus giving Theorem A, in the second instance it is combined with an estimation method due to E. LANDAU [16], thus giving Theorem B. (For the euclidean case see [7].)

§ 1

Let $\mathbf{H}_n$ be the n-dimensional real hyperbolic space of sectional curvature $-\gamma^2$ and ω a 1-form in $\mathbf{H}_n$ with $d\omega=0$, $\delta\omega=0$. There is a smooth function g with $\omega=dg$. Let A be the vector field belonging to ω and [2]

$$(1.1)\qquad L[u]:= \triangle u+4\pi iA(u)-4\pi^2\|A\|^2u.$$

We consider the following Euler—Poisson—Darboux equation for functions $u_\lambda(t, x)\in C^\infty(\mathbf{R}\times\mathbf{H}_n)$ and parameter $\lambda\in\mathbf{R}$

$$(1.2)\qquad \mathfrak{D}_\lambda[u_\lambda]:=\left\{\frac{\partial^2}{\partial t^2}+\lambda\gamma\coth\gamma t\frac{\partial}{\partial t}-L+\frac{\gamma^2}{4}(\lambda^2-(n-1)^2)\right\}[u_\lambda] = 0$$

[1] This means formulas (1.8), (1.9). By our knowledge the name was introduced by A. WEINSTEIN for the corresponding formulas in the euclidean case.

[2] Here $\|A\|^2$: A^iA_i.

with initial conditions:

$$u_\lambda(0, \cdot) = \varphi \in C^\infty(\mathbf{H}_n), \quad \frac{\partial u_\lambda}{\partial t}(0, \cdot) = 0. \tag{1.3}$$

If $\lambda > n-1$, $t>0$ the unique solution of (1.2), (1.3) is given by

$$\begin{aligned} u_\lambda(t, x) = \frac{\Gamma((\lambda+1)/2)}{\pi^{n/2}\Gamma((\lambda-n+1)/2)} \left(\frac{\sinh \gamma t}{\gamma}\right)^{1-\lambda} \times \\ \times \int_{\mathbf{H}_n} \left\{\frac{2}{\gamma^2}[\cosh \gamma t - \cosh \gamma r(x, y)]\right\}_+^{\frac{\lambda-n-1}{2}} U(x, y)\varphi(y)\, d\varrho(y). \end{aligned} \tag{1.4}$$

Here ϱ is the invariant measure on $\mathbf{H}_n$ and $r(x, y) := \operatorname{dist}(x, y)$. The function $U(x, y)$ is defined for $(x, y) \in \mathbf{H}_n \times \mathbf{H}_n$ by

$$U(x, y) = \exp 2\pi i\big(g(y) - g(x)\big) = \exp 2\pi i \int_x^y \omega \tag{1.5}$$

and has the properties

$$\begin{aligned} U(x, y) = \overline{U(y, x)}, \quad |U(x, y)| = 1, \\ U(x, y)\, U(y, z) = U(x, z). \end{aligned} \tag{1.6}$$

Finally,

$$\{\tau\}_+^\beta = \begin{cases} \tau^\beta & \text{if } \tau > 0, \\ 0 & \text{if } \tau \leqq 0. \end{cases} \tag{1.7}$$

The easy verification of (1.4) is left to the reader. Notice that the solution of (1.2), (1.3) is an even function of t: $u_\lambda(-t, x) = u_\lambda(t, x)$.

The differential operator $\mathfrak{D}_\lambda$ obeys the following *"correspondence principle"*:

$$\mathfrak{D}_\lambda\left[\left(\frac{\sinh \gamma t}{\gamma}\right)^{1-\lambda} u\right] = \left(\frac{\sinh \gamma t}{\gamma}\right)^{1-\lambda} \mathfrak{D}_{2-\lambda}[u], \tag{1.8}$$

$$\mathfrak{D}_\lambda\left[\frac{\gamma}{\sinh \gamma t}\frac{\partial u}{\partial t}\right] = \frac{\gamma}{\sinh \gamma t}\frac{\partial}{\partial t}\mathfrak{D}_{\lambda-2}[u]. \tag{1.9}$$

From (1.8), (1.9) it follows for the solution of (1.2), (1.3):

$$u_\lambda(t, x) = \left\{\frac{1}{\lambda+1}\left(\frac{\sinh \gamma t}{\gamma}\right)\frac{\partial}{\partial t} + \cosh \gamma t\right\} u_{\lambda+2}(t, x). \tag{1.10}$$

If the initial function $\varphi \in C^\infty(\mathbf{H}_n)$ is an eigenfunction of L:

$$L[\varphi] + \mu\varphi = 0, \tag{1.11}$$

we obtain

$$u_\lambda(t, x) = z(t; \lambda, \mu)\varphi(x), \tag{1.12}$$

where $z(t; \lambda, \mu)$ solves the ordinary differential equation

$$z''+\lambda\gamma \coth \gamma t\cdot z'+\left\{\mu+\frac{\gamma^2}{4}[\lambda^2-(n-1)^2]\right\} z=0 \tag{1.13}$$

and has initial values

$$z(0; \lambda, \mu)=1, \quad z'(0; \lambda, \mu)=0. \tag{1.14}$$

For $\lambda=n\pm 1$ the comparison of (1.12) with (1.4) gives the generalization of the mean value theorem for eigenfunctions of the Laplace operator. (See [10], [22].) From (1.10) we are led to

$$z(t; \lambda, \mu)\cdot=\left\{\frac{1}{\lambda+1}\left(\frac{\sinh \gamma t}{\gamma}\right)\frac{d}{dt}+\cosh \gamma t\right\} z(t; \lambda+2, \mu). \tag{1.15}$$

Moreover we have $z(-t; \lambda, \mu)=z(t; \lambda, \mu)$.

LEMMA 1.1. *Let $\lambda\geqq 1$ and $\varkappa>0$ be fixed. Then for $t>0$ and $\mu\geqq[\gamma(n-1)/2]^2+\varkappa$ the following estimation is valid:*

$$|z(t; \lambda, \mu)|\leqq \text{const}\left(\frac{\sinh \gamma t}{\gamma}\right)^{-\lambda/2}(\mu-[\gamma(n-1)/2]^2)^{-\lambda/4}. \tag{1.16}$$

The proof of this lemma follows from a more general proposition concerning ordinary differential equations of type (1.13). (See [8], proposition 2.2 and the literature quoted there.)

§ 2

Let $\mathfrak{G}$ be a properly discontinuous group of orientation preserving isometries with compact fundamental domain $\mathscr{F}$ operating freely on $\mathbf{H}_n$. We assume that the 1-form ω is $\mathfrak{G}$-invariant. Then we have

$$\forall \mathfrak{b}\in\mathfrak{G}:\ U(\mathfrak{b}(x), \mathfrak{b}(y))=U(x, y). \tag{2.1}$$

Further, the function $U(x, \mathfrak{b}(x))$ is independent from x. If $b_1, b_2, \dots, b_k$ are the periods of the harmonic form (ω relative to a suitable choosen basis of the first homology group $H_1(\mathbf{H}_n/\mathfrak{G}, \mathbf{Z})$, k is the first Betti number), then we obtain

$$U(x, \mathfrak{b}(x))=\exp 2\pi i\sum_{i=1}^{k} m_i(\mathfrak{b})\, b_i \tag{2.2}$$

with integers $m_i(\mathfrak{b})$. From (2.1) it follows:

$$\forall \mathfrak{c}\in\mathfrak{G}:\ U(x, \mathfrak{c}\mathfrak{b}\mathfrak{c}^{-1}(x))=U(x, \mathfrak{b}(x)) \tag{2.3}$$

and $U(x, \mathfrak{b}(x))$ only depends on the conjugacy class of $\mathfrak{b}$.

For $t>0$, $x, y\in\mathbf{H}_n$, $\lambda>n-1$ we define:

(2.4)

$$A(t, x, y, \lambda)=\frac{1}{\Gamma((\lambda-n+1)/2)}\sum_{\mathfrak{b}\in\mathfrak{G}}\left\{\frac{2}{\gamma^2}[\cosh \gamma t-\cosh \gamma r(x, \mathfrak{b}(y))]\right\}_+^{\frac{\lambda-n-1}{2}} U(x, \mathfrak{b}(y)).$$

In this formula the range of summation is finite. Especially, for $\lambda=n+1$, $\omega=0$, $A(t, x, y, \lambda)$ is the number of elements $\mathfrak{b}\in\mathfrak{G}$ with $r(x, \mathfrak{b}(y))<t$. We complete the definition of $A(t, x, y, \lambda)$ for integer $\lambda\geqq n+1$ by the agreements:

$$A(0, x, y, \lambda) = 0, \tag{2.4)'$$

$$A(-t, x, y, \lambda) = (-1)^{\lambda-1} A(t, x, y, \lambda), \quad \text{if} \quad t>0. \tag{2.4)''$$

For a $\mathfrak{G}$-automorphic initial function φ the solution $u_\lambda(t, x)$ of (1.2), (1.3) is $\mathfrak{G}$-automorphic, too; an easy transformation of (1.4) gives in this case for $\lambda>n-1$:

$$u_\lambda(t, x) = \pi^{-\frac{n}{2}} \Gamma\left(\frac{\lambda+1}{2}\right)\left(\frac{\sinh \gamma t}{\gamma}\right)^{1-\lambda} \int_{\mathscr{F}} A(t, x, y, \lambda)\varphi(y)\, d\varrho(y). \tag{2.5}$$

Let $L_2(\mathfrak{G})$ be the Hilbert space of the complex valued $\mathfrak{G}$-automorphic functions, which are quadratic summable over $\mathscr{F}$; the scalar product in $L_2(\mathfrak{G})$ is given by

$$(\varphi, \psi) = \int_{\mathscr{F}} \varphi(x)\overline{\psi(x)}\, d\varrho(x). \tag{2.6}$$

The differential operator $L[u]$ for $u\in C^2(\mathbf{H}_n)\cap L_2(\mathfrak{G})$ is formal self-adjoint; there is a complete orthonormal system $\{\varphi_i\}_{i\geqq 0}$ of $L_2(\mathfrak{G})$ with

$$L[\varphi_i]+\mu_i\varphi_i = 0, \quad i \geqq 0. \tag{2.7}$$

We denote the sequence of eigenvalues by Spec $\{\mathfrak{G}, \omega\}:=\{\mu_i\}_{i\geqq 0}$. The following facts are easy to prove. (See [7] for the euclidean case.)

(a) $\mu_i\geqq 0$. We can assume, that the sequence $\{\mu_i\}_{i\geqq 0}$ is monotonic increasing.

(b) $\mu_0=0$ if and only if the periods $b_1, \ldots, b_k$ of ω are integers.

(c) Let ω_1, ω_2 be two harmonic $\mathfrak{G}$-invariant forms, then Spec $\{\mathfrak{G}, \omega_1\}=$ $=$Spec $\{\mathfrak{G}, \omega_2\}$, if $\omega_1-\omega_2$ has integer periods.

From (1.12), (2.5) follows:

$$\int_{\mathscr{F}} A(t, x, y, \lambda)\varphi_i(y)\, d\varrho(y) = \pi^{\frac{n}{2}} \frac{\left(\frac{\sinh \gamma t}{\gamma}\right)^{\lambda-1}}{\Gamma\left(\frac{\lambda+1}{2}\right)} z(t; \lambda, \mu_i)\varphi_i(x). \tag{2.8}$$

Thus the φ_i prove eigenfunctions of the hermitian kernel $A(t, x, y, \lambda)$ and the domain $\mathscr{F}$.

LEMMA 2.1. *For fixed $\lambda>2n$ the eigenfunction development of the kernel $A(t, x, y, \lambda)$*

$$A(t, x, y, \lambda) = \pi^{\frac{n}{2}} \frac{\left(\frac{\sinh \gamma t}{\gamma}\right)^{\lambda-1}}{\Gamma\left(\frac{\lambda+1}{2}\right)} \sum_{i=0}^{\infty} z(t; \lambda, \mu_i)\varphi_i(x)\overline{\varphi_i(y)} \tag{2.9}$$

uniformly converges if $t\in(0, b)$, $x, y\in\mathbf{H}_n$.

PROOF. Using Lemma 1.1 we obtain for $\mu_i > [\gamma(n-1)/2]^2$ and $t \in (0, b)$:

$$\left|\frac{\sinh \gamma t}{\gamma}\right|^{1-\lambda} |z(t; \lambda, \mu_i)| \leqq c_1 \left|\frac{\sinh \gamma b}{\gamma}\right|^{(\lambda-2)/2} \frac{|\varphi_i(x)|^2 + |\varphi_i(y)|^2}{\mu_i^{\lambda/4}}.$$

The proof is finished by taking into consideration the well-known uniform convergence of the series

$$\sigma > \frac{n}{2}: \quad \sum_{i=0}^{\infty} \frac{|\varphi_i(x)|^2}{\mu_i^{\sigma}}.$$

(See [8], Satz 2.3 and the literature quoted there.)

LEMMA 2.2. *We define for* $\lambda \geqq n+1$, $t \in \mathbf{R}$:

$$A(t, \lambda) = \int_{\mathscr{F}} A(t, x, x, \lambda)\, d\varrho(x). \tag{2.10}$$

Then the expansion

$$A(t, \lambda) = \frac{\pi^{\frac{n}{2}}}{\Gamma((\lambda+1)/2)} \sum_{i=0}^{\infty} \left(\frac{\sinh \gamma t}{\gamma}\right)^{\lambda-1} z(t; \lambda, \mu_i) \tag{2.11}$$

is uniformly convergent in every bounded t-interval, if λ *is an integer with* $\lambda > 2n$.

PROOF. The assertion follows from Lemma 2.1, our agreements (2.4)′, (2.4)″ and the fact, that $z(t; \lambda, \mu)$ is an even function of t.

LEMMA 2.3. *Let* f *be an odd distribution,* $f \in \mathscr{D}'(\mathbf{R})$. *There is exactly one odd distribution* $\Lambda f \in \mathscr{D}'(\mathbf{R})$ *defined by*

$$\varphi \in C_0^\infty(\mathbf{R}) \to \langle \Lambda f, \varphi \rangle := \frac{1}{2}\left\langle f, \gamma\left(\frac{\varphi(t)-\varphi(-t)}{\sinh \gamma t}\right)'\right\rangle.$$

PROOF. It suffices to remark, that the mapping

$$\varphi \in \mathscr{D}(\mathbf{R}) \to \frac{\varphi(t)-\varphi(-t)}{\sinh \gamma t} \in \mathscr{D}(\mathbf{R})$$

is a continuous one.

PROPOSITION 2.1. *Let* $k \geqq 1$ *be an integer. The equation*

$$2^{-(n+k)} \pi^{(1-n)/2} \frac{d}{dt} \Lambda^{(n+k-1)} A(t, 2n+2k) = \sum_{i=0}^{\infty} \cos\left(\sqrt{\mu_i - [\gamma(n-1)/2]^2}\, t\right) \tag{2.12}$$

is valid in $\mathscr{D}'(\mathbf{R})$.

PROOF. According to (2.4)″ we consider $A(t, 2n+2k)$ as an odd distribution over $\mathbf{R}$. By Lemma 2.3 the left-hand side of (2.12) exists and is an even distribution. In the expansion (2.11) the operator $(d/dt)\Lambda^{(n+k-1)}$ can be applied term by term, because the convergence is uniform. The recursion formula (1.15) can be written in the following manner:

$$(2.13)\qquad \begin{aligned}(\lambda+1)\left(\frac{\sinh\gamma t}{\gamma}\right)^{\lambda-1} z(t;\lambda,\mu) =\\ = \left(\frac{\gamma}{\sinh\gamma t}\right)\frac{d}{dt}\left\{\left(\frac{\sinh\gamma t}{\gamma}\right)^{\lambda+1} z(t;\lambda+2,\mu)\right\}.\end{aligned}$$

From this we obtain by iteration

$$1\cdot 3\cdot 5\ldots(2n+2k-1)\, z(t;0,\mu) =$$

$$= \frac{d}{dt}\Lambda^{(n-k-1)}\left\{\left(\frac{\sinh\gamma t}{\gamma}\right)^{2n+2k-1} z(t;2n+2k;\mu)\right\}.$$

Finally, it is easily seen from (1.13), (1.14):

$$z(t;0,\mu) = \cos\left(\sqrt{\mu-[\gamma(n-1)/2]^2}\, t\right).$$

This proves (2.12).

§ 3

Let Ω be the set of nontrivial free homotopy classes of the compact orientable manifold $V=\mathbf{H}_n/\mathfrak{G}$; for $\theta\in\Omega$ let $l(\theta)$, $\nu(\theta)$, $\sigma(\theta)$ be the numbers defined in the introduction. As remarked in § 2 the function $U(x,\mathfrak{b}(x))$ depends only on the conjugacy class of $\mathfrak{b}\in\mathfrak{G}$; but these conjugacy classes are in 1—1 correspondence with the free homotopy classes of V. If $x\in\mathbf{H}_n$, $\mathfrak{b}\in\mathfrak{G}\setminus\{\mathrm{Id}\}$ are given, we choose any curve α connecting x with $\mathfrak{b}(x)$; to α corresponds a closed curve $\tilde\alpha$ in V and there is exactly one $\theta\in\Omega$ with $\tilde\alpha\in\theta$. Then we have

$$(3.1)\qquad U(x,\mathfrak{b}(x)) = \exp 2\pi i\int_{\tilde\alpha}\omega = f(\theta)$$

and the number $f(\theta)$ is correctly defined.

PROPOSITION 3.1. *Let $k(\tau)$ be defined for $\tau\geqq 0$ and piecewise continuous; the series*

$$(3.2)\qquad \mathscr{J} = \sum_{\mathfrak{b}\in\mathfrak{G}\setminus\{Id\}}\int_{\mathscr{F}} k(r(x,\mathfrak{b}(x)))\,U(x,\mathfrak{b}(x))\,d\varrho(x)$$

may be absolutely convergent. If $k(t)=k^*(\cosh\gamma t)$ we have

$$(3.3)\qquad \mathscr{J} = \frac{2\pi^{\frac{n-1}{2}}\gamma^{1-n}}{\Gamma\left(\frac{n-1}{2}\right)}\sum_{\theta\in\Omega} f(\theta)\,l(\theta)\,\sigma(\theta)\left(\cosh\gamma l(\theta)\right)^{\frac{n-1}{2}} I(\theta)$$

with

$$(3.4)\qquad I(\theta) = \int_0^\infty k^*\left([1+\tau^2]\cosh\gamma l(\theta)\right)\tau^{n-2}\,d\tau.$$

The proof of this proposition is analogous to the computations performed by H. HUBER [13], RIGGENBACH [21] and L. BERARD-BERGERY [1] in the case $k(\tau)=\cosh^{-s}\gamma\tau$, $\omega=0$; so it is superfluous to repeat them here. On the other hand, these computations are equivalent to the application of the Selberg trace formula.

PROPOSITION 3.2. *We have for* $t>0$, $\lambda>n+1$:

$$A(t,\lambda)=\frac{\varrho(\mathscr{F})}{\Gamma((\lambda-n+1)/2)}\left\{\frac{2}{\gamma^2}[\cosh\gamma t-1]\right\}^{(\lambda-n-1)/2}+ \tag{3.5}$$
$$+\frac{\pi^{(n-1)/2}}{2^{(n-1)/2}\Gamma(\lambda/2)}\sum_{\theta\in\Omega}f(\theta)l(\theta)\sigma(\theta)\left\{\frac{2}{\gamma^2}[\cosh\gamma t-\cosh\gamma l(\theta)]\right\}_+^{(\lambda-2)/2}.$$

PROOF. From (2.10), (2.4) follows:

$$A(t,\lambda)= \tag{3.6}$$
$$=\frac{1}{\Gamma\left(\frac{\lambda-n+1}{2}\right)}\sum_{\mathfrak{b}\in\mathfrak{G}}\int_{\mathscr{F}}\left\{\frac{2}{\gamma^2}\left[\cosh\gamma t-\cosh\gamma r(x,\mathfrak{b}(x))\right]\right\}_+^{\frac{\lambda-n-1}{2}}U(x,\mathfrak{b}(x))\,d\varrho(x).$$

We have to apply Proposition 3.1 with

$$k(\tau):=\frac{1}{\Gamma\left(\frac{\lambda-n+1}{2}\right)}\left\{\frac{2}{\gamma^2}[\cosh\gamma t-\cosh\gamma\tau]\right\}_+^{\frac{\lambda-n-1}{2}}. \tag{3.7}$$

The arising integral $I(\theta)$ is reduced to an Euler integral (Beta-function) by the substitution

$$\tau=\tau'\left\{\frac{\cosh\gamma t-\cosh\gamma l(\theta)}{\cosh\gamma l(\theta)}\right\}^{1/2}.$$

The first term in the right-hand side of equation (3.5) corresponds to the term with $\mathfrak{b}=\mathrm{Id}$ in equation (3.6).

COROLLARY. *Let k be an integer with* $0\leq k<(\lambda-2)/2$, $\lambda>n+1$, *then* $A(\cdot,\lambda)\in C^k(\mathbf{R}_+)$.

THEOREM A. *In the sense of distributions over* $\mathbf{R}$ *we have*

$$2D_n:=2\sum_{i=0}^{\infty}\cos\left(\sqrt{\mu_i-[\gamma(n-1)/2]^2}\,t\right)= \tag{3.8}$$
$$=\varrho(\mathscr{F})T_n+2^{(1-n)/2}\sum_{\theta\in\Omega}f(\theta)l(\theta)\sigma(\theta)\{\delta_{l(\theta)}+\delta_{-l(\theta)}\}.$$

T_n is given by formulas (0.5), (0.6) *of the introduction.*

PROOF. We combine (3.6) for $\lambda=2n+2k$, $k\in\mathbf{Z}$, $k\geqq 1$ with (2.12). According to our agreements $A(t, 2n+2k)$ is an odd function of t. Firstly we have to compute the distribution

$$\frac{d}{dt}\Lambda^{(n+k-1)}\operatorname{sig}(t)\left\{\frac{2}{\gamma^2}[\cosh\gamma t-\cosh\gamma l(\theta)]\right\}_+^{n+k-1}.$$

For it we easily obtain

$$(3.9)\qquad \begin{aligned}&(n+k-1)!\,2^{n+k-1}\frac{d}{dt}\{\eta_{l(\theta)}(t)-\eta_{-l(\theta)}(-t)\}=\\ &=(n+k-1)!\,2^{n+k-1}\{\delta_{l(\theta)}+\delta_{-l(\theta)}\}.\end{aligned}$$

This yields the terms of the sum in the right-hand side of (3.8). Further we have to consider the function

$$A_0(t, 2n+2k):=\frac{\operatorname{sig} t}{\Gamma\left(\dfrac{n+2k+1}{2}\right)}\left\{\frac{2}{\gamma^2}[\cosh\gamma t-1]\right\}_+^{(n+2k-1)/2},$$

and to apply the operator $(d/dt)\Lambda^{(n+k-1)}$ on it. Let φ be an even test function, $\varphi\in C_0^\infty(\mathbf{R})$; then we get

$$(3.10)\qquad K:=2^{-(n+k)}\pi^{(1-n)/2}\left\langle\frac{d}{dt}\Lambda^{(n+k-1)}A_0(\cdot, 2n+2k),\varphi\right\rangle=$$

$$=\frac{(-1)^{n+k}2^{1-n-k}\pi^{(1-n)/2}}{\Gamma((n+2k+1)/2)}\int_0^\infty\left\{\frac{2}{\gamma^2}[\cosh\gamma t-1]\right\}^{\frac{(n+2k-1)}{2}}\frac{d}{dt}\Lambda^{(n+k-1)}\varphi(t)\,dt.$$

The last integral makes sense for integer k with $n+2k-1\geqq -1$; the expression K is independent of k. If n is odd we choose $k=-(n-1)/2$ and obtain

$$(3.11)\qquad K=(-2\pi)^{\frac{1-n}{2}}\Lambda^{\left(\frac{n-1}{2}\right)}\varphi(t)|_{t=0}=\frac{1}{2}\langle T_n,\varphi\rangle.$$

If n is even we choose in (3.10): $k=-n/2$. This leads to

$$(3.12)\qquad K=\frac{2}{(-2\pi)^{n/2}}\int_0^\infty\cosh\gamma t/2\cdot\Lambda^{(n/2)}\varphi(t)\,dt=\frac{1}{2}\langle T_n,\varphi\rangle.$$

Finally, if φ is an odd test function, then both K and $\langle T_n,\varphi\rangle$ vanish. Theorem A is proved.

REMARK 1. We have

$$(3.13)\qquad \operatorname{sing\,supp} D_n=\{\pm l(\theta)|\theta\in\Omega\}\cup\{0\}.$$

REMARK 2. If n is odd the regular part of D_n vanishes. In this case we can prove an alternative expression for T_n $(n\geqq 3)$:

$$(3.14)\qquad \langle T_n,\varphi\rangle=\frac{(-\pi)^{\frac{1-n}{2}}\left(\dfrac{n-3}{2}\right)!}{(n-2)!}\prod_{\nu=0}^{(n-3)/2}\left(\frac{d^2}{dt^2}-\nu^2\gamma^2\right)\varphi(0).$$

REMARK 3. If n is even outside its singular support D_n equals the function

$$1\cdot 3\cdot 5\dots(n-1)\frac{2\varrho(\mathscr{F})}{(-2\pi)^{n/2}}\left(\frac{\gamma}{2}\right)^n\frac{\cosh\gamma t/2}{(\sinh\gamma t/2)^n}. \tag{3.15}$$

LEMMA 3.1. *For $\mu>0$, $\alpha>0$ and $t\in\mathbf{R}$ we have*

$$\begin{aligned}\cos(\sqrt{\mu}t)-\cos(\sqrt{\mu-\alpha}\,t) &= \\ &= -\alpha t\int_0^t \frac{\mathscr{J}_1(\sqrt{\alpha(t^2-z^2)})}{\sqrt{\alpha(t^2-z^2)}}\cos(\sqrt{\mu-\alpha}\,z)\,dz.\end{aligned} \tag{3.16}$$

$\mathscr{J}_1$: Bessel function with index 1.

PROOF. If one uses the power expansions of the involved functions, the proof is easily obtained.

LEMMA 3.2. *We define for $\varphi\in C_0^\infty(\mathbf{R})$, $z\in\mathbf{R}$, $\alpha:=[\gamma(n-1)/2]^2$:*

$$L_n[\varphi](z) = -\alpha\int_{|z|}^\infty \frac{\varphi(t)+\varphi(-t)}{2}\,\frac{\mathscr{J}_1(\sqrt{\alpha(t^2-z^2)})}{\sqrt{\alpha(t^2-z^2)}}\,t\,dt. \tag{3.17}$$

Then we have for $\mu>0$:

$$\begin{aligned}\langle\cos\sqrt{\mu}t-\cos(\sqrt{\mu-\alpha}\,t),\ \varphi(t)\rangle &= \\ &= \langle\cos(\sqrt{\mu-\alpha}\,t),\ L_n[\varphi](t)\rangle.\end{aligned} \tag{3.18}$$

PROOF. One has to use (3.16), (3.17) and to interchange the order of integration.

PROPOSITION 3.3. *We define the distribution*

$$E_n:=\sum_{i=0}^\infty \cos\sqrt{\mu_i}t\in\mathscr{D}'(\mathbf{R}). \tag{3.19}$$

Then we have

$$E_n = D_n+D_n\circ L_n. \tag{3.20}$$

PROOF. From the definition of L_n it follows, that the mapping $\mathscr{D}(\mathbf{R})\ni\varphi\to\to L_n[\varphi]\in\mathscr{D}(\mathbf{R})$ is a continuous one. Then (3.20) follows from (3.18).

§ 4

DEFINITION 4.1. We set for $t\geqq 0$:

$$G(t) = \sum_{\theta\in\Omega} f(\theta)l(\theta)\sigma(\theta)\eta_{l(\theta)}(t). \tag{4.1}$$

We make the transformation of the variable t setting $\xi=\cosh\gamma t$ and write $G(t)= =G^*(\xi)$, $A(t,\lambda)=A^*(\xi,\lambda)$. Finally we define:

$$G_n^*(\xi) = \int_1^\xi\int_1^{\eta_n}\dots\int_1^{\eta_2} G^*(\eta_1)\,d\eta_1\dots d\eta_n. \tag{4.2}$$

LEMMA 4.1. *For* $\xi \geqq 1$:

$$G_n^*(\xi) = \frac{\gamma^{2n}}{\pi^{(n-1)/2} 2^{(n+1)/2}} A^*(\xi, 2n+2) - \frac{\varrho(\mathscr{F})\gamma^{n-1}(\xi-1)^{(n+1)/2}}{\Gamma\big((n+3)/2\big)\pi^{(n-1)/2}}. \tag{4.3}$$

PROOF. From (4.1), (4.2) it follows:

$$G_n^*(\xi) = \frac{1}{n!} \sum_{\theta \in \Omega} f(\theta) l(\theta) \sigma(\theta) \{\xi - \cosh \gamma l(\theta)\}_+^n. \tag{4.4}$$

On the other hand, Proposition 3.2 with $\lambda = 2n+2$ gives:

$$\begin{aligned} A_n^*(\xi, 2n+2) &= \frac{\varrho(\mathscr{F}) 2^{(n+1)/2}}{\Gamma\big((n+3)/2\big)\gamma^{n+1}} (\xi-1)^{(n+1)/2} + \\ &+ \frac{2^{(n+1)/2}\pi^{(n-1)/2}}{n!\,\gamma^{2n}} \sum_{\theta \in \Omega} f(\theta) l(\theta) \sigma(\theta) \{\xi - \cosh \gamma l(\theta)\}_+^n. \end{aligned} \tag{4.5}$$

These formulas prove our assertion.

DEFINITION 4.2. Let $\mu_0, \mu_1, \ldots, \mu_l$ be those of the eigenvalues for which $0 \leqq \mu_i \leqq \leqq [\gamma(n-1)/2]^2$ is valid. We set $\gamma\delta_i := \big([\gamma(n-1)/2]^2 - \mu_i\big)^{1/2}$, $i = 0, 1, \ldots, l$. Further we define

$$H(t) := 2^{(n+1)/2} \sum_{i=0}^{l} \frac{\sinh \gamma\delta_i t}{\gamma\delta_i}, \quad H(t) = H^*(\xi); \tag{4.6}$$

$$R(t) := G(t) - H(t), \quad R(t) = R^*(\xi). \tag{4.7}$$

Finally, let $H_n^*(\xi)$, $R_n^*(\xi)$ be defined analogous to $G_n^*(\xi)$ by the iterated integral (4.2).

LEMMA 4.2. *For* $\xi \geqq 1$ *we have*

$$\begin{aligned} G_n^*(\xi) &= H_n^*(\xi) + \frac{\pi^{1/2}(\xi^2-1)^{(2n+1)/2}}{2^{(n+1)/2}\gamma\Gamma\big((2n+3)/2\big)} \sum_{i=l+1}^{\infty} z^*(\xi; 2n+2, \mu_i) - \\ &- \frac{\varrho(\mathscr{F})\gamma^{n-1}}{\Gamma\big((n+3)/2\big)\pi^{(n-1)/2}} (\xi-1)^{(n+1)/2}. \end{aligned} \tag{4.8}$$

(*Of course:* $z^*(\xi; \ldots) = z(t; \ldots)$.)

PROOF. We carry in formula (4.3) the development (2.11) of $A(t, 2n+2)$. From (2.13) it follows by iteration

$$\begin{aligned} &1 \cdot 3 \cdot 5 \ldots (2n+1) \frac{\sinh \gamma t}{\gamma} z(t; 2, \mu_i) = \\ &= \Lambda^{(n)} \left\{ \left(\frac{\sinh \gamma t}{\gamma} \right)^{2n+1} z(t; 2n+2, \mu_i) \right\}. \end{aligned} \tag{4.9}$$

It is easily seen that

$$\frac{\sinh \gamma t}{\gamma} z(t; 2, \mu_i) = \frac{\sinh \gamma\delta_i t}{\gamma\delta_i}, \qquad i = 0, 1, 2, \ldots, l;$$

$$\Lambda^{(n)} = \gamma^{2n} d^n/d\xi^n.$$

An n-fold integration of (4.9) then gives

$$\frac{\pi^{1/2}(\xi^2-1)^{(2n+1)/2}}{2^{(n+1)/2}\gamma\Gamma(2n+3)/2)} \sum_{i=0}^{l} z^*(\xi; 2n+2, \mu_i) = H_n^*(\xi). \tag{4.10}$$

This completes the proof of (4.8).

Definition 4.3. For $a\in(0, 1)$ we set $\eta=\xi^a$. For a complex valued function $F(\xi)$, defined for $\xi\geqq 1$, let the difference operators $\nabla^\pm$ be given by

$$\nabla^+ F(\xi) = \sum_{\nu=0}^{n} \binom{n}{\nu} (-1)^{n-\nu} F(\xi+\nu\eta),$$

$$\nabla^- F(\xi) = \sum_{\nu=0}^{n} \binom{n}{\nu} (-1)^{\nu} F(\xi-\nu\eta).$$

(Certainly ∇^- is defined if $\xi\geqq(n+1)^{1/(1-a)}$.) In the sequel $c_1, c_2, \ldots$ are positive numbers independent of ξ and μ_i.

Lemma 4.3.

$$\nabla^\pm R_n^*(\xi) = O(\xi^{n+a-1}). \tag{4.11}$$

Proof. From Lemma 4.2 follows

$$R_n^*(\xi) = R_n^1(\xi)+R_n^2(\xi)$$

with

$$R_n^1(\xi) = c_1(\xi^2-1)^{(2n+1)/2} \sum_{i=l+1}^{\infty} z^*(\xi; 2n+2, \mu_i), \tag{4.12}$$

$$R_n^2(\xi) = c_2(\xi-1)^{(n+1)/2}. \tag{4.13}$$

In order to estimate $\nabla^\pm R_n^1(\xi)$ we use Lemma 1.1 in the following form:

$$i \geqq l+1: \; |z^*(\xi; 2n+2, \mu_i)| \leqq c_3(\xi^2-1)^{-(n+1)/2}\mu_i^{-(n+1)/4}.$$

This gives

$$|\nabla^\pm\{(\xi^2-1)^{(2n+1)2} z^*(\xi; 2n+2, \mu_i)\}| \leqq$$

$$\leqq \sum_{\nu=0}^{n} \binom{n}{\nu} |([\xi\pm\nu\eta]^2-1)^{(2n+1)/2} z^*(\xi\pm\nu\eta; 2n+2, \mu_i)| \leqq$$

$$\leqq \frac{c_4}{\mu_i^{(n+1)/2}} \sum_{\nu=0}^{n} \binom{n}{\nu} ([\xi\pm\nu\eta]^2-1)^{n/2} \leqq \tag{4.14}$$

$$\leqq c_5\xi^n/\mu_i^{(n+1)/2}.$$

On the other hand we have

$$|\nabla \pm \{(\xi^2-1)^{(2n+1)/2} z^*(\xi;\, 2n+2,\, \mu_i)\}| =$$

$$= \eta^n \left| \frac{d^n}{d\xi^n} \{(\xi^2-1)^{(2n+1)/2} z^*(\xi;\, 2n+2,\, \mu_i)\} \right|_{\xi=\tilde{\xi}},$$

with $\tilde{\xi}\in(\xi,\, \xi \pm n\eta)$. Using (4.9) we obtain

$$= c_6 \eta^n |(\xi^2-1)^{1/2} z^*(\xi;\, 2,\, \mu_i)|_{\xi=\tilde{\xi}},$$

and applying Lemma 1.1 once more

$$\leqq c_7 \eta^n / \mu_i^{1/2} = c_7 \xi^{na} / \mu_i^{1/2}. \tag{4.15}$$

Let b be a positive number; we can write

$$|\nabla \pm R_n^1(\xi)| \leqq c_1 \sum_{i \geqq l+1,\, \mu_i < \xi^b} |\nabla \pm \{(\xi^2-1)^{(2n+1)/2} z^*(\xi;\, 2n+2,\, \mu_i)\}| +$$

$$+ c_1 \sum_{i \geqq l+1,\, u_i \geqq \xi^b} |\nabla \pm \{(\xi^2-1)^{(2n+1)/2} z^*(\xi;\, 2n+2,\, \mu_i)\}|.$$

In the first sum we estimate according to (4.15), in the second sum according to (4.14). This gives

$$\leqq c_8 \xi^{na} \sum_{i \geqq l+1,\, \mu_i < \xi^b} \mu_i^{-1/2} + c_8 \xi^n \sum_{i \geqq l+1,\, \mu_i \geqq \xi^b} \mu_i^{-(n+1)/2}.$$

Now we use the well-known formulas valid for $\zeta \to \infty$:

$$\sum_{\mu_i < \zeta} \mu_i^{-\varrho} = O(\zeta^{\frac{n}{2}-\varrho}) \quad \text{if} \quad \varrho < \frac{n}{2},$$

$$\sum_{\mu_i \geqq \zeta} \mu_i^{-\varrho} = O(\zeta^{\frac{n}{2}-\varrho}) \quad \text{if} \quad \varrho > \frac{n}{2}.$$

The result is

$$|\nabla \pm R_n^1(\xi)| = O(\xi^{na+(n-1)b/2}) + O(\xi^{n-b/2}).$$

We choose $b=2(1-a)$, then the exponents on the right-hand side equals. This gives the assertion related to $R_n^1(\xi)$. Further we have

$$|\nabla \pm R_n^2(\xi)| \leqq c_9 \eta^n \left| \frac{d^n}{d\xi^n} (\xi-1)^{(n+1)/2} \right|_{\xi=\tilde{\xi}},$$

with another $\tilde{\xi}\in(\xi,\, \xi \pm n\eta)$. A simple estimation proves the last factor $O(1)$. We obtain

$$|\nabla \pm R_n^2(\xi)| \leqq c_{10} \xi^{na} = O(\xi^{n-1+a}).$$

Lemma 4.3 is completely proved.

PROPOSITION 4.1. *For* $t \to \infty$:

$$G(t) = 2^{(n+1)/2} \sum_{\frac{n-1}{2} \geqq \delta_i > \frac{(n-1)^2}{2n}} \frac{\sinh \gamma \delta_i t}{\gamma \delta_i} + O(e^{\gamma \frac{(n-1)^2}{2n} t}). \tag{4.16}$$

PROOF. From (4.2) and Definition 4.3 follows with $\eta=\xi^a$:

$$\nabla^+ G_n^*(\xi) = \int_{\xi}^{\xi+\eta}\int_{\eta_n}^{\eta_n+\eta}\dots\int_{\eta_2}^{\eta_2+\eta} G^*(\eta_1)\,d\eta_1\dots d\eta_n. \tag{4.17}$$

An analogous formula is valid for $\nabla^+ H_n^*(\xi)$.

Firstly, we treat the case $\omega=0$; in this case $f(\theta)=1$ and $G(t)$ is real valued and monotonic increasing. The same is true for $H(t)$. Then we obtain from

$$\nabla^+ G_n^*(\xi) = \nabla^+ H_n^*(\xi)+\nabla^+ R_n^*(\xi)$$

at once

$$G^*(\xi) \leqq H^*(\xi+n\eta)+\eta^{-n}\nabla^+ R_n^*(\xi).$$

A simple application of the mean value theorem shows:

$$H^*(\xi+n\eta) = H^*(\xi)+O(\xi^{a+(n-3)/2}).$$

Lemma 4.3 gives

$$G^*(\xi) \leqq H^*(\xi)+O(\xi^{a+(n-3)/2})+O(\xi^{(n-1)(1-a)}).$$

Now we choose $a=(n+1)/2n$. With this value both exponents in the right-hand side of the last inequality equals. So we have shown

$$G^*(\xi) \leqq H^*(\xi)+O(\xi^{\frac{(n-1)^2}{2n}}).$$

Using the difference operator ∇^- we obtain the inequality

$$G^*(\xi) \geqq H^*(\xi-n\eta)+\eta^{-n}\nabla^- R_n^*(\xi),$$

and from it with the same value of a:

$$G^*(\xi) \geqq H^*(\xi)+O(\xi^{\frac{(n-1)^2}{2n}}).$$

Together with the above inequality we get

$$G^*(\xi) = H^*(\xi)+O(\xi^{\frac{(n-1)^2}{2n}}).$$

Returning to the variable t and omitting those members in the sum for $H(t)$ which have the same order like the remainder term we obtain the assertion (4.16) in the case $\omega=0$.

Now we treat the general case $\omega\neq 0$. For the purpose of our proof we denote the functions G, H, R related to $\omega=0$ by $\tilde{G}$, $\tilde{H}$, $\tilde{R}$. From (4.17) it follows with $\eta=\xi^a$:

$$|\nabla^+ G_n^*(\xi)-\eta^n G^*(\xi)| \leqq$$

$$\leqq \int_{\xi}^{\xi+\eta}\int_{\eta_n}^{\eta_n+\eta}\dots\int_{\eta_2}^{\eta_2+\eta} |G^*(\eta_1)-G^*(\xi)|\,d\eta_1\dots d\eta_n \leqq$$

$$\leqq \int_{\xi}^{\xi+\eta}\int_{\eta_n}^{\eta_n+\eta}\dots\int_{\eta_2}^{\eta_2+\eta} \{\tilde{G}^*(\eta_1)-\tilde{G}^*(\xi)\}\,d\eta_1\dots d\eta_n.$$

The transition from G^* to $\tilde{G}^*$ is legitimate because $|f(\theta)|=1$. We obtain

$$|\nabla^+ G_n^*(\xi)-\eta^n G^*(\xi)| \leqq \nabla^+ \tilde{G}_n^*(\xi)-\eta^n \tilde{G}^*(\xi) =$$
$$= \nabla^+ \tilde{H}_n^*(\xi)-\eta^n \tilde{H}^*(\xi)+\nabla^+ \tilde{R}_n^*(\xi)-\eta^n \tilde{R}^*(\xi).$$

These considerations together with the first part of the proof show

$$|G^*(\xi)-\eta^{-n}\nabla^+ G_n^*(\xi)| = O(\xi^{\frac{(n-1)^2}{2n}}). \tag{4.18}$$

Further we have

$$\begin{aligned} \eta^{-n}\nabla^+ G_n^*(\xi) &= \eta^{-n}\nabla^+ H_n^*(\xi)+\eta^{-n}\nabla^+ R_n^*(\xi) = \\ &= H^*(\xi)+\{\eta^{-n}\nabla^+ H_n^*(\xi)-H^*(\xi)\}+\eta^{-n}\nabla^+ R_n^*(\xi). \end{aligned} \tag{4.19}$$

The last two terms are of order $O(\xi^{\frac{(n-1)^2}{2n}})$. The formulas (4.18), (4.19) give

$$G^*(\xi) = H^*(\xi)+O(\xi^{\frac{(n-1)^2}{2n}}).$$

This proves Proposition 4.1.

DEFINITION 4.4. For every $\theta\in\Omega$ we set

$$\mu(\theta) = \frac{1}{\nu(\theta)}\prod_{i=1}^{n-1}\left(1-\frac{\alpha_i}{\cosh \gamma l(\theta)}\right)^{-1/2}. \tag{4.20}$$

(Compare the definitions of $\nu(\theta)$, $\sigma(\theta)$, given in the introduction.) Obviously, we have $\mu(\theta)=(\cosh \gamma l(\theta))^{(n-1)/2}\sigma(\theta)$. Further we define for $t\geqq 0$:

$$\varphi(t) = \sum_{\theta\in\Omega} f(\theta)\mu(\theta)\eta_{l(\theta)}(t). \tag{4.21}$$

Finally, we set for $\alpha>0$, $t>0$:

$$E(t,\alpha) = \int_1^t e^{\alpha\tau}/\tau \, d\tau. \tag{4.22}$$

PROPOSITION 4.2. *For* $t\to\infty$

$$\varphi(t) = \sum_{\frac{n-1}{2}\geqq\delta_i\geqq\frac{(n-1)^2}{2n}} E\left(t,\gamma\left[\frac{n-1}{2}+\delta_i\right]\right)+O\left(E\left(t,\gamma\left[\frac{n-1}{2}+\frac{(n-1)^2}{2n}\right]\right)\right). \tag{4.23}$$

PROOF. From (4.18) and (4.1) it follows with a sufficient small positive number ε:

$$\varphi(t) = \int_\varepsilon^t \frac{(\cosh \gamma\tau)^{(n-1)/2}}{\tau}\, dG(\tau). \tag{4.24}$$

We denote the remainder term in (4.16) by $r(t)$; then it is easily shown by partial integration

$$\int_{\varepsilon}^{t} \frac{(\cosh \gamma\tau)^{(n-1)/2}}{\tau}\, dr(\tau) = O\left(\frac{1}{t} e^{\gamma t\left[\frac{n-1}{2}+\frac{(n-1)^2}{2n}\right]}\right). \tag{4.25}$$

Owing to

$$\lim_{t\to\infty} \frac{e^{\alpha t}}{tE(t, \alpha)} = \alpha$$

we can write

$$\int_{\varepsilon}^{t} \frac{(\cosh \gamma\tau)^{(n-1)/2}}{\tau}\, dr(\tau) = O\left(E\left(t, \gamma\left[\frac{n-1}{2}+\frac{(n-1)^2}{2n}\right]\right)\right).$$

Further we have to consider the integrals

$$I_i := \frac{2^{(n+1)/2}}{\delta_i \gamma} \int_{\varepsilon}^{t} \frac{(\cosh \gamma\tau)^{(n-1)/2}}{\tau}\, d \sinh \gamma\delta_i\tau =$$

$$= 2^{(n+1)/2} \int_{\varepsilon}^{t} \frac{(\cosh \gamma\tau)^{(n-1)/2}}{\tau} \cosh \gamma\delta_i\tau \cdot d\tau.$$

In the integrand is

$$2^{(n+1)/2}(\cosh \gamma\tau)^{(n-1)/2} \cosh \gamma\delta_i\tau = e^{\gamma\tau\left(\frac{n-1}{2}+\delta_i\right)} + O(e^{\gamma\tau\varepsilon_i})$$

with $\varepsilon_i = (n-1)/2+\delta_i-2\min\{1, \delta_i\}$. Taking into account that only such values δ_i are relevant for which $(n-1)^2/2n \leqq \delta_i \leqq (n-1)/2$, we obtain $\varepsilon_i \leqq (n-1)/2+(n-1)^2/2n$. This gives

$$I_i = E\left(t, \gamma\left[\frac{n-1}{2}+\delta_i\right]\right) + O\left(E\left(t, \gamma\left[\frac{n-1}{2}+\frac{(n-1)^2}{2n}\right]\right)\right). \tag{4.26}$$

From (4.24), (4.25), (4.26) Proposition 4.2 follows.

Proposition 4.3. *For* $t\to\infty$:

$$\psi(t) := \sum_{\theta\in\Omega} \frac{f(\theta)}{\nu(\theta)} \eta_{l(\theta)}(t) = \varphi(t) + O\left(E\left(t, \gamma\left[\frac{n-1}{2}+\frac{(n-1)^2}{2n}\right]\right)\right). \tag{4.27}$$

Proof. For every $\theta\in\Omega$ the following inequality is valid:

$$\left|\mu(\theta) - \frac{1}{\nu(\theta)}\right| = \mu(\theta)\left|1 - \prod_{i=1}^{n-1}\left(1 - \frac{\alpha_i}{\cosh \gamma l(\theta)}\right)^{1/2}\right| \leqq c_{11}\mu(\theta)[\cosh \gamma l(\theta)]^{-1}.$$

Thereby we have used that $|\alpha_i| \leqq 1$; this leads to

$$|\varphi(t)-\psi(t)| \leqq c_{11} \sum_{\theta\in\Omega} \frac{|f(\theta)|\,\mu(\theta)}{\cosh \gamma l(\theta)} \eta_{l(\theta)}(t) \leqq c_{11} \int_{0}^{t} \frac{d\tilde{\varphi}(\tau)}{\cosh \gamma\tau}. \tag{4.28}$$

In (4.28) $\tilde{\varphi}$ is the function φ related to the case $\omega=0$. From Proposition 4.2 it follows at once

$$\int_0^t \frac{d\tilde{\varphi}(\tau)}{\cosh \gamma\tau} = O\big(E(t, \gamma(n-2))\big).$$

Because $n-2 \leqq (n-1)/2+(n-1)^2/2n$ the assertion follows.

Now we are able to prove Theorem B formulated in the introduction. These last steps are the same as in the papers of H. HUBER [12], [13].

PROOF of Theorem B. Firstly we treat the case $\omega=0$. The above considered function $\psi(t)$ and the functions $P(t)$, $N(t)$ defined in the introduction satisfy the relations

$$\psi(t) = P(t) + \sum_{m=2}^{\infty} \frac{1}{m} P\left(\frac{t}{m}\right), \tag{4.29}$$

$$N(t) = \sum_{m=1}^{\infty} P\left(\frac{t}{m}\right). \tag{4.30}$$

If $l_0 = \min\{l(\theta) | \theta \in \Omega\}$ then $P(t/m)=0$ for $t/m \leqq l_0$. From (4.29) it follows

$$0 \leqq \psi(t) - P(t) \leqq \sum_{2 \leqq m \leqq t/l_0} \frac{1}{m} \psi\left(\frac{t}{m}\right) \leqq \frac{t}{l_0} \psi\left(\frac{t}{2}\right).$$

Proposition 4.2 and 4.3 gives

$$P(t) = \psi(t) + O\left(e^{\frac{\gamma(n-1)}{2}t}\right) \tag{4.31}$$

and this is the assertion related to $P(t)$. From (4.30), (4.31) it follows the assertion related to $N(t)$.

Now we treat the general case $\omega \neq 0$. The functions ψ, P, N related to the special case $\omega=0$ are denoted by $\tilde{\psi}$, $\tilde{P}$, $\tilde{N}$. Then we have according to (4.31) (taken for the special case)

$$|\psi(t) - P(t)| \leqq \tilde{\psi}(t) - \tilde{P}(t) = O\left(e^{\frac{\gamma(n-1)}{2}t}\right)$$

and this gives (4.31) in the general case. Finally we have

$$|N(t) - P(t)| \leqq \tilde{N}(t) - \tilde{P}(t) = \sum_{m=2}^{\infty} \tilde{P}\left(\frac{t}{m}\right) = O\left(e^{\frac{\gamma(n-1)}{2}t}\right).$$

From this the last assertion follows for the general case.

REFERENCES

[1] BERARD-BERGERY, L.: Laplacien et géodésiques fermées sur les formes d'espace hyperboliques compactes. *Séminaire Bourbaki*, 1971—72, 107—122 (Lectures Notes in Mathematics, 317).

[2] CHAZARAIN, J.: Formule de Poisson pour les variétés riemanniennes, *Inventiones Math.* **24** (1974), 65—82.

[3] DONELLY, H.: On the wave equation asymptotics of a compact negatively curved surface, *Inventiones Math.* **45** (1978), 115—137.

[4] DE GEORGE, D. L.: Length spectrum for compact locally symmetric spaces of strictly negative curvature, *Ann. Ec. Norm. Sup.* **10** (1977), 133—152.

[5] DUISTERMAAT, J. J.—GUILLEMIN, V. W.: The spectrum of positive elliptic operators and periodic bicharacteristics, *Inventiones Math.* **29** (1975), 39—79.

[6] DUISTERMAAT, J. J.—KOLK, J. A. C.—VARADARAJAN, V. S.: Spectra of compact locally symmetric spaces of negative curvature, *Inventiones Math.* (to appear).

[7] GÜNTHER, P.: Eine Funktionalgleichung für den Gitterrest, *Math. Nachr.* **76** (1977), 5—27.

[8] GÜNTHER, P.: Gitterpunktprobleme in Riemannschen symmetrischen Räumen vom Rang 1, *Math. Nachr.* **94** (1980), 5—27.

[9] GÜNTHER, P.: Problème de réseaux dans les espaces hyperboliques, *C. R. Acad. Sci. Paris. Sér. A—B* (1979), A49—A52.

[10] HELGASON, S.: *Differential geometry and symmetric spaces,* Academic Press, New York, 1962.

[11] HÖLDER, E.: Poissonsche Wellenformel in nichteuklidischen Räumen, *Berichte Verh. Sächs. Acad. Wiss. Leipzig* **99** (1938), 55—66.

[12] HUBER, H.: Zur analytischen Theorie hyperbolischer Raumformen und Bewegungsgruppen I, *Math. Ann.* **138** (1959), 1—26.

[13] HUBER, H.: Zur analytischen Theorie hyperbolischer Raumformen und Bewegungsgruppen II, *Math. Ann.* **142** (1961), 385—398; **143** (1961), 463—464.

[14] KOLK, J.: Formule de Poisson et distribution asymptotique du spectre simultané d'opérateurs différentiels, *C. R. Acad. Sci. Paris. Sér. A—B* **284** (1977), 1045—1048.

[15] KOLK, J.: The Selberg trace formula and asymptotic behaviour of spectra, Thesis, Rijksuniversiteit Utrecht, 1977.

[16] LANDAU, E.: Über eine Aufgabe aus der Theorie der quadratischen Formen, *Sitzungsberichte Akad. Wien* **24** (1915), 445—468.

[17] MCKEAN, H. P.—SINGER, J. M.: Curvature and the eigenvalues of the Laplacian, *J. Differential Geometry* **1** (1967), 43—69.

[18] OLEVSKY, M. N.: On a generalization of Bessel functions, *Doklady Akad. Nauk SSSR* **40,** No. 1 (1943), 5—10.

[19] OLEVSKY, M. N.: Solution du problème de Cauchy pour l'équation des ondes dans un espace à n dimensions à courbure constante, *Doklady Akad. Nauk SSSR* **46**, No. 1, (1945), 3—6.

[20] RANDOL, B.: On the analytic continuation of the Minakshisundaram—Pleijel zeta function for compact Riemannian surfaces, *Trans. Amer. Math. Soc.* **201** (1975), 241—246.

[21] RIGGENBACH, H.: Freie Homotopieklassen und das Eigenwertspektrum des Laplace-Operators bei hyperbolischen Raumformen, Inauguraldissertation, Universität Basel, 1975.

[22] RUSE, H.S.—WALKER, A. G.—WILLMORE, T. J.: *Harmonic spaces,* Edizioni Cremonese, Roma, 1961.

[23] SELBERG, A.: Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series, *J. Indian Math. Soc.* **20** (1956), 47—87.

(Received October 28, 1979)

SEKTION MATHEMATIK DER
KARL-MARX-UNIVERSITÄT
KARL-MARX-PLATZ
DDR—701 LEIPZIG
GERMAN DEMOCRATIC REPUBLIC

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 125—129

TWO METHODS OF CONSTRUCTION OF FREE BOOLEAN ALGEBRAS

by

ALEXANDER ABIAN

Abstract

Let $\{g_0, g_1, g_2, \ldots\}$ be an infinite set of preassigned cardinality. We give below two methods of associating with every g_i a dyadic sequence g_i^* (of an appropriate infinite type) such that $\{g_0^*, g_1^*, g_2^*, \ldots\}$ becomes a set of free generators of a free Boolean algebra whose elements are (finite) sums of (finite) products of g_i^*'s and the corresponding unit dyadic sequence and where addition and multiplication among these elements are performed coordinatewise Mod 2. One of the two methods yields the Stone representation of the corresponding free Boolean algebra and both methods coincide when $\{g_0, g_1, g_2, \ldots\}$ is a finite set.

As usual, a sequence (of any finite or infinite type) is called *dyadic* if and only if its terms are 0 or 1. We recall that a ring with a multiplicative unit is called a *Boolean algebra* if and only if every element of it is idempotent (i.e., is equal to its own square). It is well-known [1, p. 19] that every Boolean algebra is isomorphic to a subset A of the set of all dyadic sequences of an appropriate type such that the *unit dyadic sequence* (i.e., the sequence all of whose terms are 1) is an element of A and such that A is closed under coordinatewise addition and multiplication Mod 2 which constitute the addition and the multiplication of the Boolean algebra A. In fact, it is always possible that the elements of A be of the type of the cardinal of the set of all the prime ideals of A.

We recall also [1, p. 47] that a (not necessarily countable) set $\{g_0, g_1, g_2, \ldots\}$ is called a *set of free generators* of a Boolean algebra A with unit u if and only if $\{g_0, g_1, g_2, \ldots\} \cup \{u\}$ generates A and every mapping of $\{g_0, g_1, g_2, \ldots\}$ into a Boolean algebra B can be extended to a homomorphism from A into B.

Based on the above, we describe below our First Method of constructing a free Boolean algebra on a set of preassigned (finite or infinite) cardinality of free generators.

First Method. Let k be any (finite or infinite) cardinal (i.e., k is an initial ordinal number). To construct a free Boolean algebra on the set $\{g_i | i<k\}$ of k free generators $g_0, g_1, g_2, \ldots$ we set up a k by 2^k matrix M with k rows $g_0^*, g_1^*, g_2^*, \ldots$ (or for the sake of simplicity, with k rows $g_0, g_1, g_2, \ldots$) whose columns are *all the possible k-termed* dyadic sequences (i.e., sequences of ordinal type k whose terms are

AMS (MOS) subject classifications (1970). Primary 06A40; Secondary 16A06.
Key words and phrases. Free Boolean algebras, free generators.

This work is partly supported by the Iowa State University SHRI.

0 or 1). Thus:

$$
\begin{array}{ll}
& g_0 = 0\,1\,0\,1\,0\,1\,0\,1\,0\,1\,0\,1\,0\,1\,0\,1 \ldots \\
& g_1 = 0\,0\,1\,1\,0\,0\,1\,1\,0\,0\,1\,1\,0\,0\,1\,1 \ldots \\
& g_2 = 0\,0\,0\,0\,1\,1\,1\,1\,0\,0\,0\,0\,1\,1\,1\,1 \ldots \\
(1) & g_3 = 0\,0\,0\,0\,0\,0\,0\,0\,1\,1\,1\,1\,1\,1\,1\,1 \ldots \\
& \ldots \\
& g_i = 0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \ldots \\
& \ldots
\end{array}
$$

As we see shortly, the rows g_i of the matrix M serve as a set of k free generators of the free Boolean algebra A whose elements consist of all the (finite) sums of the (finite) products of the rows g_i of M together with the *unit* sequence u of type 2^k (i.e., a sequence of length 2^k all of whose terms are 1) where addition and multiplication among g_i's and u are performed coordinatewise Mod 2.

Let us observe that the *zero* sequence of type 2^k is readily obtained as the sum of any of the g_i's with itself. On the other hand, however, the *unit* sequence of type 2^k cannot be obtained as the sum of products of g_i's. Indeed, let us assume to the contrary that the unit sequence u of type 2^k is a sum of products of some g_i's. Without loss of generality (since $g_i^2 = g_i$ for every $i<k$) let

$$(2) \qquad g_3 g_\omega g_2 + g_7 g_{\omega+\omega} = u.$$

Since by our construction of matrix M every dyadic sequence of type k appears as a column of M, clearly there exists a column c_i of M such that 0 is located at the intersection of c_i with each of the rows $g_3, g_\omega, g_2, g_7, g_{\omega+\omega}$ of M. Clearly, this implies that the left side of the equality sign in (2) is a dyadic sequence of type 2^k whose i-th coordinate is 0. But this contradicts the fact that the right side of the equality sign in (2) is the unit dyadic sequence of type 2^k. Thus, indeed u cannot be obtained as a sum of the products of g_i's.

Accordingly, we amend (at the beginning) matrix M as given by (1) with a row of 1's (which corresponds to the unit sequence u of type 2^k) yielding matrix M' as given by (3).

Now, let A be the set of all the (finite) sums of the (finite) products of the rows $u, g_0, g_1, g_2, g_3, \ldots, g_i, \ldots$ of matrix M' where addition and multiplication are performed coordinatewise Mod 2.

$$
\begin{array}{ll}
& u \;\; = 1\,1\,1\,1\,1\,1\,1\,1\,1\,1\,1\,1\,1\,1\,1\,1 \ldots \\
& g_0 = 0\,1\,0\,1\,0\,1\,0\,1\,0\,1\,0\,1\,0\,1\,0\,1 \ldots \\
& g_1 = 0\,0\,1\,1\,0\,0\,1\,1\,0\,0\,1\,1\,0\,0\,1\,1 \ldots \\
(3) & g_2 = 0\,0\,0\,0\,1\,1\,1\,1\,0\,0\,0\,0\,1\,1\,1\,1 \ldots \\
& g_3 = 0\,0\,0\,0\,0\,0\,0\,0\,1\,1\,1\,1\,1\,1\,1\,1 \ldots \\
& \ldots \\
& g_i = 0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \ldots \\
& \ldots
\end{array}
$$

Clearly, A is a Boolean algebra (since A is a ring with unit u and $x^2=x$ for every $x\in A$) generated by $\{g_0, g_1, g_2, \dots\}\cup\{u\}$.

We prove that A is a free Boolean algebra on the set $\{g_0, g_1, g_2, \dots\}$ of free generators g_i with $i<k$.

Indeed, let $(B, +, \cdot)$ be a Boolean algebra with unit u'. Let h be a mapping from $\{g_0, g_1, g_2, \dots\}$ into B. We show that h can be extended to a homomorphism from A into B. Since every element a of A is a sum of products of the rows $g_0, g_1, g_2, \dots$ together with the unit row u, we extend h to A in the obvious way. Thus, if

$$a = g_4 g_\omega g_{41}+u+g_{\omega 2+1} g_5$$

then, as expected, we set:

$$h(a) = h(g_4)h(g_\omega)h(g_{41})+u'+h(g_{\omega 2+1})h(g_5) \quad \textit{with} \quad h(u) = u'. \tag{4}$$

The fact that h is a homomorphism will follow trivially from (4) *as soon as we establish that* h (given by (4)) *is well defined.* To this end we must show that if a and b are elements of A such that $a=b$ then $h(a)=h(b)$.

Without loss of generality, let $a=g_3g_7+u+g_2g_5$ and $b=g_7g_9$. Thus, assuming $g_3g_7+u+g_2g_5=g_7g_9$ we must prove that $h(g_3)h(g_7)+u'+h(g_2)h(g_5)=h(g_7)h(g_9)$. In other words, assuming

$$g_3 g_7+u+g_2 g_5+g_7 g_9 = 0 \quad (\textit{the zero sequence in } (3)) \tag{5}$$

we must prove that

$$h(g_3)h(g_7)+u'+h(g_2)h(g_5)+h(g_7)h(g_9) = 0 \quad (\textit{the zero of } B). \tag{6}$$

Clearly, $h(g_2)$, $h(g_3)$, $h(g_5)$, $h(g_7)$, $h(g_9)$ and u' are a finite number of elements of the Boolean algebra B. Let B' be the finite (Boolean) subalgebra of B generated by $h(g_2)$, $h(g_3)$, $h(g_5)$, $h(g_7)$, $h(g_9)$ and u'. As such, every element of B' can be represented as a dyadic sequence of length less than 2^k (in fact, of finite length). Thus, in particular, we may assume that:

$$\begin{aligned} u' &= 1\,1\,1\,1\,1\,1\,1\,1 \dots 1\\ h(g_2) &= 1\,0\,0\,1\,0\,1\,1\,0 \dots 0\\ h(g_3) &= 0\,0\,1\,1\,1\,0\,0\,1 \dots 0\\ h(g_5) &= 1\,1\,0\,0\,1\,1\,0\,0 \dots 1\\ h(g_7) &= 0\,1\,0\,1\,0\,1\,0\,1 \dots 0\\ h(g_9) &= 1\,1\,1\,0\,0\,1\,1\,1 \dots 1. \end{aligned} \tag{7}$$

However, since in (1) every possible dyadic sequence of type k appears as a column (of matrix M), it is clear that in (3) there are columns $c_i, c_j, c_k, \dots$ such that the intersection of c_i with the rows $u, g_2, g_3, g_5, g_7, g_9$ of (3) is identical with the first column of (7). Similarly, in (3), the intersection of column c_j with the rows $u, g_2, g_3, g_5, g_7, g_9$ of (3) is identical with the second column of (7). Likewise, in (3), the intersection of c_k with the rows $u, \ldots, g_9$ of (3) is identical with the third column of (7), ... But then from (5) it follows that the left side of the equality sign in (6) when interpreted as rows of (7) yields the zero row of (7), i.e., the zero element of B' and therefore the zero element of B. Thus, (6) is established.

Hence, indeed the rows g_i of the k by 2^k matrix M as given in (1) form a set of k free generators of the free Boolean algebra A mentioned above.

The fact that A given as the set of all the (finite) sums of the (finite) products (with coordinatewise addition and multiplication Mod 2) of the rows $u, g_0, g_1, g_2, g_3, \ldots, \ldots, g_i, \ldots$ of matrix M' is the Stone representation of A can be shown as follows. Let P be a prime ideal of A. Then P is uniquely defined by a homomorphism p from A onto the two-element Boolean algebra $\{0, 1\}$. Clearly, p in turn is uniquely defined by an assignment of 0 or 1 to the generators $g_0, g_1, g_2, \ldots$. Since by our construction of matrix M as given in (1), every dyadic sequence of type k appears precisely once as a column of M, we see that to every prime ideal P of A there corresponds a column of M and vice versa. But this ensures that the representation of A by all those dyadic sequences of length 2^k each of which is obtained as a sum of products of the rows of matrix M' as given in (3), is the Stone representation [1, p. 18] of A.

Remark. It is easily seen that if k is a finite cardinal then the sums of the products of the rows of the corresponding k by 2^k matrix M' yield every possible dyadic sequence of length 2^k. Thus, in this case the corresponding algebra A is a (finite) free Boolean algebra on a finite set of k free generators and has 2^{2^k} elements. Clearly, in this case, A is isomorphic to the direct product of 2^k copies of 2-element field $\{0, 1\}$.

Recalling that every two free Boolean algebras with equipollent sets of free generators are isomorphic [1, p. 48] from the above it follows that we have proved:

Theorem 1. *The free Boolean algebra A on a set of k (where k is a finite or an infinite cardinal) free generators is isomorphic to a subalgebra of the direct product of 2^k copies of 2-element field* $\{0, 1\}$. *Moreover, a set of k free generators of A is given by the rows of a k by 2^k matrix M whose columns are all possible dyadic sequences of length k. Thus, every element of A is represented as a sum of products (coordinatewise* Mod 2*) of the rows of matrix M together with the unit sequence of length* 2^k. *Furthermore, this representation of A is the Stone representation of A. Clearly, if k is an infinite cardinal then A has k elements and if k is a finite cardinal then A has 2^{2^k} elements in which case it is isomorphic to the direct product of 2^k copies of 2-element field* $\{0, 1\}$.

Second Method. If we examine our proof of Theorem 1, and, in particular, our reasoning in connection with (2), (5), (6), we see that all the reasonings remain valid if instead of matrix M as given in (1), we consider a k-rowed matrix M_1 whose columns are all possible k-termed dyadic sequences in each of which *all but at most a finite number of terms are equal to* 0. This is because in the left side of equality sign in each of (2), (5), (6), only a finite number of g_i's occur. Clearly, if k is an infinite cardinal then M_1 is a k by k matrix. Therefore, in this case, the elements of the ensuing free Boolean algebra A on a set of k free generators are represented by dyadic sequences of length k. Thus, this representation is not the Stone representation of A. On the other hand, if k is a finite cardinal, then in view of the above Remark, M_1 is a k by 2^k matrix in which case A is isomorphic to the direct product of 2^k copies of 2-element field $\{0, 1\}$, just as in the case of Theorem 1. In short, we have:

THEOREM 2. *The free Boolean algebra A on a set of k (where k is a finite or an infinite cardinal) free generators is isomorphic to a subalgebra of the direct product of h copies of 2-element field* $\{0, 1\}$. *Moreover, a set of k free generators of A is given by the rows of a k by h matrix M_1 whose columns are all possible k-termed dyadic sequences in each of which all but at most a finite number of terms are equal to* 0. *Thus, every element of A is represented as a sum of products (coordinatewise* Mod 2*) of the rows of matrix M_1 together with the unit sequence of length h. Furthermore, this representation of A is not the Stone representation of A if k is an infinite cardinal, If k is infinite then A has k elements and $h=k$. If k is finite then A has 2^{2^k} elements and $h=2^k$ and A is isomorphic to the direct product 2^k copies of 2-element field* $\{0, 1\}$ *and the representation of A is the Stone representation.*

REFERENCE

[1] DWINGER, PH.: *Introduction to Boolean Algebras,* Hamburger Math. Einzelschr. **40** (1961), Physica-Verlag, Würzburg.

(Received November 12, 1979)

DEPARTMENT OF MATHEMATICS
IOWA STATE UNIVERSITY
AMES, IA 50011
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 131—137

EINE ALGEBRAISCHE CHARAKTERISIERUNG DES NORMALFORMENPROBLEMS FÜR AUTONOME DIFFERENTIALSYSTEME

von

J. SCHWAIGER

§ 1. Einleitung

Das Problem der Bestimmung von Normalformen analytischer autonomer Differentialgleichungssysteme ist in der Literatur wohlbekannt und oft und in vielfacher Hinsicht behandelt worden (vgl. z. B. [1], [2], [3]). Hier sollen formale Überlegungen im Vordergrund stehen. Formale Normalformen sind von großen Interesse in der Theorie der analytischen (und kontinuierlichen) Iterationen formaler Potenzreihen (vgl. [4], [5]). Im folgenden werden nachstehende Bezeichnungen verwendet (man vergleiche wieder mit [4], [5]).

$X={}^t(X_1, X_2, \ldots, X_n)$ ist ein n-dimensionaler Vektor der Unbestimmten $X_1, X_2, \ldots, X_n$ und $\mathbf{C}[X]$ der Ring der formalen Potenzreihen mit komplexen Koeffizienten in diesen Unbestimmten. Für ein $F\in\mathbf{C}[X]$ und für $T={}^t(T_1, \ldots, T_n)\in(\mathbf{C}[X])^n$ ist $F\circ T$ als formale Reihe erklärt, wenn nur die Ordnung aller T_i größer als 0 ist (d. h., wenn alle T_i nur Glieder $a_\nu X^\nu$ enthalten, deren Grad größer als 0 ist). Eine Reihe $S\in(\mathbf{C}[X])^m$ schreiben wir in der Form

$$S = \sum_{|\nu|\geqq 0} p_\nu X^\nu, \tag{1}$$

mit $p_\nu\in\mathbf{C}^m$, $\nu=(\nu_1, \ldots, \nu_n)\in\mathbf{N}_0^n$, $|\nu|:=\sum_{j=1}^n \nu_j$, $X^\nu:=X_1^{\nu_1}\cdot\ldots\cdot X_n^{\nu_n}$ und definieren $(S)_\nu:=p_\nu$.

Definiert man für diese S als Ordnung $\operatorname{ord} S:=\min\{|\nu|\,|\,(S)_\nu\neq 0\}$, und ist $\operatorname{ord} S\geqq 1$, so schreibt man S in der Form $S=AX+\mathfrak{S}(X)$, wobei A eine komplexe (m, n)-Matrix ist und die Ordnung von $\mathfrak{S}$ größer als 1 ist.

Zwei Differentialsysteme $\big(F, G\in(\mathbf{C}[X])^n$, $\operatorname{ord} F$, $\operatorname{ord} G\geqq 1$, $f(t)=\sum_{|\nu|\geqq 1} p_\nu(t)X^\nu$, $g(t)=\sum_{|\nu|\geqq 1} q_\nu(t)X^\nu\big)$

$$\dot f = F(f) \tag{2a}$$

und

$$\dot g = G(g) \tag{2b}$$

heißen äquivalent, wenn für ein $T\in(\mathbf{C}[X])^n$ der Form $T=AX+\mathfrak{T}(X)$ mit $\det A\neq 0$

$$F\circ T = \frac{\partial T}{\partial X}G \tag{3}$$

AMS (MOS) subject classifications (1970). Primary 34A25, 34A30; Secondary 39B70.
Key words and phrases. Normalforms for autonomous systems of differential equations.

gilt. Dabei ist

$$\frac{\partial T}{\partial X}=\begin{pmatrix}\frac{\partial T_1}{\partial X_1}\cdots\frac{\partial T_1}{\partial X_n}\\ \vdots \\ \frac{\partial T_n}{\partial X_1}\cdots\frac{\partial T_n}{\partial X_n}\end{pmatrix}$$

die Jacobische Matrix von T.

Ist nun $F=AX+\mathfrak{P}(X)$ gegeben, so seien die Jordan-Blöcke $J_1, \ldots, J_r$ in der Jordan'schen Normalform J von A so angeordnet, daß der Block

$$J_i=\begin{pmatrix}\lambda_i & 1 & & 0\\ & \ddots & \ddots & \\ & & \ddots & 1\\ 0 & & & \lambda_i\end{pmatrix}$$

in den Zeilen $t_{i-1}+1, \ldots, t_i$ der Matrix J stehe.

Ein Monom X^ν heiße glatt für λ_i (bezüglich $\Lambda:=(\lambda_1, \ldots, \lambda_r)$), falls mit $|\nu|_j:=\sum_{k=t_{j-1}+1}^{t_j}\nu_k$ gilt:

$$\lambda_i=\nu\cdot\lambda:=\sum_{j=1}^{r}|\nu|_j\lambda_j. \tag{4}$$

Mit $L^{(i)}$ wird die Menge der Multiindizes $\nu\in\mathbf{N}_0^n$ bezeichnet, für die X^ν glatt für λ_i ist.

Definition 1. $N=BX+\mathfrak{N}(X)\in(\mathbf{C}[X])^n$ wird (glatte) Normalform genannt, falls gilt:

i) $B=J$ ist eine Jordan'sche Normalform.

ii) In $\mathfrak{N}_k$, der k-ten Komponente von $\mathfrak{N}$, treten für $t_{i-1}+1\leq k\leq t_i$ nur für λ_i glatte Monome X^ν wirklich auf.

In [4] findet sich das folgende Ergebnis:

Satz. *Zu jedem $F=AX+\mathfrak{P}(X)$ existiert eine Normalform N und eine formalbiholomorphe Transformation $T=CX+\mathfrak{T}(X)$, sodaß*

$$F\circ T=\frac{\partial T}{\partial X}N$$

gilt.

Mit anderen Worten, es ist also jedes formale Differentialsystem

$$\dot{f}=F(f) \tag{5}$$

äquivalent zu einem System

$$\dot{g}=N(g), \tag{6}$$

wobei N die Struktur einer Normalform aufweist.

Das analoge Problem der Bestimmung von Normalformen für formalholomorphe Abbildungen wurde ebenfalls ausführlich behandelt (vgl. [6]). In [7] findet sich ein in [8] näher ausgeführter Hinweis, wie die Zusatzmonome im Normalformenproblem für Abbildungen algebraisch deutbar sind mit Hilfe gewisser Vektorraumendomorphismen. In [9] schließlich wurde u. a. gezeigt, daß (5) eine lineare Normalform besitzt, falls gewisse kanonisch durch den Linearteil von F bestimmte Endomorphismen auf dem Raum der homogen Polynome eines Grades i $(i=2, 3, \ldots)$ in den Unbestimmten $X_1, \ldots, X_n$ mit Koeffizienten aus $\mathbf{C}^n$ Automorphismen sind. Hier sollen nun unter Verwendung der eben erwähnten Endomorphismen die Zusatzmonome im Normalformenproblem für Differentialsysteme ebenfalls algebraisch gedeutet werden.

§ 2. Das Normalformenproblem

Im folgenden wird der sogenannte „Fitting'sche Zerlegungssatz" benötigt (vgl. [10], p. 378) Dieser lautet folgendermaßen:

Es sei V ein endlich dimensionaler K-Vektorraum und $f\colon V\to V$ ein K-linearer Endomorphismus. Dann gilt für die Unterräume $U := \bigcup_{i\geqq 0} \operatorname{Ker} f^i$ und $W := \bigcap_{i\geqq 0} \operatorname{Im} f^i$ von V:

i) V ist die direkte Summe von U und W.

ii) Die Einschränkung von f auf W liefert einen Automorphismus von W.

Außerdem sind auch die Relationen $\operatorname{Ker} f^i \subset \operatorname{Ker} f^{i+1}$ und $\operatorname{Im} f^{i+1} \subset \operatorname{Im} f^i$ erfüllt. Nun betrachten wir $\mathscr{P}_l$, den Vektorraum der homogen Polynome von Grad l in den Unbestimmten $X_1, \ldots, X_n$ mit Koeffizienten aus $\mathbf{C}^n$. Für eine eingangs erwähnte Jordan'sche Normalform J betrachtet man dann die folgende lineare Abbildung $\Phi_l = \Phi_{l,J}$:

$$\Phi_l\colon \mathscr{P}_l \to \mathscr{P}_l,\ \Phi_l(q) = \frac{\partial q}{\partial X} JX - Jq. \tag{7}$$

Dabei wird $q \in \mathscr{P}_l$ aufgefaßt als Vektor $q = {}^t(q_1, \ldots, q_n) \in (\mathbf{C}[X])^n$. Daß Φ_l wirklich eine Abbildung von $\mathscr{P}_l$ in sich ist, ist klar deshalb, weil $\frac{\partial q}{\partial X}$ eine Matrix ist, deren Spalten in $\mathscr{P}_{l-1}$ liegen und weil $J \cdot q$ ein Element von $\mathscr{P}_l$ ist. Es ist auch leicht einzusehen, daß Φ_l $\mathbf{C}$-linear auf $\mathscr{P}_l$ operiert.

Das weitere Vorgehen gestaltet sich folgendermaßen: Schreibt man eine Reihe $N \in (\mathbf{C}[X])^n$ als Summe ihrer homogenen Bestandteile: $N = AX + \sum_{i\geqq 2} n_i$, und ist e_k der k-te Einheitsvektor des $\mathbf{C}^n$, so kann man eine Normalform im Sinne von Definition 1 auch so beschreiben:

Definition 1'. $N = AX + \sum_{i\geqq 2} n_i$ ist Normalform genau dann, wenn

i) $A = J$ eine Jordansche Normalform ist und

ii) für $i \geqq 2$ n_i in dem Unterraum $\mathscr{Q}_i$ von $\mathscr{P}_i$ liegt, der erzeugt wird von der Menge $\{e_k X^\nu \mid 1 \leqq j \leqq r,\ t_{j-1}+1 \leqq k \leqq t_j,\ \nu \in L^{(j)},\ |\nu| = i\}$.

Es wird sich nun zeigen — und das ist die Hauptabsicht der vorliegenden Arbeit —, daß $\mathcal{Q}_i$ gegeben ist durch $\mathcal{Q}_i = \bigcup_{l \geqq 0} \operatorname{Ker} \Phi_i^l$. Außerdem fällt im Zuge dieser Betrachtungen ein neuer Beweis für die Existenz von Normalformen der obigen Art an.

DEFINITION 2. Sind $\mu, \nu \in \mathbf{N}_0^n$, so heißt $\nu < \mu$ (ν kleiner als μ) genau dann, wenn für ein l_0 mit $1 \leqq l_0 \leqq n$ folgendes gilt:

i) $\nu_l = \mu_l$ für $l_0 < l \leqq n$,

ii) $\nu_{l_0} > \mu_{l_0}$.

Dies ist eine lineare Ordnung auf $\mathbf{N}_0^n$, die also speziell auf der Menge der Multiindizes ν eines festen Betrages $k = |\nu|$ eine Wohlordnung induziert. Es gilt der folgende Hilfssatz.

HILFSSATZ 1. *Für* $t_{i-1} + 1 \leqq k \leqq t_i$, $\nu \in \mathbf{N}_0^n$, $|\nu| = s$ *und* $u \in \mathbf{N}$ *gilt:*

$$\Phi_s(e_k X^\nu) = (\nu\lambda - \lambda_i)^u e_k X^\nu + \sum_{k < l \leqq t_i} c_{l,\nu} e_l X^\nu + \sum_{\substack{k \leqq l \leqq t_i \\ \mu < \nu,\ |\mu| = s}} c_{l,\mu} e_l X^\mu, \tag{8}$$

wobei die $c_{l,\mu}$ *gewisse durch* $e_k X^\nu$ *bestimmte komplexe Zahlen sind und in der zweiten Summe nur solche* μ *auftreten können, für die* $\mu\lambda = \nu\lambda$ *gilt.*

BEWEIS. Induktion nach u. Im Falle des Induktionsanfanges $u = 1$ ergibt eine explizite Berechnung

$$\Phi_s(e_k X^\nu) = e_k \sum_{l=1}^{n} \nu_l X^{\nu - e_l} s_l(X) - \lambda_i e_k X^\nu - \varepsilon e_{k+1} X^\nu. \tag{9}$$

Dabei ist $\varepsilon = 0$, falls $k = t_i$ ist, und $\varepsilon = 1$ sonst. Für $s_l(X)$ gilt: $s_{t_j}(X) = \lambda_j X_j$ sowie $s_l(X) = \lambda_j X_l + X_{l+1}$, wenn $t_{j-1} + 1 \leqq l < t_j$ ist. Eine weitere Auswertung von (9) ergibt dann

$$\Phi_s(e_k X^\nu) = e_k(\lambda\nu - \lambda_i) X^\nu + e_k \sum_{l \neq t_1, \ldots, t_n} \nu_l X^{\nu - e_l + e_{l+1}} - \varepsilon e_{k+1} X^\nu,$$

oder

$$\Phi_s(e_k X^\nu) = (\lambda\nu - \lambda_i) e_k X^\nu + \sum_{\substack{\mu < \nu,\ |\mu| = s \\ \mu \cdot \lambda = \nu \cdot \lambda}} d_{\mu,k} e_k X^\mu - \varepsilon e_{k+1} X^\nu. \tag{10}$$

(10) gilt deswegen, weil für $l \neq t_1, \ldots, t_n$ $\mu = \nu - e_l + e_{l+1}$, $\mu < \nu$ und $|\mu|_j = |\nu|_j$ ist. Damit ist der Fall $u = 1$ erledigt. Der Induktionsschluß bereitet dann, wenn auf (10) Φ_s^u angewendet wird, keine Mühe mehr.

Nun zur angekündigten Behauptung:

SATZ 1. *Für* $i \geqq 2$ *ist eine Basis von* $\mathcal{Q}_i = \bigcup_{u \geqq 0} \operatorname{Ker} \Phi_i^u$ *gegeben durch die Vektoren der Menge*

$$R_i := \{e_k X^\nu \mid 1 \leqq j \leqq r,\ t_{j-1} + 1 \leqq k \leqq t_j,\ \nu \in L^{(j)},\ |\nu| = i\}.$$

BEWEIS. R_i ist als Teilmenge der „kanonischen“ Basis

$$K = \{e_k X^\nu \mid 1 \leqq k \leqq n,\ \nu \in \mathbf{N}_0^n,\ |\nu| = i\}$$

sicher ein freies System von Vektoren des Raumes $\mathcal{P}_i$. Wenn man nun zeigen kann, daß

$$R_i \subset \mathcal{Q}_i \tag{11}$$

und

$$S_i := K \setminus R_i \subset \mathcal{R}_i := \bigcap_{u \geqq 0} \operatorname{Im} \Phi_i^u \tag{12}$$

gilt, so folgt für die von R_i bzw. S_i erzeugten Räume U bzw. V, daß U ein Unterraum von $\mathcal{Q}_i$ und V ein Unterraum von $\mathcal{R}_i$ ist. Außerdem gilt dann $\dim U \leqq \dim \mathcal{Q}_i$ und $\dim V \leqq \dim \mathcal{R}_i$. Weil aber auch — aus dem Fitting'schen Zerlegungssatz folgend — $\mathcal{P}_i = \mathcal{Q}_i \oplus \mathcal{R}_i$ gilt und weil außerdem $\mathcal{P}_i$ auch die direkte Summe von U und V ist, folgt aus einem Dimensionsvergleich die Gleichheit der Dimensionen und U und $\mathcal{Q}_i$ bzw. von V und $\mathcal{R}_i$. Dann stimmen aber auch die entsprechenden Räume überein; und damit wäre alles gezeigt.

Es bleibt also noch die Richtigkeit von (11) und (12) nachzuweisen. Ich beweise zunächst (11). Dazu werde $e_k X^\nu \in R_i$ fest gewählt und $\Phi := \Phi_i$ gesetzt. Ist dann $k = t_j$ und ν minimal bezüglich „$<$" in

$$L^{(j,i)} := L^{(j)} \cap \{\nu | \nu \in \mathbf{N}_0^n,\ |\nu| = i\},$$

so folgt aus Hilfssatz 1

$$\Phi(e_k X^\nu) = (\lambda\nu - \lambda_j) e_k X^\nu + 0 + 0 = 0 e_k X = 0.$$

Daher liegt $e_k X^\nu$ in $\mathcal{Q}_i$. Falls nun $t_{j-1} + 1 \leqq k < t_j$ ist und angenommen wird, daß ein $u \in \mathbf{N}$ gefunden ist, sodaß $\Phi^u(e_l X^\mu) = 0$ ist, wenn für $e_l X^\mu$ entweder $k < l \leqq t_j$ und $\mu \in L^{(j,i)}$ oder $l = k$, $\mu \in L^{(j,i)}$ sowie $\mu < \nu$ gilt, so folgt:

$$\Phi^{u+1}(e_k X^\nu) = \Phi^u\Big((\lambda\nu - \lambda_j) e_k X^\nu + \sum_{k < l \leqq t_j} c_{l,\nu} e_l X^\nu + \sum_{\substack{k \leqq l \leqq t_j \\ \mu < \nu,\ \mu \in L^{(j,i)}}} c_{l,\mu} e_l X^\mu\Big) =$$
$$= \Phi^u(0) + 0 + 0 = 0.$$

Das heißt aber, $e_k X^\nu$ liegt in $\operatorname{Ker} \Phi^{u+1} \subset \mathcal{Q}_i$, womit (11) gezeigt ist.

(12) beweist man ähnlich.

Ist wieder $k = t_j$ und ν minimal in

$$K^{(j,i)} = \{\nu | \nu \in \mathbf{N}_0^n,\ |\nu| = i,\ \nu\lambda \neq \lambda_j\},$$

so ergibt Hilfssatz 1 für $u \geqq 0$:

$$\Phi^u(e_k X^\nu) = (\nu\lambda - \lambda_j)^u e_k X^\nu;$$

d. h.

$$e_k X^\nu = \Phi^u\big((\nu\lambda - \lambda_j)^{-u} e_k X^\nu\big) \in \operatorname{Im} \Phi^u.$$

Nimmt man schließlich an, daß alle $e_l X^\mu \in S_i$ in $\mathcal{R}_i$ liegen, für die entweder $t_j \geqq l > k$ und $\mu \in K^{(j,i)}$ oder $l = k$, $\mu \in K^{(j,i)}$ sowie $\mu < \nu$ gilt, so kommt man bei nochmaliger Verwendung von Hilfssatz 1 zu einer Relation der Form

$$\Phi^u(e_k X^\nu) = (\lambda\nu - \lambda_j)^u e_k X^\nu + \Phi^u(p_u)$$

mit einem geeigneten $p_u \in \mathcal{P}_i$. Daher gilt schließlich

$$e_k X^\nu = \Phi^u\big((\lambda\nu - \lambda_j)^{-u} e_k X^\nu - p_u\big) \in \operatorname{Im} \Phi^u$$

für alle $u \geqq 0$. Verwendet man den soeben bewiesenen Satz, so kann man erneut die Existenz von Normalformen zeigen.

SATZ 2. *Es sei* $F = AX + \mathfrak{P}(X) = AX + \sum_{i \geqq 2} f^i \in (\mathbf{C}[X])^u$, *wobei* A *eine* (n, n)*-Matrix mit Jordan'scher Normalform* $J = \mathrm{diag}(J_1, \ldots, J_r)$ *ist und die* f_i *in* $\mathscr{P}_i$ *liegen. Dann gilt: Es existiert eine formal-biholomorphe Transformation* $U = SX + \sum_{i \geqq 2} u_i = SX + \mathfrak{U}(X)$ $(u_i \in \mathscr{P}_i)$ *und es existiert eine Normalform* $N = JX + \sum_{i \geqq 2} n_i$ $(n_i \in \mathscr{Q}_i)$, *sodaß*

$$F \circ U = \frac{\partial U}{\partial X} N \tag{13}$$

erfüllt ist.

BEWEIS. Indem man gegebenenfalls von F zu $R^{-1}FR$ übergeht, wobei die (n, n)-Matrix R $R^{-1}AR = J$ erfüllt, kann o. B. d. A. $A = J$ und $S = E$ angenommen werden. (13) ist dann äquivalent zu

$$JU + \mathfrak{P} \circ U = N + \frac{\partial U}{\partial X} N. \tag{14}$$

Vergleicht man nun in dieser Gleichung die homogenen Bestandteile vom Grad i, so erhält man für $i=1$ die Identität $JX = JX$ und für $i \geqq 2$

$$Ju_i + (\mathfrak{P} \circ U)_i = n_i + \frac{\partial u_i}{\partial X} JX + \sum_{j=2}^{i-1} \frac{\partial u_j}{\partial X} n_{i+1-j}, \tag{15}$$

wobei $(\mathfrak{F} \circ U)_i$ den homogenen Bestandteil vom Grad i in dieser Reihe bedeutet. Das bedeutet aber:

$$\Phi_i(u_i) + n_i = L_i(u_1, \ldots, u_{i-1}, n_2, \ldots, n_{i-1}) = L_i. \tag{16}$$

Dabei ist $L_i(\ldots)$ ein homogenes Polynom des Grades i, das — abgesehen von den f_i — nur abhängt von solchen u_j und n_k, für die $j, k < i$ gilt. Nun sei für $i \geqq 2$ $u_i = u_i^{(1)} + u_i^{(2)}$, $L_i = L_i^{(1)} + L_i^{(2)}$. $u_i^{(1)}$, $L_i^{(1)}$ liegen dabei in $\mathscr{Q}_i$ und $u_i^{(2)}$ und $L_i^{(2)}$ in $\mathscr{R}_i$. Wählt man jetzt für $i \geqq 2$ $u_i^{(1)}$ beliebig und schreibt $\Phi_i(u_i^{(1)})$ als Summe $v_i^{(1)} + v_i^{(2)}$ mit $v_i^{(1)} \in \mathscr{Q}_i$ und $v_i^{(2)} \in \mathscr{R}_i$, so ist (14) äquivalent zu

$$\Phi_i(u_i^{(2)}) + n_i = (L_i^{(1)} - v_i^{(1)}) + (L_i^{(1)} - v_i^{(2)}) \qquad (i \geqq 2). \tag{17}$$

Bei beliebiger Vorgabe der $u_i^{(1)}$ folgt aus dem Fitting'schen Zerlegungssatz und wegen dem über L_i Gesagten die eindeutige Auflösbarkeit von (16) nach den $u_i^{(2)}$ und n_i gemäß

$$u_i^{(2)} = \Phi_i^{-1}(L_i^{(2)} - v_i^{(2)}) \tag{18}$$

und

$$n_i = L_i^{(1)} - v_i^{(1)}, \tag{19}$$

weil ja $L_i^{(1)} - v_i^{(1)}$ in $\mathscr{Q}_i$, $L_i^{(2)} - v_i^{(2)}$ in $\mathscr{R}_i$ liegt, weil $\mathscr{P}_i$ die direkte Summe von $\mathscr{Q}_i$ und $\mathscr{R}_i$ ist und weil schließlich $\Phi_i|_{\mathscr{R}_i}$: $\mathscr{R}_i \to \mathscr{R}_i$ ein Isomorphismus ist.

§ 3. Anwendung

Als Anwendung sei der folgende Satz genannt, dessen Beweis ich, da er dem seines Analogons in [8] ziemlich genau entspricht, übergehen möchte.

SATZ 3. *Es seien* $N=JX+\sum_{i\geq 2} n_i$, $M=JX+\sum_{i\geq 2} m_i$ $(n_i, m_i\in\mathcal{Q}_i)$ *Normalformen und es sei* U *eine formal-biholomorphe Abbildung der Form* $U=BX+\sum_{i\geq 2} u_i$ *mit* $u_i\in\mathcal{P}_i$. *Dann gilt: Ist die Gleichung* $N\circ U=\frac{\partial U}{\partial X}M$ *erfüllt, so folgt daraus notwendigerweise:* i) $BJ=JB$; ii) $u_i\in\mathcal{Q}_i$ *für* $i=2, 3, \dots$.

LITERATURVERZEICHNIS

[1] BRJUNO, A.: *The analytic form of differential equations,* Trans. Moscow. Math. Soc., Vol. **26**, 1972.

[2] SIEGEL, C. L.: Über die Normalform analytischer Differentialgleichungen in der Nähe einer Gleichgewichtslösung, *Nachrichten der Akademie der Wissenschaften in Göttingen, Math.-phys. Klasse,* (1952), Nr. 5, 21—30.

[3] REICH, L.: Zur Majorantenmethode im Normalformenproblem für analytische Differentialgleichungssysteme, *J. Reine Angew. Math.* **239/240** (1970), 78—87.

[4] REICH, L. und SCHWAIGER, J.: Über einen Satz von Shl. Sternberg in der Theorie der analytischen Iterationen, *Mh. Math.* **83** (1977), 207—231.

[5] REICH, L. und SCHWAIGER, J.: Linearisierung formal-biholomorpher Abbildungen und Iterationsprobleme, *Aequationes Math.* **20** (1980), 224—243.

[6] REICH, L.: Das Typenproblem bei formal-biholomorphen Abbildungen mit anziehenden Fixpunkt, *Math. Ann.* **179** (1969), 227—250.

[7] WARE, B.: The Siegel Centerstable Theorem in Hilbertspace, *Ber. d. Math.-stat. Sektion im Forschungszentrum Graz,* Ber. Nr. **37**, 1975.

[8] REICH, L. und SCHWAIGER, J.: Über die Funktionalgleichung $N\circ T=T\circ M$ für formale Potenzreihen, *Aequationes Math.* **19** (1979), 66—78.

[9] WARE, B.: A proof of the Siegel linearization theorem by Diliberto bounded dominants, *Studia Sci. Math. Hungar.* **10** (1975), 197—202.

[10] GREUB, W.: *Linear Algebra* I, Springer-Verlag, Berlin—Heidelberg—New York, 1974.

(Eingegangen am 20. Dezember 1979)

INSTITUT FÜR MATHEMATIK
UNIVERSITÄT GRAZ
BRANDHOFGASSE 18
A—8010 GRAZ
AUSTRIA

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 139—144

REMARK ON COMPATIBLE AND ORDER-PRESERVING FUNCTION ON LATTICES

by

E. TAMÁS SCHMIDT

1. Introduction

Let k be a positive integer. If L is a lattice then $F_k(L)$ denotes the lattice of all functions f: $L^k \to L$. A function $f \in F_k(L)$ is called *compatible* if for any congruence relation Θ of L $a_i \equiv b_i(\Theta)$, $i=1, 2, \ldots, k$ imply $f(a_1, \ldots, a_k) = f(b_1, \ldots, b_h)(\Theta)$, f is called *order-preserving*, if $a_i \leqq b_i$, $i=1, \ldots, k$, implies $f(a_1, \ldots, a_k) \leqq f(b_1, \ldots, b_k)$. The set of all k-place compatible functions on L denoted by $C_k(L)$, the set of all k-place order-preserving functions on L is $OF_k(L)$. In this paper we deal with the following problem: which modular lattices satisfy $C_k(L) \subset OF_k(L)$?

For distributive lattices this problem was solved by D. DORNINGER and G. EIGENTHALER [1]:

THEOREM 1. *Let L be a distributive lattice. Then $C_k(L) \subsetneqq OF_k(L)$ if and only if L contains a proper interval which is a Boolean lattice.*

In [2] we have given a simple modular lattice M with the property that none of its proper intervals is complemented. In a simple lattice every function f: $L \to L$ is of course compatible, hence Theorem 1 cannot be generalized for an arbitrary modular lattice. In [5] R. WILLE and the author has proved the following statement:

Let L be a modular lattice of finite primitive length. Then $C_k(L) \subsetneqq OF_k(L)$ if and only if L contains a proper interval which is a Boolean lattice.

In connection to this theorem we prove the following

THEOREM 2. *There exists a modular lattice L satisfying the following conditions:*

(i) *L is the subdirect product of finite lattices;*

(ii) *there is a 1-place compatible function $\varphi \in C_1(L)$ on L which is not order-preserving;*

(iii) *none of the proper intervals $[a, b]$ of L is complemented.*

We give the proof in two steps. Let $[a, b]$ and $[c, d]$ be two isomorphic intervals of a chain C. The isomorphism f: $[a, b] \to [c, d]$ is called an interval-isomorphism, f is of course a partial operation on C, and can be extended to a unary operation $\bar{f}$ as follows: $\bar{f}(x) = f(x)$ for all $a \leqq x \leqq b$, $\bar{f}(x) = f(a)$ if $x \leqq a$ and finally $\bar{f}(x) = f(b)$ for $x \geqq b$. We say that $\bar{f}$ is the operation induced by f. The congruence relations of the partial algebra $\langle C; f \rangle$ are exactly the congruence relations of $\langle C, \bar{f} \rangle$. f is determined by $\bar{f}$, hence we can use the same letter for both.

First we construct an algebra $\mathscr{C} = \langle R, \vee, \wedge, f_i \rangle$ $i \in I$ where R denotes the bounded chain of rationals, the f_i-s are special interval-isomorphism. This will be a subdirect product of finite algebras and $\mathscr{C}$ satisfies (ii), (iii). The second step is the construction of a modular lattice L which contains $\mathscr{C}$ as a sublattice and satisfies the given properties.

2. The construction of $\mathscr{C}$

It is well-known that a bounded countable chain which is dense-in-itself is determined up to isomorphism. Therefore we can start with the chain R of rational numbers $\frac{k}{2^n}$ where $-2^n \leqq k \leqq 2^n$, $n=0, 1, \dots$. We take every $r \in R$, $r \neq \pm 1$ in two copies r and r', i.e., we split the elements. We set $1=1'$, $-1=(-1)'$.

Then we define on the set $K=\{r, r'; r\in R\}$ an ordering

(1) r is covered by r' $(r \neq \pm 1)$;

(2) $r' \leqq s$ if and only if $r \leqq s$ in R.

K is a chain and R is a subchain of K. The prime intervals of K are the following: $[r, r']$ $(r \neq \pm 1)$. The function defined by $f_{10}(r)=r+1$, $f_{10}(r')=(r+1)'$, $f_{10}(-1)=0'$ maps $[-1, 0]$ onto $[0', 1]$. f_{10} is an interval-isomorphism. On the same way we get functions f_{21}, f_{20} defined on $\left[-1, -\frac{1}{2}\right]$ resp. $\left[\left(-\frac{1}{2}\right)', 0\right]$, $f_{21}(r)=r+\frac{3}{2}$, $f_{21}(r')=$ $=\left(r+\frac{3}{2}\right)'$, $f_{21}(-1)=\left(\frac{1}{2}\right)'$; $f_{20}(r)=r+\frac{1}{2}$, $f_{20}(r')=\left(r+\frac{1}{2}\right)'$. Figure 1 helps to visualize these functions.

On the same way we can define for $n \geqq 1$, $0 \leqq k < 2^{n-1}$ the function f_{nk}:

$$f_{nk}\colon \left[\left(-\frac{k+1}{2^{n-1}}\right)', -\frac{k}{2^{n-1}}\right] \to \left[\left(\frac{k}{2^{n-1}}\right)', \frac{k+1}{2^{n-1}}\right],$$

$$f_{nk}(r) = r+\frac{2k+1}{2^{n-1}}.$$

If $r\in R$, $r \notin \{0, +1, -1\}$ then we define

$$g_r\colon [-r, (-r)'] \to [r, r'].$$

Finally, let φ be defined by $\varphi(1)=-1$, $\varphi(-1)=1$, $\varphi(-r)=r'$, $\varphi((-r)')=r$. The next step is the description of the congruence relations of $\langle K, \vee, \wedge, f_{nk}, g_r\rangle = \mathscr{K}_0$. First we define for each natural number $n \geqq 1$ an equivalence relation Θ_n on K: $x \equiv y(\Theta_n)$ if and only if there exists a k, $0 \leqq k < 2^{n-1}$ such that either $\left(\frac{k}{2^{n-1}}\right)' \leqq x, y \leqq \frac{k+1}{2^{n-1}}$ or $\left(-\frac{(k+1)}{2^{n-1}}\right)' \leqq x, y \leqq -\frac{k}{2^{n-1}}$. In Figure 1 the wavy lines denote the Θ_2-classes. There are two Θ_1-classes: $\{x, x \geqq 0'\}$ $\{x; x \leqq 0\}$. It is easy to show that Θ_n is a congruence relation of $\mathscr{K}_0$, $\mathscr{K}_0/\Theta_n$ is finite and $\bigwedge_{n=0}^{\infty} \Theta_n = \omega$. By an easy computation — applying the operations f_{nk} — we get that the principal congruence

$$\Theta\left(\left(\frac{k}{2^{n-1}}\right)', \frac{k+1}{2^{n-1}}\right)$$

is Θ_n.

Principal congruences of $\mathscr{K}_0$ are the congruence relations Θ_n and the congruence relations $\Theta(r, r')$. All these are compatible with φ, consequently φ is a congruence-preserving mapping, and φ of course is not order-preserving. $\mathscr{K}_0$ satisfies (i). $\mathscr{K}_0$ contains complemented intervals, these are the prime intervals $[r, r']$.

For each natural number i we take an isomorphic copy $\mathscr{K}_i$ of $\mathscr{K}_0$ such that $i \neq j$ implies $\mathscr{K}_i \cap \mathscr{K}_j = \emptyset$. We put an isomorphic copy of $\mathscr{K}_1$ into the prime interval $[r, r']$ of $\mathscr{K}_0$, i.e., we have an isomorphism $\varphi_r^1 \colon \mathscr{K}_1 \to [r, r']$, satisfying $\varphi_r^1(1_1) = r'$, $\varphi_r^1(0_1) = r$, where 0_1 resp. 1_1 are the zero resp. unit elements of $\mathscr{K}_1$ (see Figure 2).

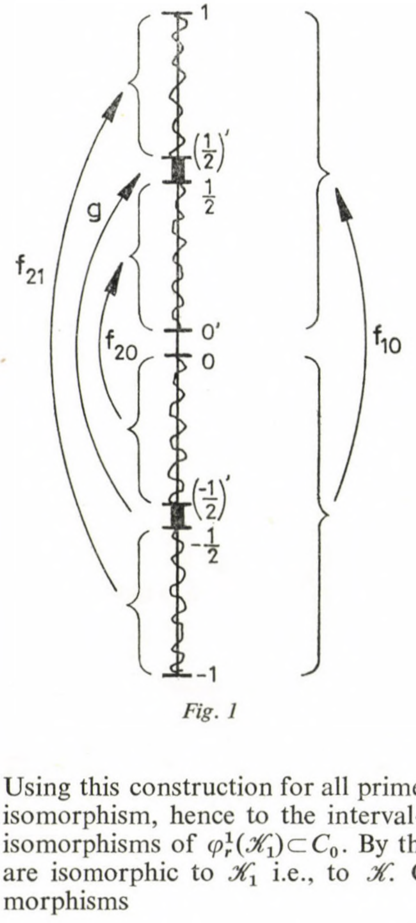


Fig. 1

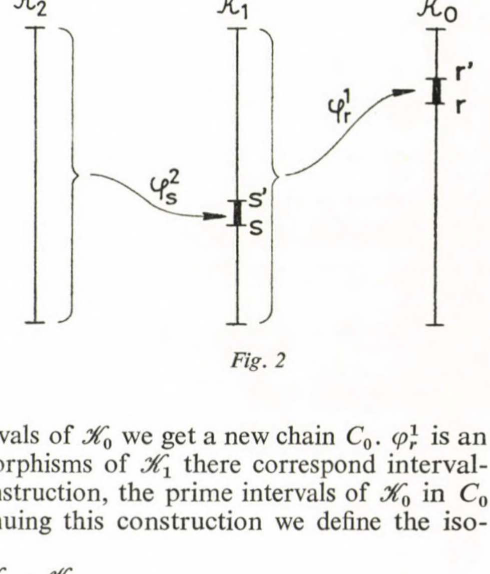


Fig. 2

Using this construction for all prime intervals of $\mathscr{K}_0$ we get a new chain C_0. φ_r^1 is an isomorphism, hence to the interval-isomorphisms of $\mathscr{K}_1$ there correspond interval-isomorphisms of $\varphi_r^1(\mathscr{K}_1) \subset C_0$. By this construction, the prime intervals of $\mathscr{K}_0$ in C_0 are isomorphic to $\mathscr{K}_1$ i.e., to $\mathscr{K}$. Continuing this construction we define the isomorphisms

$$\varphi_r^2 \colon \mathscr{K}_2 \to \mathscr{K}_1$$

then $\bigcup_{r,s \in R} \left(\varphi_s^2(\mathscr{K}_2) \cup \varphi_r^1(\mathscr{K}_1) \cup \mathscr{K}_0\right)$ is a chain C_1. On this way we get a sequence of chains $C_0 \subset C_1 \subset \dots$. Let $\mathscr{C}$ be the chain $\bigcup_{i=0} C_i$, i.e., the direct limit of the C_i-s.

The conditions (ii) and (iii) are obviously satisfied. We prove that $\mathscr{C}$ is the subdirect product of finite algebras. We define special congruences Φ_n on $\mathscr{C}$ $(n=0, 1, \dots)$ such that $\mathscr{C}/\Phi_n$ is finite and $\bigwedge_{n=0}^{\infty} \Phi_n = \omega$.

By the isomorphism $\mathscr{K}_0\cong\mathscr{K}_i$ the image of the congruence relation Θ_n is denoted by Θ_n^i. Let us take Θ_1^j on $\mathscr{K}_j$ for $j>n$ and Θ_n^i on $\mathscr{K}_i$ for $i\leqq n$. By the construction of $\mathscr{C}$ the image of these congruences defines a congruence relation Φ_n on $\mathscr{C}$, Φ_n is the transitive hull of all Θ_1^j and Θ_n^i $(i\leqq n<j)$. Then it is easy to show that $\bigwedge\Phi_n=\omega$. On the other hand from $\mathscr{K}_0/\Theta_1\cong 2$ it follows that $\mathscr{C}/\Phi_n$ is finite.

3. The construction of the modular lattice L

We denote the chain of all non-negative rational numbers by Q^+ and Q^- is the chain of all non-positive rationals. We define a sublattice D of $Q^+\times Q^-$. Let $A=\{(x,y);\ 0\leqq x<1,\ -1<y\leqq 0\}\subseteq Q^+\times Q^-$, $B=\{(r,-1);\ r\geqq 1\}$ and

$$C=\{(1,r);\ r\leqq -1\}.$$

Then $(Q^+\times Q^{-1})\setminus\{A\cup B\cup C\}$ is a sublattice D (see Figure 3).

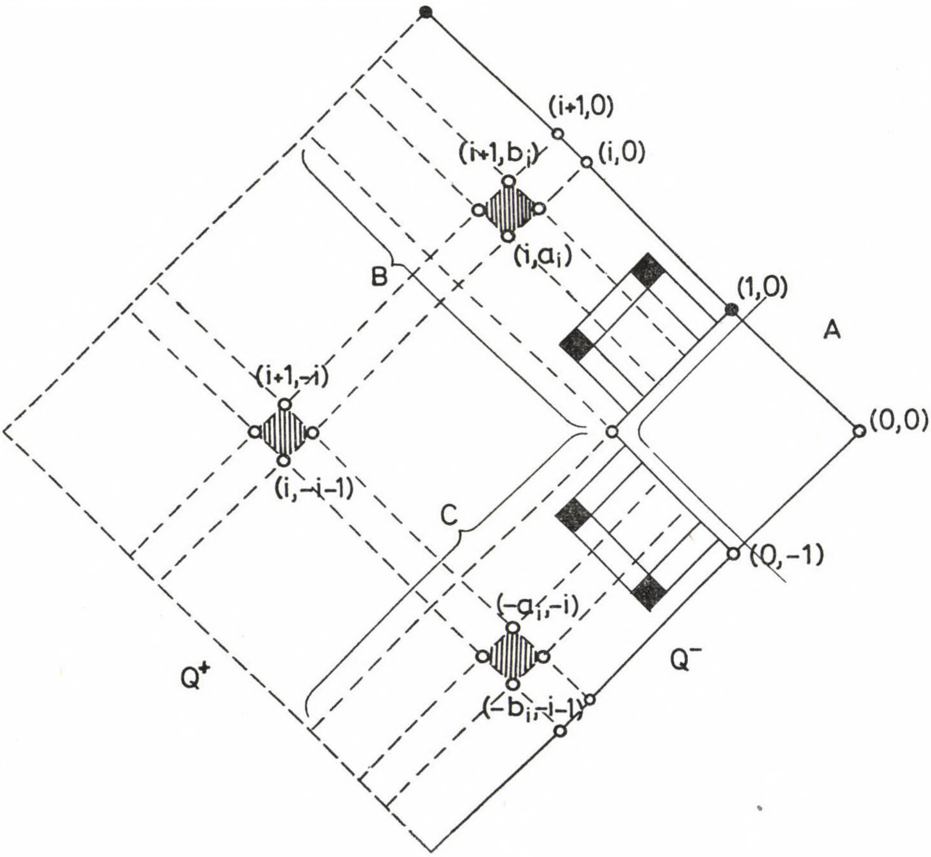


Fig. 3

The elements $\{(1, r);\ -1<r\leqq 0\}\cup\{(r, -1);\ 0\leqq r<1\}$ form a bounded countable chain which is dense-in-itself, hence we can identify this with the chain $\mathscr{C}$. If φ is the congruence-preserving function which is not order-preserving then we can assume: $\varphi((1, r))=(-r, -1)$. Then we can extend this φ to D: $\varphi(x, y)=(-y, -x)$.

Let M_3 be the five-element non-distributive modular lattice and let R be the bounded chain of rationals. Then there exists a modular lattice $M_3[R]$ having the following properties:

(a) $M_3[R]$ contains a $\{0, 1\}$-sublattice $\{0, a_1, a_2, a_3, 1\}$ isomorphic to M_3;

(b) the interval $[0, a_1]$ is isomorphic to R.

This lattice is determined up to isomorphism (see [4]).

An important property of $M_3[R]$ is that Con $(M_3[R])$ is isomorphic to Con (R). The intervals $[0, a_1]$ and $[0, a_3]$ are projective.

On $\mathscr{C}$ we have two different types of operations. Let $r=[0, 0']$. Take the operation of K_0, $\varphi_r^1(K_1)$, $\varphi_r^2\varphi_r^1(K_2)$, ... These are countable many unary operations therefore these may be enumerated, as $f_1, f_2, \ldots$. Let us assume that the corresponding interval isomorphism is

$$f_i\colon [a_i, b_i]\to[-b_i, -a_i].$$

To i we can associate three intervals of D

$$I_{i1} = [(-b_i, -2(i+1)),\ (-a_i, 2i)]$$

$$I_{i2} = [(2i, -2(i+1)),\ (2(i+1), -2i)]$$

$$I_{i3} = [(2i, ai),\ (2(i+1), b_i)].$$

All these intervals are isomorphic to $R\times R$.

The other type of the operations are the operations of the chains K_j, $j>0$. These may be enumerated as $g_1, g_2, \ldots$. Let us assume that the corresponding interval isomorphism is

$$g_i\colon [u_i, v_i]\to[w_i, z_i].$$

Then we can assume that $u_i, v_i, w_i, z_i\geqq 0'\in K_0\subseteqq\mathscr{C}$. Let g_i be defined by

$$g_i'\colon [-v_i, -u_i]\to[-z_i, -w_i].$$

To each g_i (resp. g_i') we associate the following two intervals

$$J_{i1} = [(z_i+1, v_i),\ (z_i+2, u_i)]$$

$$J_{i2} = [(2i+1, z_i),\ (2i+1, w_i)].$$

Now, we change each I_{ik}, J_{ik} to the lattice $L_{ik}\cong M_3[R]$. The elements $0\leqq x\leqq a_1$, $0\leqq y\leqq a_3$ generates a sublattice of $M_3[R]$ isomorphic to I_{ik}, i.e., I_{ik} is a sublattice of L_{ik}. This technique was developed in [3]. On this way we get from D a lattice $L(\subseteqq D)$ in which the intervals $[a_i, b_i]$, $[-b_i, -a_i]\subseteqq C$ (resp. $[u_i, v_i]$, $[w_i, z_i]$) are projective, we say that this projectivity realize the functions f_i, g_i. Then Con $(L)\cong$ $\cong$ Con $(\mathscr{C})$ hence L satisfies the three conditions (i)—(iii).

REFERENCES

[1] Dorninger, D. and Eigenthaler, G.: On compatible and order-preserving functions on lattices, Preprint, 1978.
[2] Schmidt, E. T.: On finitely generated simple modular lattices, *Period. Math. Hungar.* **6** (1975), 213—216.
[3] Schmidt, E. T.: Every finite distributive lattice is the congruence lattice of a modular lattice, *Algebra Universalis* **4** (1974), 49—57.
[4] Mitschke, A. und Wille, R.: Freie modulare Verbände $FM(M_3)$, *Proc. Univ. Houston,* Lattice Theory Conference, pp. 383—396.
[5] Schmidt, E. T. and Wille, R.: Note on compatible operations on modular lattices, *Algebra Universalis* (to appear).

(Received January 31, 1980)

MTA MATEMATIKAI KUTATÓ INTÉZETE
REÁLTANODA U. 13—15
H—1053 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 145—167

ON ADDITIVE REPRESENTATIONS OF INTEGERS I

by

A. SÁRKÖZY

1. Throughout this paper, we write $e^{2\pi i\alpha}=e(\alpha)$.

The Hardy—Littlewood method is often used in order to show that a certain sequence A of positive integers $a_1<a_2<\dots$ is *asymptotic basis of order* k, i.e., the equation

$$a_{x_1}+a_{x_2}+\dots+a_{x_k}=N \tag{1}$$

can be solved for $N>N_0$. In all these applications, we use the following facts:

(A) We have some information concerning the accounting function

$$A(u)=\sum_{a_i\leqq u}1.$$

(In fact, e.g., the condition $A(N)-A(N/k-1)>0$ is necessary for the solvability of (1).)

(B) For small moduli q, say for $1\leqq q\leqq P$, and for $0\leqq h<q$, $1\leqq u\leqq N$, we have an approximating formula of the type

$$\sum_{\substack{a_i\leqq u\\ a_i\equiv h\,(\mathrm{mod}\,q)}}1\sim m(q,h)\sum_{a_i\leqq u}1. \tag{2}$$

This approximation must be such that the absolute value of the so called "singular series"

$$T=\sum_{q=1}^{P}\sum_{\substack{0\leqq b<q\\ (b,q)=1}}\left(\sum_{h=0}^{q-1}m(q,h)e(hb/q)\right)^k e(-Nb/q) \tag{3}$$

should be large, i.e., we have a lower bound of the form

$$|T|>\eta.$$

(C) Write

$$S_u(\alpha)=\sum_{a_i\leqq u}e(a_i\alpha).$$

Then for large moduli q, say for $P<q\leqq Q$, and for all b, u with $0\leqq b<q$, $(b,q)=1$, $1\leqq u\leqq N$, we have an upper bound for $|S_u(b/q)|$.

The aim of this paper is to analyze and generalize the most important applications of the Hardy—Littlewood method, by proving a theorem which says that possibly general conditions of the type (A)—(C) imply the solvability of (1). In fact, Sections 2, 3, 4 and 5 will be devoted to the proof of the following theorem:

THEOREM 1. *Let $P, Q, N, k, a_1, a_2, \dots, a_y$ be positive integers such that*

$$2P < Q \leqq N, \tag{4}$$

$$PQ \leqq N, \tag{5}$$

$$a_1 < a_2 < \dots < a_y \leqq N \tag{6}$$

and

$$k \geqq 3. \tag{7}$$

Let $d_1, d_2, \dots, d_y$ be non-negative real numbers such that $d_i \neq 0$ holds for at least one i with $1 \leqq i \leqq y$ and let us put

$$\max_{1 \leqq j \leqq y} d_j = \Delta,$$

$$D(u, q, h) = \sum_{\substack{a_i \leqq u \\ a_i \equiv h \pmod q}} d_i,$$

$$D(u) = D(u, 1, 0) = \sum_{a_i \leqq u} d_i,$$

$$S_u(\alpha) = \sum_{a_i \leqq u} d_i e(a_i \alpha),$$

$$S(\alpha) = S_N(\alpha) = \sum_{i=1}^{y} d_i e(a_i \alpha),$$

and for $j = 1, 2, \dots,$ denote the number of solutions of the equation

$$a_{x_1} + a_{x_2} + \dots + a_{x_j} = a_{z_1} + a_{z_2} + \dots + a_{z_j}, \quad 1 \leqq x_1, \dots, x_j, z_1, \dots, z_j \leqq y$$

by R_{2j}.

Let us assume the following conditions hold:

(i) *There exist real numbers K, A and a real function $F(u)$ (defined for $0 < u < +\infty$) such that*

$$|D(u) - F(u)| < K \quad \textit{for} \quad 1 \leqq u \leqq N, \tag{8}$$

$$K \geqq 1, \tag{9}$$

$F(u)$ has continuous second derivative in $(0, N]$,

$$F(1) = 1, \tag{10}$$

$$F'(N) > 0, \tag{11}$$

$$F''(u) \leqq 0 \quad \textit{for} \quad 1 \leqq u \leqq N \tag{12}$$

and

$$uF'(u) \leqq AF(u) \quad \textit{for} \quad 1 \leqq u \leqq N. \tag{13}$$

(ii) *There exist positive real numbers $L_1, L_2, \ldots, L_P$, and for $1 \leqq q \leqq P$, $0 \leqq h \leqq q$, there exist real numbers $m(q, h)$ such that*

$$\sum_{h=0}^{q-1} |D(u, q, h) - m(q, h) D(u)| < L_q \quad \textit{for} \quad 1 \leqq q \leqq P \textit{ and } 0 \leqq u < N. \tag{14}$$

Write

$$T(q, b) = \sum_{h=0}^{q-1} m(q, h) e(hb/q) \qquad (\textit{for } 1 \leqq q \leqq P \quad \textit{and all} \quad b)$$

$$T_q = \max_{\substack{0 \leqq b < q \\ (b,q)=1}} |T(q, b)| \qquad (\textit{for } 1 \leqq q \leqq P)$$

and

$$T = T^{(N)} = \sum_{q=1}^{P} \sum_{\substack{0 \leqq b < q \\ (b,q)=1}} (T(q, b))^k e(-Nb/q).$$

(iii) *There exists a real number H such that if b, q are integers satisfying $P < q \leqq Q$, $(b, q) = 1$, and $0 < u \leqq N$, then we have*

$$|S_u(b/q)| < H. \tag{15}$$

If all the conditions above hold then we have

$$\left| \sum_{\substack{1 \leqq x_1, \ldots, x_k \leqq y \\ a_{x_1} + \ldots + a_{x_k} = N}} \prod_{i=1}^{k} d_{x_i} - T \sum_{\substack{1 \leqq n_1, \ldots, n_k \leqq N \\ n_1 + \ldots + n_k = N}} \prod_{j=1}^{k} F'(n_j) \right| < \tag{16}$$

$$< \min_{1 \leqq j < k/2} \left(8 \frac{HN}{PQ} \right)^{k-2j} \Delta^{2j} R_{2j} + \sum_{q=1}^{P} 2^{3k+1} k (A+5)^k q (L_q + T_q K) \left\{ L_q^{k-1} N (qQ)^{-2} + \right.$$

$$\left. + T_q^{k-1} \left((K + F(N))^{k-1} N^{-1} + K^{k-1} N^k (qQ)^{-(k+1)} + N \int_{qQ}^{N} x^{-3} (F(x))^{k-1} dx \right) \right\} +$$

$$+ \sum_{q=1}^{P} 2(2A+1)^k q T_q^k \int_{2}^{qQ} x^{-2} (F(x))^k dx.$$

Note that we usually have $d_1 = \ldots = d_y = 1$ in the applications (and in this case, the first term on the left-hand side of (16) is equal to the number of solutions of (1)); however, e.g., in the case of Goldbach's problem, it is more convenient to put $d_i = \log a_i = \log p_i$.

In Section 6, we shall deduce from Theorem 1

THEOREM 2. *Assume that all the conditions* (4)—(15) *in Theorem* 1 *hold, furthermore, we have*

$$|T|(F'(N))^k\binom{N-1}{k-1} > \min_{1\le j<k/2}\left(8\frac{HN}{PQ}\right)^{k-2j}\Delta^{2j}R_{2j}+ \tag{17}$$

$$+\sum_{q=1}^{P}2^{3k+1}k(A+5)^k q(L_q+T_qK)\Big\{L_q^{k-1}N(qQ)^{-2}+$$

$$+T_q^{k-1}\Big((K+F(N))^{k-1}N^{-1}+K^{k-1}N^k(qQ)^{-(k+1)}+N\int_{qQ}^{N}x^{-3}\big(F(x)\big)^{k-1}dx\Big)\Big\}+$$

$$+\sum_{q=1}^{P}2(2A+1)^k qT_q^k\int_{2}^{qQ}x^{-2}\big(F(x)\big)^k dx.$$

Then N can be represented in the form (1).

In order to illustrate the applicability of our theorems, we show in Sections 7 and 8 that these theorems cover the Goldbach—Waring problem, by proving that for fixed s, the numbers p^s form an asymptotic basis (of order less than $5s^2\log s$ for large s). In the proof of this fact, we need, of course, also Vinogradov's and Hua's deep estimates (Lemmas 13—15 below); the point of our discussion is that our theorems provide a simple and quick way to deduce the final result from these estimates (and Theorems 1 and 2 can be used analogously also in other fields of the additive number theory). Furthermore, this application of our theorems shows that the great generality in the conditions of our theorems does not lead to a considerable loss in the final results.

One might like to seek for possibly general conditions which imply conditions (i), (ii) and (iii) of our Theorems but which can be verified more easily. This discussion will be carried out in the next parts of this series.

2. Proof of Theorem 1. Following the Hardy—Littlewood—Vinogradov method, we start out from the integral

$$J=\int_0^1\big(S(\alpha)\big)^k e(-N\alpha)\,d\alpha=\sum_{\substack{1\le x_1,\dots,x_k\le y\\ a_{x_1}+\dots+a_{x_k}=N}}\prod_{i=1}^{k}d_{x_i}. \tag{18}$$

If q is a positive integer, b an integer, then let $I_{q,b}$ denote the interval $\left[\frac{b}{q}-\frac{1}{qQ},\ \frac{b}{q}+\frac{1}{qQ}\right]$. Let us form all the intervals $I_{q,b}$ with $1\le q\le P$, $0\le b\le q-1$, $(b,q)=1$. (These intervals are the so called "basic intervals".) These intervals do not overlap. In fact, if $1\le q_1, q_2\le P$,

$$0\le b_1\le q_1-1,\ 0\le b_2\le q_2-1,\ (b_1,q_1)=(b_2,q_2)=1\quad\text{and}\quad\frac{b_1}{q_1}\ne\frac{b_2}{q_2}$$

then by (4), we have

$$\left|\frac{b_1}{q_1}-\frac{b_2}{q_2}\right|=\frac{|b_1q_2-q_1b_2|}{q_1q_2}\ge\frac{1}{q_1q_2}=\frac{Q}{Qq_1q_2}>\frac{2P}{Qq_1q_2}\ge\frac{q_1+q_2}{Qq_1q_2}=\frac{1}{q_1Q}+\frac{1}{q_2Q}$$

which yields that $I_{q_1,b_1}\cap I_{q_2,b_2}=\emptyset$.

Let

$$E = \left[-\frac{1}{Q},\, 1-\frac{1}{Q}\right] - \bigcup_{q=1}^{P} \bigcup_{\substack{0\leqq b\leqq q-1\\(b,q)=1}} I_{q,b}.$$

Then we have

$$J = \int_{-1/Q}^{1-1/Q} (S(\alpha))^k e(-N\alpha)\,d\alpha =$$

$$= \sum_{q=1}^{P} \sum_{\substack{0\leqq b\leqq q-1\\(b,q)=1}} \int_{I_{q,b}} (S(\alpha))^k e(-N\alpha)\,d\alpha + \int_E (S(\alpha))^k e(-N\alpha)\,d\alpha$$

hence

(19) $$\left| J - \sum_{q=1}^{P} \sum_{\substack{0\leqq b\leqq q-1\\(b,q)=1}} \int_{I_{q,b}} (S(\alpha))^k e(-N\alpha)\,d\alpha \right| =$$

$$= \left| J - \sum_{q=1}^{P} \sum_{\substack{0\leqq b\leqq q-1\\(b,q)=1}} J_{q,b} \right| = \left| \int_E (S(\alpha))^k e(-N\alpha)\,d\alpha \right| \leqq \int_E |S(\alpha)|^k\,d\alpha = J^*$$

where

$$J_{q,b} = \int_{I_{q,b}} (S(\alpha))^k e(-N\alpha)\,d\alpha = \int_{-1/qQ}^{+1/qQ} \left(S\left(\frac{b}{q}+\beta\right)\right)^k e\left(-N\left(\frac{b}{q}+\beta\right)\right) d\beta$$

and

$$J^* = \int_E |S(\alpha)|^k\,d\alpha.$$

Thus in order to prove (16), we have to estimate the integrals $J_{q,b}$, J^*.

3. In this section, we estimate the integrals $J_{q,b}$ (i.e., the contribution of the "basic intervals").

LEMMA 1. *If q, b are integers such that $1\leqq q\leqq P$, $(b,q)=1$ and u is a real number satisfying $0<u\leqq N$ then we have*

$$|S_u(b/q) - T(q,b)F(u)| < L_q + T_q K.$$

PROOF. We have

$$S_u(b/q) = \sum_{1\leqq a_i\leqq u} d_i e(a_i b/q) = \sum_{h=0}^{q-1} \Big(\sum_{\substack{a_i\leqq u\\ a_i\equiv h\,(\mathrm{mod}\,q)}} d_i \Big) e(hb/q) = \sum_{h=0}^{q-1} D(u,q,h) e(hb/q)$$

thus by (8) and (14),

$$|S_u(b/q) - T(q,b)F(u)| \leqq |S_u(b/q) - T(q,b)D(u)| + |T(q,b)|\,|D(u)-F(u)| \leqq$$

$$\leqq \left| \sum_{h=0}^{q-1} D(u,q,h)e(hb/q) - D(u) \sum_{h=0}^{q-1} m(q,h)e(hb/q) \right| + T_q K =$$

$$= \left| \sum_{h=0}^{q-1} \big(D(u,q,h) - m(q,h)D(u)\big) e(hb/q) \right| + T_q K \leqq$$

$$\leqq \sum_{h=0}^{q-1} |D(u,q,h) - m(q,h)D(u)| + T_q K < L_q + T_q K$$

which proves Lemma 1.

LEMMA 2. *Let q, b be integers such that $1 \leqq q \leqq P$ and $(b, q)=1$, and let β be any real number. Let*

$$U(\beta) = \sum_{n=1}^{N} F'(n)e(n\beta). \tag{20}$$

Then we have

$$\left| S\left(\frac{b}{q}+\beta\right) - T(q, b)U(\beta)\right| < \begin{cases} 2(A+5)(L_q+T_qK) & \text{for } 0 \leqq |\beta| \leqq 1/N \\ 2(A+5)(L_q+T_qK)|\beta| N & \text{for } 1/N \leqq |\beta|. \end{cases}$$

PROOF. By Lemma 1, (9) and (10), and with respect to the inequality

$$|1-e(\beta)| \leqq 2\pi|\beta|, \tag{21}$$

we have

$$\left| S\left(\frac{b}{q}+\beta\right) - T(q, b) \sum_{n=2}^{N} (F(n)-F(n-1))e(n\beta)\right| = \tag{22}$$

$$= \left| \sum_{n=1}^{N} (S_n(b/q)-S_{n-1}(b/q))e(n\beta) - T(q, b) \sum_{n=2}^{N} (F(n)-F(n-1))e(n\beta)\right| =$$

$$= \left| \sum_{n=1}^{N} S_n(b/q)\big(e(n\beta)-e((n+1)\beta)\big) + S_N(b/q)e\big((N+1)\beta\big) - \right.$$

$$-T(q, b) \sum_{n=1}^{N} F(n)\big(e(n\beta)-e((n+1)\beta)\big) - T(q, b)F(N)e\big((N+1)\beta\big) +$$

$$\left. + T(q, b)F(1)e(\beta)\right| =$$

$$= \left| \sum_{n=1}^{N} \big(S_n(b/q)-T(q, b)F(n)\big)\big(e(n\beta)-e((n+1)\beta)\big) + \right.$$

$$\left. + \big(S_N(b/q)-T(q, b)F(N)\big)e\big((N+1)\beta\big) + T(q, b)e(\beta)\right| \leqq$$

$$\leqq \sum_{n=1}^{N} |S_n(b/q)-T(q, b)F(n)|\,|1-e(\beta)| + |S_N(b/q)-T(q, b)F(n)| + |T(q, b)| \leqq$$

$$\leqq \sum_{n=1}^{N} (L_q+T_qK)2\pi|\beta| + (L_q+T_qK) + T_q <$$

$$< (L_q+T_qK)2\pi|\beta|N + 2(L_q+T_qK) = 2(L_q+T_qK)(1+\pi|\beta|N).$$

Furthermore, by (10), (11), (12) and (13), we have

$$\left|\sum_{n=2}^{N}(F(n)-F(n-1))e(n\beta)-U(\beta)\right| = \tag{23}$$

$$= \left|\sum_{n=2}^{N}(F(n)-F(n-1))e(n\beta)-\sum_{n=1}^{N}F'(n)\,e(n\beta)\right| =$$

$$= \left|\sum_{n=2}^{N}\Big(\int_{n-1}^{n}F'(u)\,du\Big)e(n\beta)-\sum_{n=2}^{N}\Big(\int_{n-1}^{n}F'(n)\,du\Big)\,e(n\beta)-F'(1)e(\beta)\right| =$$

$$= \left|\sum_{n=2}^{N}\Big(\int_{n-1}^{n}(F'(u)-F'(n))\,du\Big)\,e(n\beta)-F'(1)\,e(\beta)\right| \leqq$$

$$\leqq \sum_{n=2}^{N}\int_{n-1}^{n}|F'(u)-F'(n)|\,du+F'(1) =$$

$$= \sum_{n=2}^{N}\int_{n-1}^{n}(F'(u)-F'(n))\,du+F'(1) \leqq$$

$$\leqq \sum_{n=2}^{N}\int_{n-1}^{n}(F'(n-1)-F'(n))\,du+F'(1) = \sum_{n=2}^{N}(F'(n-1)-F'(n))+F'(1) =$$

$$= 2F'(1)-F'(N) < 2F'(1) \leqq 2AF(1) = 2A.$$

(22) and (23) yield that

$$\left|S\left(\frac{b}{q}+\beta\right)-T(q,\,b)\,U(\beta)\right| \leqq \left|S\left(\frac{b}{q}+\beta\right)-T(q,\,b)\sum_{n=2}^{N}(F(n)-F(n-1))e(n\beta)\right| +$$

$$+|T(q,\,b)|\left|\sum_{n=2}^{N}(F(n)-F(n-1))e(n\beta)-U(\beta)\right| <$$

$$< 2(L_q+T_qK)(1+\pi|\beta|\,N)+T_q\cdot 2A <$$

$$< 2(L_q+T_qK)(1+A+\pi|\beta|\,N) < \begin{cases} 2(A+5)(L_q+T_qK) & \text{for } \ 0 \leqq |\beta| \leqq 1/N \\ 2(A+5)(L_q+T_qK)|\beta|N & \text{for } \ 1/N \leqq |\beta| \end{cases}$$

which proves Lemma 2.

LEMMA 3. *If $U(\beta)$ is defined by* (20) *then we have*

$$|U(\beta)| < \begin{cases} (A+1)\,F(N) & \textit{for } \ 0 \leqq |\beta| \leqq 1/N \\ (2A+1)\,F(|\beta|^{-1}) & \textit{for } \ 1/N \leqq |\beta| \leqq 1/2. \end{cases} \tag{24}$$

PROOF. If $1 \leqq v \leqq N$ then with respect to (10), (11), (12) and (13) we have

$$\sum_{1\leqq n\leqq v}F'(n) = \sum_{n=2}^{[v]}\int_{n-1}^{n}F'(n)\,du+F'(1) \leqq \tag{25}$$

$$\leqq \sum_{n=2}^{[v]}\int_{n-1}^{n}F'(u)\,du+F'(1) = \int_{1}^{[v]}F'(u)\,du+F'(1) =$$

$$= F([v])-F(1)+F'(1) < F(v)+F'(1) \leqq F(v)+AF(1) \leqq (A+1)\,F(v).$$

Furthermore, we have

$$(26)\qquad |U(\beta)| = \left|\sum_{n=1}^{N} F'(n)\, e(n\beta)\right| \leqq \sum_{n=1}^{N} |F'(n)| = \sum_{n=1}^{N} F'(n)$$

(again by (11) and (12)). (25) (with $v=N$) and (26) yield the first inequality in (24) (for all β).

For $n=0, 1, 2, \dots$, let us write

$$G_n(\beta) = \sum_{j=0}^{n} e(j\beta).$$

Then for all n and $0<|\beta|\leqq 1/2$, we have

$$(27)\qquad |G_n(\beta)| = \left|\frac{1-e((n+1)\beta)}{1-e(\beta)}\right| \leqq \frac{1+|e((n+1)\beta)|}{|1-e(\beta)|} =$$

$$= \frac{2}{|e(-\beta/2)-e(\beta/2)|} = \frac{1}{\sin|\pi\beta|} \leqq \frac{1}{\frac{2}{\pi}\cdot\pi|\beta|} = \frac{1}{2|\beta|}.$$

Using (11), (12), (13), (25) and (27), we obtain for $1/N\leqq|\beta|\leqq 1/2$ that

$$|U(\beta)| = \left|\sum_{n=1}^{N} F'(n)\, e(n\beta)\right| \leqq \left|\sum_{1\leqq n\leqq|\beta|^{-1}} F'(n) e(n\beta)\right| + \left|\sum_{|\beta|^{-1}<n\leqq N} F'(n) e(n\beta)\right| \leqq$$

$$\leqq \sum_{1\leqq n\leqq|\beta|^{-1}} |F'(n)| + \left|\sum_{|\beta|^{-1}<n\leqq N} F'(n)\big(G_n(\beta)-G_{n-1}(\beta)\big)\right| =$$

$$= \sum_{1\leqq n\leqq|\beta|^{-1}} F'(n) + \Big|\sum_{|\beta|^{-1}<n\leqq N-1} \big(F'(n)-F'(n+1)\big) G_n(\beta) -$$

$$- F'([|\beta|^{-1}]+1)\, G_{[|\beta|^{-1}]}(\beta) + F'(N)\, G_N(\beta)\Big| \leqq$$

$$\leqq (A+1)\, F(|\beta|^{-1}) + \sum_{|\beta|^{-1}<n\leqq N-1} |F'(n)-F'(n+1)|\, |G_n(\beta)| +$$

$$+ F'([|\beta|^{-1}]+1)\, |G_{[|\beta|^{-1}]}(\beta)| + F'(N)\, |G_N(\beta)| \leqq$$

$$\leqq (A+1)\, F(|\beta|^{-1}) + \sum_{|\beta|^{-1}<n\leqq N-1} \big(F'(n)-F'(n+1)\big)\frac{1}{2|\beta|} +$$

$$+ F'([|\beta|^{-1}]+1)\frac{1}{2|\beta|} + F'(N)\frac{1}{2|\beta|} =$$

$$= (A+1)\, F(|\beta|^{-1}) + 2F'([|\beta|^{-1}]+1)\frac{1}{2|\beta|} \leqq$$

$$\leqq (A+1)\, F(|\beta|^{-1}) + F'(|\beta|^{-1})\, |\beta|^{-1} \leqq (A+1)\, F(|\beta|^{-1}) + AF(|\beta|^{-1}) =$$

$$= (2A+1)\, F(|\beta|^{-1})$$

which completes the proof of (24).

LEMMA 4. *Let q, b be integers such that $1 \leqq q \leqq P$ and $(b, q)=1$, and let β be any real number. Then we have*

$$\max\left\{\left|S\left(\frac{b}{q}+\beta\right)\right|,\ |T(q, b)U(\beta)|\right\} <$$

$$< \begin{cases} 2(A+5)\big(L_q+T_q(K+F(N))\big) & \text{for } 0 \leqq |\beta| \leqq 1/N \\ 2(A+5)\big(L_q+T_q(K|\beta|N+F(|\beta|^{-1}))\big) & \text{for } 1/N \leqq |\beta| \leqq 1/2. \end{cases}$$

PROOF. By Lemmas 2 and 3, we have

$$\max\left\{\left|S\left(\frac{b}{q}+\beta\right)\right|, |T(q, b)U(\beta)|\right\} \leqq$$

$$\leqq \left|S\left(\frac{b}{q}+\beta\right)-T(q, b)U(\beta)\right|+|T(q, b)U(\beta)| \leqq$$

$$\leqq \left|S\left(\frac{b}{q}+\beta\right)-T(q, b)U(\beta)\right|+T_q|U(\beta)| <$$

$$< \begin{cases} 2(A+5)(L_q+T_qK)+T_q(A+1)F(N) & \text{for } 0 \leqq |\beta| \leqq 1/N \\ 2(A+5)(L_q+T_qK)|\beta|N+T_q(2A+1)F(|\beta|^{-1}) & \text{for } 1/N \leqq |\beta| \leqq 1/2 \end{cases}$$

hence (with respect to (9))

$$\max\left\{\left|S\left(\frac{b}{q}+\beta\right)\right|, |T(q, b)U(\beta)|\right\} < \tag{28}$$

$$< 2(A+5)(L_q+T_qK)+T_q(A+1)F(N) < 2(A+5)\big(L_q+T_q(K+F(N))\big)$$

for $0 \leqq |\beta| \leqq 1/N$ and

$$\max\left\{\left|S\left(\frac{b}{q}+\beta\right)\right|, |T(q, b)U(\beta)|\right\} < \tag{29}$$

$$< 2(A+5)(L_q+T_qK)|\beta|N+T_q(2A+1)F(|\beta|^{-1}) <$$

$$< 2(A+5)\big(L_q+T_q(K|\beta|N+F(|\beta|^{-1}))\big)$$

for $1/N \leqq |\beta| \leqq 1/2$. (28) and (29) yield the desired inequality.

LEMMA 5. *If q, b are integers such that $1 \leqq q \leqq P$, $(b, q)=1$ then we have*

$$\left|J_{q,b}-(T(q, b))^k e(-Nb/q) \int_{-1/2}^{+1/2} (U(\beta))^k e(-N\beta)\, d\beta\right| < \tag{30}$$

$$< 2^{3k+1}k(A+5)(L_q+T_qK)\Big\{L_q^{k-1}N(qQ)^{-2}+T_q^{k-1}\Big((K+F(N))^{k-1}N^{-1}+$$

$$+K^{k-1}N^k(qQ)^{-(k+1)}+N\int_{qQ}^{N} x^{-3}\big(F(x)\big)^{k-1}\,dx\Big)\Big\}+$$

$$+2(A+1)^kT_q^k\int_2^{qQ} x^{-2}\big(F(x)\big)^k\,dx.$$

PROOF. For any real numbers a, b, we have

$$|a^k-b^k| = |a-b|\left|\sum_{j=0}^{k-1} a^j b^{k-1-j}\right| \leqq k|a-b|(\max\{|a|, |b|\})^{k-1}.$$

Thus

$$\left|J_{q,b}-(T(q, b))^k e(-Nb/q)\int_{-1/2}^{+1/2}(U(\beta))^k e(-N\beta)\,d\beta\right| = \tag{31}$$

$$= \left|\int_{-1/qQ}^{+1/qQ}\left(S\left(\frac{b}{q}+\beta\right)\right)^k e\left(-N\left(\frac{b}{q}+\beta\right)\right)d\beta -\right.$$

$$\left.-\int_{-1/2}^{+1/2}(T(q, b)U(\beta))^k e\left(-N\left(\frac{b}{q}+\beta\right)\right)d\beta\right| =$$

$$= \left|\int_{-1/qQ}^{+1/qQ}\left(\left(S\left(\frac{b}{q}+\beta\right)\right)^k-(T(q, b)U(\beta))^k\right)e\left(-N\left(\frac{b}{q}+\beta\right)\right)d\beta -\right.$$

$$\left.-\int_{1/qQ\leqq|\beta|\leqq 1/2}(T(q, b)U(\beta))^k e\left(-N\left(\frac{b}{q}+\beta\right)\right)d\beta\right| \leqq$$

$$\leqq \int_{-1/qQ}^{+1/qQ}\left|\left(S\left(\frac{b}{q}+\beta\right)\right)^k-(T(q, b)U(\beta))^k\right|d\beta +$$

$$+\int_{1/qQ\leqq|\beta|\leqq 1/2}|T(q, b)|^k|U(\beta)|^k\,d\beta \leqq$$

$$\leqq k\int_{-1/qQ}^{+1/qQ}\left|S\left(\frac{b}{q}+\beta\right)-T(q, b)U(\beta)\right|\left(\max\left\{\left|S\left(\frac{b}{q}+\beta\right)\right|, |T(q, b)U(\beta)|\right\}\right)^{k-1}d\beta +$$

$$+T_q^k\int_{1/qQ\leqq|\beta|\leqq 1/2}|U(\beta)|^k\,d\beta.$$

By Lemmas 2 and 4, and with respect to (5) and the inequality

$$(a+b)^t = \sum_{j=0}^{t}\binom{t}{j}a^j b^{t-j} \leqq \sum_{j=0}^{t}\binom{t}{j}\max\{|a|^t, |b|^t\} \leqq$$

$$\leqq \sum_{j=0}^{t}\binom{t}{j}(|a|^t+|b|^t) = 2^t(|a|^t+|b|^t),$$

we have

$$(32)\quad \int_{-1/qQ}^{+1/qQ}\left|S\left(\frac{b}{q}+\beta\right)-T(q,b)U(\beta)\right|\left(\max\left\{\left|S\left(\frac{b}{q}+\beta\right)\right|,\ |T(q,b)U(\beta)|\right\}\right)^{k-1}d\beta<$$

$$<2\int_0^{1/N}2(A+5)(L_q+T_qK)\{2(A+5)(L_q+T_q(K+F(N)))\}^{k-1}\,d\beta+$$

$$+2\int_{1/N}^{1/qQ}2(A+5)(L_q+T_qK)\beta N\{2(A+5)(L_q+T_q(K\beta N+F(\beta^{-1})))\}^{k-1}\,d\beta<$$

$$<2^{2k}(A+5)^k(L_q+T_qK)(L_q^{k-1}+T_q^{k-1}(K+F(N))^{k-1})N^{-1}+$$

$$+2^{2k}(A+5)^k(L_q+T_qK)N\int_{1/N}^{1/qQ}\beta\{L_q^{k-1}+T_q^{k-1}(K\beta N+F(\beta^{-1}))^{k-1}\}\,d\beta<$$

$$<2^{3k}(A+5)(L_q+T_qK)\{(L_q^{k-1}+T_q^{k-1}(K+F(N))^{k-1})N^{-1}+$$

$$+NL_q^{k-1}\int_0^{1/qQ}\beta\,d\beta+T_q^{k-1}K^{k-1}N^k\int_0^{1/qQ}\beta^k\,d\beta+$$

$$+NT_q^{k-1}\int_{1/N}^{1/qQ}\beta(F(\beta^{-1}))^{k-1}\,d\beta\}<$$

$$<2^{3k}(A+5)^k(L_q+T_qK)\{(L_q^{k-1}+T_q^{k-1}(K+F(N))^{k-1})N^{-1}+NL_q^{k-1}(qQ)^{-2}+$$

$$+T_q^{k-1}K^{k-1}N^k(qQ)^{-(k+1)}+NT_q^{k-1}\int_{qQ}^{N}x^{-3}(F(x))^{k-1}\,dx\}=$$

$$=2^{3k}(A+5)^k(L_q+T_qK)\{L_q^{k-1}N(qQ)^{-2}((qQ/N)^2+1)+$$

$$+T_q^{k-1}((K+F(N))^{k-1}N^{-1}+K^{k-1}N^k(qQ)^{-(k+1)}+N\int_{qQ}^{N}x^{-3}(F(x))^{k-1}\,dx)\}<$$

$$<2^{3k+1}(A+5)^k(L_q+T_qK)\{L_q^{k-1}N(qQ)^{-2}+T_q^{k-1}((K+F(N))^{k-1}N^{-1}+$$

$$+K^{k-1}N^k(qQ)^{-(k+1)}+N\int_{qQ}^{N}x^{-3}(F(x))^{k-1}\,dx)\}$$

and by Lemma 3,

$$(33)\quad \int\limits_{1/qQ\leqq|\beta|\leqq 1/2}|U(\beta)|^k\,d\beta<2\int_{1/qQ}^{1/2}((2A+1)F(\beta^{-1}))^k\,d\beta=$$

$$=2(2A+1)^k\int_2^{qQ}x^{-2}(F(x))^k\,dx.$$

(31), (32) and (33) yield (30) and this completes the proof of Lemma 5.

4. In this section, we estimate the integral J^* (i.e., the contribution of the so called "supplementary intervals").

LEMMA 6. *If* $\alpha \in E$ *then we have*

$$|S(\alpha)| < 8\frac{HN}{PQ}.$$

PROOF. By Dirichlet's theorem, there exist integers q, b such that $1 \leqq q \leqq Q$, $(b, q) = 1$ and

$$\left|\alpha - \frac{b}{q}\right| < \frac{1}{qQ}. \tag{34}$$

By the definition of the set E, $\alpha \in E$ implies that

$$P < q. \tag{35}$$

Let us write $\alpha - \dfrac{b}{q} = \beta$. Then by (5), (15), (21), (34) and (35), we have

$$|S(\alpha)| = \left|S_N\left(\frac{b}{q}+\beta\right)\right| = \left|\sum_{n=1}^{N}\left(S_n\left(\frac{b}{q}\right) - S_{n-1}\left(\frac{b}{q}\right)\right)e(n\beta)\right| =$$

$$= \left|\sum_{n=1}^{N} S_n\left(\frac{b}{q}\right)\bigl(e(n\beta) - e((n+1)\beta)\bigr) + S_N\left(\frac{b}{q}\right)e\bigl((N+1)\beta\bigr)\right| \leqq$$

$$\leqq \sum_{n=1}^{N}\left|S_n\left(\frac{b}{q}\right)\right| |1 - e(\beta)| + \left|S_N\left(\frac{b}{q}\right)\right| \leqq \sum_{n=1}^{N} H\cdot 2\pi|\beta| + H =$$

$$= H(1+2\pi N|\beta|) < H\left(1+2\pi N\frac{1}{qQ}\right) < H\left(1+\frac{2\pi N}{PQ}\right) \leqq H\left(\frac{N}{PQ}+\frac{2\pi N}{PQ}\right) < 8\frac{HN}{PQ}$$

which proves Lemma 6.

LEMMA 7. *We have*

$$|J^*| < \min_{1 \leqq j < k/2}\left\{\left(8\frac{HN}{PQ}\right)^{k-2j} \Delta^{2j} R_{2j}\right\}. \tag{36}$$

PROOF. If j is an integer satisfying $1 \leqq j < k/2$ then by Lemma 6, we have

$$|J^*| = \int_E |S(\alpha)|^k \, d\alpha = \int_E |S(\alpha)|^{2j} |S(\alpha)|^{k-2j} \, d\alpha \leqq$$

$$\leqq \int_E |S(\alpha)|^{2j} \bigl(\max_{\alpha\in E} |S(\alpha)|\bigr)^{k-2j} \, d\alpha \leqq$$

$$\leqq \bigl(\max_{\alpha\in E} |S(\alpha)|\bigr)^{k-2j} \int_0^1 |S(\alpha)|^{2j} \, d\alpha <$$

$$< \left(8\frac{HN}{PQ}\right)^{k-2j} \int_0^1 \Bigl|\sum_{1\leqq x\leqq y} d_x e(a_x\alpha)\Bigr|^{2j} \, d\alpha =$$

$$= \left(8\frac{HN}{PQ}\right)^{k-2j} \sum_{\substack{a_{x_1}+\ldots+a_{x_j}-a_{z_1}-\ldots-a_{z_j}=0 \\ 1\le x_1,\ldots,x_j,\, z_1,\ldots,z_j\le y}} d_{x_1}\ldots d_{x_j} d_{z_1}\ldots d_{z_j} \le$$

$$\le \left(8\frac{HN}{PQ}\right)^{k-2j} \sum_{\substack{a_{x_1}+\ldots+a_{x_j}=a_{z_1}+\ldots+a_{z_j} \\ 1\le x_1,\ldots,x_j,\, z_1,\ldots,z_j\le y}} \Delta^{2j} =$$

$$= \left(8\frac{HN}{PQ}\right)^{k-2j} \Delta^{2j} R_{2j}.$$

This holds for all j (with $1\le j<k/2$) which proves (36).

5. In this section, we complete the proof of Theorem 1.

By (18), (19) and Lemmas 5 and 7, we have

$$\left| \sum_{\substack{1\le x_1,\ldots,x_k\le y \\ a_{x_1}+\ldots+a_{x_k}=N}} \left(\prod_{i=1}^{k} d_{x_i}\right) - T \sum_{\substack{1\le n_1,\ldots,n_k\le N \\ n_1+\ldots+n_k=N}} \left(\prod_{j=1}^{k} F'(n_j)\right)\right| =$$

$$= \left| J - \left(\sum_{q=1}^{P} \sum_{\substack{0\le b<q \\ (b,q)=1}} (T(q,b))^k e(-Nb/q)\right) \int_{-1/2}^{+1/2} (U(\beta))^k e(-N\beta)\, d\beta \right| =$$

$$= \left| \left(J - \sum_{q=1}^{P} \sum_{\substack{0\le b\le q-1 \\ (b,q)=1}} J_{q,b}\right) + \right.$$

$$\left. + \sum_{q=1}^{P} \sum_{\substack{0\le b\le q-1 \\ (b,q)=1}} \left(J_{q,b} - (T(q,b))^k e(-Nb/q) \int_{-1/2}^{+1/2} (U(\beta))^k e(-N\beta)\, d\beta\right) \right| \le$$

$$\le \left| J - \sum_{q=1}^{P} \sum_{\substack{0\le b\le q-1 \\ (b,q)=1}} J_{q,b} \right| +$$

$$+ \sum_{q=1}^{P} \sum_{\substack{0\le b\le q-1 \\ (b,q)=1}} \left| J_{q,b} - (T(q,b))^k e(-Nb/q) \int_{-1/2}^{+1/2} (U(\beta))^k e(-N\beta)\, d\beta \right| \le$$

$$\le J^* + \sum_{q=1}^{P} \sum_{\substack{0\le b\le q-1 \\ (b,q)=1}} \left| J_{q,b} - (T(q,b))^k e(-Nb/q) \int_{-1/2}^{+1/2} (U(\beta))^k e(-N\beta)\, d\beta \right| <$$

$$< \min_{1\le j<k/2} \left(8\frac{HN}{PQ}\right)^{k-2j} \Delta^{2j} R_{2j} + \sum_{q=1}^{P} 2^{3k+1} k(A+5)^k q(L_q + T_q K)\Big\{L_q^{k-1} N(qQ)^{-2} +$$

$$+ T_q^{k-1}\Big((K+F(N))^{k-1} N^{-1} + K^{k-1} N^k (qQ)^{-(k+1)} + N \int_{qQ}^{N} x^{-3}(F(x))^{k-1}\, dx\Big)\Big\} +$$

$$+ \sum_{q=1}^{P} 2(2A+1)^k q T_q^k \int_{2}^{qQ} x^{-2}(F(x))^k\, dx$$

which completes the proof of Theorem 1.

6. In this section, we deduce Theorem 2 from Theorem 1.

PROOF of Theorem 2.

In order to show the solvability of (1), it is sufficient to show that the first sum on the left-hand side of (16) is different from 0:

$$\sum_{\substack{1<x_1,\dots,x_k\le y\\ a_{x_1}+\dots+a_{x_k}=N}} \prod_{i=1}^{k} d_{x_i} \ne 0.$$

By (16) in Theorem 1, this would follow from

$$(37)\qquad \left|T \sum_{\substack{1\le n_1,\dots,n_k\le N\\ n_1+\dots+n_k=N}} \prod_{j=1}^{k} F'(n_j)\right| > \min_{1\le j<k/2}\left(8\frac{HN}{PQ}\right)^{k-2j} \Delta^{2j} R_{2j}+$$

$$+\sum_{q=1}^{P} 2^{3k+1} k(A+5)^k q(L_q+T_qK)\Big\{L_q^{k-1}N(qQ)^{-2}+$$

$$+T_q^{k-1}\Big(\big(K+F(N)\big)^{k-1}N^{-1}+K^{k-1}N^k(qQ)^{-(k+1)}+N\int_{qQ}^{N} x^{-3}\big(F(x)\big)^{k-1}dx\Big)\Big\}+$$

$$+\sum_{q=1}^{P} 2(2A+1)^k qT_q^k \int_{2}^{qQ} x^{-2}\big(F(x)\big)^k dx.$$

By (11) and (12), we have

$$(38)\qquad \left|T \sum_{\substack{1\le n_1,\dots,n_k\le N\\ n_1+\dots+n_k=N}} \prod_{j=1}^{k} F'(n_j)\right| \ge |T| \sum_{\substack{1\le n_1,\dots,n_k\le N\\ n_1+\dots+n_k=N}} \prod_{j=1}^{k} F'(N) =$$

$$= |T|\big(F'(N)\big)^k \sum_{\substack{1\le n_1,\dots,n_k\le N\\ n_1+\dots+n_k=N}} 1.$$

It is well-known that the number of solutions of the equation

$$n_1+n_2+\dots+n_k = N, \quad 1\le n_1, n_2, \dots, n_k \le N$$

is equal to $\binom{N-1}{k-1}$ (see e.g. [2], p. 36) thus we obtain from (38) that

$$(39)\qquad \left|T \sum_{\substack{1\le n_1,\dots,n_k\le N\\ n_1+\dots+n_k=N}} \prod_{j=1}^{k} F'(n_j)\right| \ge |T|\big(F'(N)\big)^k \binom{N-1}{k-1}.$$

(17) and (39) yield (37) and this completes the proof of Theorem 2.

7. In Sections 7 and 8, we show that Theorems 1 and 2 cover the Goldbach—Waring problem.

COROLLARY 1. *Let s be a positive integer, k a positive integer satisfying*

$$(40)\qquad k \ge \begin{cases} 2^s+1 & \textit{for}\quad 1\le s\le 11\\ s^2(4\log s+2\log\log s+5)-1 & \textit{for}\quad s\ge 12. \end{cases}$$

If p is a prime number satisfying $(p-1)/s$ then define the positive integer γ_p by

$$p^{\gamma_p - 1}|s,\ p^{\gamma_p}\nmid s \quad for \quad p \neq 2 \ or \ p = 2,\ 2\nmid s$$

and

$$2^{\gamma_2 - 2}|s,\ 2^{\gamma_2 - 1}\nmid s \quad for \quad 2|s,$$

and put

$$V = \prod_{(p-1)/s} p^{\gamma_p}. \tag{41}$$

Let N be a positive integer satisfying $N > N_0(k, s)$ and

$$N \equiv k \pmod V). \tag{42}$$

Then N can be represented in the form

$$q_1^s + q_2^s + \ldots + q_k^s = N$$

where $q_1, q_2, \ldots, q_k$ are prime numbers.

COROLLARY 2. *If s is a positive integer, k is a positive integer satisfying*

$$k \geqq \begin{cases} V + 2^s & for \quad 1 \leqq s \leqq 11 \\ V + s^2(4\log s + 2\log\log s + 5) - 2 & for \quad s \geqq 12 \end{cases}$$

(where V is defined by (41)) and $N > N_1(k, s)$ then N can be represented in the form

$$r_1^s + r_2^s + \ldots + r_l^s = N$$

where $r_1, r_2, \ldots, r_l$ are prime numbers and $l \leqq k$.

PROOF of Corollary 1.

Throughout the proof, we use the following notations: $c_1, c_2, \ldots$ denote positive absolute constants which may depend on certain parameters k, s, ε but which are independent of the parameter N. $\varphi(n)$ denotes Euler's phi function. We put

$$\psi(u, q, a) = \sum_{\substack{n \leqq u \\ n \equiv a \pmod q}} \Lambda(n)$$

and

$$\psi(u) = \psi(u, 1, 0) = \sum_{n \leqq u} \Lambda(n).$$

Then by the prime number theorem (or a more elementary theorem) we have

$$\Big|\psi(u, q, a) - \sum_{\substack{p \leqq u \\ p \equiv a \pmod q}} \log p\Big| = \Big|\sum_{\substack{p^\alpha \leqq u,\ \alpha \geqq 2 \\ p^\alpha \equiv a \pmod q}} \log p^\alpha\Big| < c_1 u^{1/2} \tag{43}$$

for all u, q and a.

If q, N are positive integers and b is an integer then we write

$$W(q, b) = \sum_{\substack{0 \leqq h < q \\ (h, q) = 1}} e(bh^s/q)$$

and

$$U^{(N)} = \sum_{q=1}^{+\infty} \left(\sum_{\substack{0 \leqq b < q \\ (b,q)=1}} \left(\frac{W(q, b)}{\varphi(q)} \right)^k e(-bN/q) \right).$$

We will apply Theorem 2 with the prime powers $p_1^s, p_2^s, \dots, p_{\pi(N^{1/s})}^s$ (where p_i denotes the i^{th} prime number) in place of the sequence $a_1, a_2, \dots, a_y$ and with $d_i = \log p_i$ so that

(44) $$a_i = p_i^s,$$

(45) $$y = \pi(N^{1/s}),$$

(46) $$\Delta = \max_{1 \leqq j \leqq \pi(N^{1/s})} d_j = \max_{p \leqq N^{1/s}} \log p \leqq \log N$$

and

$$S_u(\alpha) = \sum_{p \leqq u^{1/s}} \log p\, e(p^s \alpha),$$

$$S(\alpha) = S_N(\alpha) = \sum_{p \leqq N^{1/s}} \log p\, e(p^s \alpha).$$

In order to show that Theorem 2 can be applied, we need the following lemmas:

LEMMA 8. *There exist absolute constants c_2, c_3 such that for $x \geqq 2$, we have*

$$|\psi(x) - x| < c_2 x e^{-c_3 \sqrt{\log x}}.$$

This lemma is equivalent to the prime number theorem; see e.g. [3], pp. 61—70. We need also the Siegel—Walfisz theorem (see e.g. [3], p. 144):

LEMMA 9. *If A is an arbitrary large but fixed positive number then there exist absolute constants c_4, c_5 such that if $x \geqq 2$ and q, a are any integers satisfying $1 \leqq q \leqq (\log x)^A$ and $(a, q) = 1$ then we have*

$$\left| \psi(x, q, a) - \frac{x}{\varphi(q)} \right| < c_4 x e^{-c_5 \sqrt{\log x}}.$$

LEMMA 10. *We have*

$$\varphi(n) > c_6 \frac{n}{\log \log n}$$

for all $n \geqq 3$.

This lemma is well-known; see e.g. [3], p. 24.

LEMMA 11. *For all $\varepsilon > 0$, there exists an absolute constant $c_7 = c_7(s, \varepsilon)$ such that for $q = 1, 2, \dots$ and $(b, q) = 1$ we have*

$$|W(q, b)| < c_7 q^{1 - \frac{1}{s} + \varepsilon}.$$

See [1], p. 9.

LEMMA 12. *Let s, k be positive integers such that*

$$k \geqq 3s.$$

Define the integer V by (41) *and assume that the positive integer N satisfies* (42). *Then the infinite series $U^{(N)}$ is absolutely convergent and we have*

$$|U^{(N)}| \geqq c_8 = c_8(s, k) > 0$$

(where c_8 is independent of N).

See [1], p. 99—106.

LEMMA 13. *Let σ_1, σ_2 be any positive numbers satisfying*

$$\sigma_1 \geqq 2^{6s}(\sigma_2+2).$$

Then for $u > c_9(s, \sigma_1, \sigma_2)$,

$$(\log u)^{\sigma_1} < q \leqq u(\log u)^{-\sigma_1}$$

and

$$(b, q) = 1,$$

we have

$$\Big|\sum_{p \leqq u^{1/s}} e(bp^s/q)\Big| < u^{1/s}(\log u)^{-\sigma_2}.$$

For a more general form of this theorem, see [1], p. 65.

LEMMA 14. *For $s = 1, 2, \dots, u \geqq 2$ we have*

$$\int_0^1 \Big|\sum_{1 \leqq j \leqq u} e(j^s\alpha)\Big|^{2^s} d\alpha \leqq u^{2^s-s}(\log u)^{c_9(s)}.$$

See [1], p. 19.

LEMMA 15. *For $s \geqq 12$,*

$$j > s^2(2\log s + \log\log s + 2{,}5) - 2$$

and $u \geqq 1$ we have

$$\int_0^1 \Big|\sum_{1 \leqq j \leqq u} e(j^s\alpha)\Big|^{2j} d\alpha < c_{10}(s, j)u^{2j-s}.$$

See [1], p. 89.

8. Now we are ready to complete the proof of Corollary 1. Let us write

$$\varrho = \max\{1, c_9(1), c_9(2), \ldots, c_9(11)\}$$

where $c_9(s)$ denotes the constant defined in Lemma 14 and put $\sigma_2 = k + \varrho + 1$, $\sigma_1 = 2^{6s}(\sigma_2 + 2)$, $\sigma = \sigma_1 + s(\sigma_2 - 1)$. We apply Theorem 2 with

$$P = [(\log N)^\sigma] \quad \text{and} \quad Q = [N(\log N)^{-\sigma}]. \tag{47}$$

We have to show that if k satisfies (40), $N > N_0(k, s)$ and $P, Q, a_1, a_2, \ldots, a_y$ are defined by (44), (45) and (47), then all the conditions in Theorem 2 hold.

For large N, (4), (5) and (6) hold obviously. (47) implies (7) trivially.

Now we are going to show that for large N, (i) in Theorem 1 holds with

$$F(u) = u^{1/s}, \tag{48}$$

$$K = N^{1/s} e^{-c_{11}\sqrt{\log N}} \tag{49}$$

(where $c_{11} = c_{11}(s)$ is a small positive constant) and

$$A = 1. \tag{50}$$

In fact, by (43) and Lemma 8, we have

$$|D(u) - F(u)| = \Big|\sum_{p^s \leqq u} \log p - u^{1/s}\Big| \leqq \tag{51}$$

$$\leqq \Big|\sum_{p \leqq u^{1/s}} \log p - \psi(u^{1/s})\Big| + |\psi(u^{1/s}) - u^{1/s}| <$$

$$< c_1 u^{1/2s} + c_2 u^{1/s} e^{-c_3\sqrt{\log u^{1/s}}} < N^{1/s} e^{-c_{11}\sqrt{\log N}} = K$$

for $c_{11} = c_3/2s^{1/2}$, $u_0 = u_0(s) < u \leqq N$, and

$$|D(u) - F(u)| < N^{1/s} e^{-c_{11}\sqrt{\log N}} = K \tag{52}$$

holds trivially also for $1 \leqq u \leqq u_0$ and large N. (9) (for large N), (10), (11), (12) and (13) hold trivially.

Now we show that also (ii) in Theorem 1 holds with

$$m(q, h) = \frac{1}{\varphi(q)} \sum_{\substack{0 \leqq j < q \\ (j,q)=1 \\ j^s \equiv h \pmod{q}}} 1$$

and

$$L_q = N^{1/s} e^{-c_{12}\sqrt{\log N}}. \tag{53}$$

In fact, by (43), (48), (51), (52) and Lemma 9 (and since $q \leqq P \leqq (\log N)^\sigma$), we have

$$\sum_{h=0}^{q-1} |D(u, q, h) - m(q, h)D(u)| =$$

$$= \sum_{h=0}^{q-1} \left| \sum_{\substack{0 \leqq j < q \\ j^s \equiv h \pmod q}} \sum_{\substack{p \leqq u^{1/s} \\ p \equiv j \pmod q}} \log p - \sum_{\substack{0 \leqq j < q \\ (j, q) = 1 \\ j^s \equiv h \pmod q}} \frac{D(u)}{\varphi(q)} \right| \leqq$$

$$\leqq \sum_{h=0}^{q-1} \Bigg\{ \sum_{\substack{0 \leqq j < q \\ (j, q) > 1 \\ j^s \equiv h \pmod q}} \sum_{\substack{p \leqq u^{1/s} \\ p \equiv j \pmod q}} \log p + \sum_{\substack{0 \leqq j < q \\ (j, q) = 1 \\ j^s \equiv h \pmod q}} \Bigg(\Bigg| \sum_{\substack{p \leqq u^{1/s} \\ p \equiv j \pmod q}} \log p - \psi(u^{1/s}, q, j) \Bigg| +$$

$$+ \left| \psi(u^{1/s}, q, j) - \frac{u^{1/s}}{\varphi(q)} \right| + \frac{1}{\varphi(q)} |u^{1/s} - D(u)| \Bigg) \Bigg\} \leqq$$

$$\leqq \sum_{\substack{p \leqq u^{1/s} \\ (p, q) > 1}} \log p + \sum_{h=0}^{q=1} \sum_{\substack{0 \leqq j < q \\ (j, q) = 1 \\ j^s \equiv h \pmod q}} \left(c_1 u^{1/2s} + c_4 N^{1/s} e^{-c_5 \sqrt{\log N^{1/s}}} + \frac{1}{\varphi(q)} N^{1/s} e^{-c_{11} \sqrt{\log N}} \right) \leqq$$

$$\leqq \log \prod_{p \mid q} p + \sum_{\substack{0 \leqq j < q \\ (j, q) = 1}} N^{1/s} e^{-c_{13} \sqrt{\log N}} \leqq \log q + q N^{1/s} e^{-c_{13} \sqrt{\log N}} < N^{1/s} e^{-c_{14} \sqrt{\log N}}$$

if N is sufficiently large (in terms of s) which proves (14).

Furthermore, we have

$$T(q, b) = \sum_{h=0}^{q-1} m(q, h) e(hb/q) = \tag{54}$$

$$= \frac{1}{\varphi(q)} \sum_{h=0}^{q-1} \Big(\sum_{\substack{0 \leqq j < q \\ (j, q) = 1 \\ j^s \equiv h \pmod q}} 1 \Big) e(hb/q) = \frac{1}{\varphi(q)} \sum_{\substack{0 \leqq j < q \\ (j, q) = 1}} e(bj^s/q) = \frac{1}{\varphi(q)} W(q, b)$$

thus by Lemmas 10 and 11, we have

$$T_q = \max_{\substack{0 \leqq b < q \\ (b, q) = 1}} |T(q, b)| = \max_{\substack{0 \leqq b < q \\ (b, q) = 1}} \frac{|W(q, b)|}{\varphi(q)} < c_{15}(s, \varepsilon) q^{-1/s + \varepsilon}. \tag{55}$$

(40) implies trivially that

$$k \geqq 3s. \tag{56}$$

By (42), (54), (56) and Lemma 12, the infinite series

$$U^{(N)} = \sum_{q=1}^{+\infty} \Bigg\{ \sum_{\substack{0 \leqq b < q \\ (b, q) = 1}} \left(\frac{W(q, b)}{\varphi(q)} \right)^k e(-bN/q) \Bigg\} = \sum_{q=1}^{+\infty} \Big\{ \sum_{\substack{0 \leqq b < q \\ (b, q) = 1}} (T(q, b))^k e(-bN/q) \Big\}$$

is absolutely convergent and we have

$$|U^{(N)}| \geqq c_8(s, k) > 0. \tag{57}$$

By (47), (55) (with $\varepsilon=1/2k$), (56) and (57), for $N>N_0(s,k)$ we have

$$(58)\qquad |T| = |T^{(N)}| = \left|\sum_{q=1}^{P} \sum_{\substack{0\leq b<q\\(b,q)=1}} (T(q,b))^k e(-Nb/q)\right| \geqq$$

$$\geqq |U^{(N)}| - \sum_{q=P+1}^{+\infty} \sum_{\substack{0\leq b<q\\(b,q)=1}} T_q^k \geqq c_8 - \sum_{q=P+1}^{+\infty} qT_q^k >$$

$$> c_8 - \sum_{q=P+1}^{+\infty} c_{15}^k q^{\frac{3}{2}-\frac{k}{s}} > c_8/2 = c_{16} = c_{16}(s,k) > 0$$

(provided that (42) holds).

Now we are going to show that (15) holds (for $P<q\leqq Q$, $(b,q)=1$ and $0<u\leqq \leqq N$) with

$$(59)\qquad H = c_{17} N^{1/s} (\log N)^{-\sigma_2+1}.$$

Assume first that

$$2N(\log N)^{-s(\sigma_2-1)} < u \leqq N.$$

Then for large N, $P<q\leqq Q$ implies that

$$q > P \geqq (\log N)^{\sigma} - 1 > (\log u)^{\sigma_1}$$

and

$$q \leqq Q \leqq N(\log N)^{-\sigma} = 2N(\log N)^{-s(\sigma_2-1)} \cdot \frac{1}{2} (\log N)^{s(\sigma_2-1)-\sigma} =$$

$$= 2N(\log N)^{-s(\sigma_2-1)} \cdot \frac{1}{2} (\log N)^{-\sigma_1} < u (\log u)^{-\sigma_1}.$$

Thus Lemma 13 can be applied (in order to estimate $\Big|\sum_{p\leqq u^{1/s}} e(bp^s/q)\Big|$) and we obtain for large N that

$$(60)\quad |S_u(b/q)| = \Big|\sum_{p\leqq u^{1/s}} \log p\, e(bp^s/q)\Big| = \Big|\sum_{1\leqq n\leqq u^{1/s}} \log n \Big(\sum_{p\leqq n} e(bp^s/q) - \sum_{p\leqq n-1} e(bp^s/q)\Big)\Big| =$$

$$= \Big|\sum_{1\leqq n\leqq u^{1/s}} (\log n - \log(n+1)) \sum_{p\leqq n} e(bp^s/q) + \log([u^{1/s}]+1) \sum_{p\leqq u^{1/s}} e(bp^s/q)\Big| \leqq$$

$$\leqq \sum_{1\leqq n\leqq u^{1/s}} |\log n - \log(n+1)| \Big|\sum_{p\leqq n} e(bp^s/q)\Big| + \log([u^{1/s}]+1) \Big|\sum_{p\leqq u^{1/s}} e(bp^s/q)\Big| \leqq$$

$$\leqq \sum_{1\leqq n\leqq u^{1/s}} \frac{c_{18}}{n} \Big|\sum_{p\leqq n} e(bp^s/q)\Big| + c_{19} \log N \Big|\sum_{p\leqq u^{1/s}} e(bp^s/q)\Big| \leqq$$

$$\leqq c_{18} \sum_{1\leqq n\leqq (2N)^{1/s}(\log N)^{-\sigma_2+1}} \frac{1}{n} \sum_{p\leqq n} 1 + c_{18} \sum_{(2N)^{1/s}(\log N)^{-\sigma_2+1} < n\leqq u^{1/s}} \frac{1}{n} \Big|\sum_{p\leqq n} e(bp^s/q)\Big| +$$

$$+ c_{19} \log N \Big|\sum_{p\leqq u^{1/s}} e(bp^s/q)\Big| < c_{20} \sum_{2\leqq n\leqq (2N)^{1/s}(\log N)^{-\sigma_2+1}} \frac{1}{\log n} +$$

$$+ c_{21} \sum_{2\leqq n\leqq u^{1/s}} (\log n)^{-\sigma_2} + c_{19} \log N \cdot u^{1/s} (\log u)^{-\sigma_2} <$$

$$< c_{22} N^{1/s} (\log N)^{-\sigma_2} + c_{23} u^{1/s} (\log u)^{-\sigma_2} + c_{19} \log N \cdot u^{1/s} (\log u)^{-\sigma_2} <$$

$$< c_{24} N^{1/s} (\log N)^{-\sigma_2} + c_{25} N^{1/s} (\log N)^{-\sigma_2+1} < c_{26} N^{1/s} (\log N)^{-\sigma_2+1}$$

while for $1 \leqq u \leqq 2N(\log N)^{-s(\sigma_2-1)}$, we have trivially

$$(61)\quad |S_u(b/q)| = \Big|\sum_{p \leqq u^{1/s}} \log p\, e(bp^s/q)\Big| \leqq \sum_{p \leqq u^{1/s}} \log p < c_{27} u^{1/s} \leqq c_{28} N^{1/s} (\log N)^{-\sigma_2+1}.$$

(60) and (61) yield (15) (with the number H defined by (59)).
Finally, define $j_0 = j_0(s)$ by

$$(62)\qquad j_0 = \begin{cases} [s^2(2\log s + \log\log s + 2{,}5)] - 1 & \text{for} \quad s \geqq 12 \\ 2^{s-1} \quad \text{for} \quad s \leqq 11. \end{cases}$$

Then Lemmas 14 and 15 yield that for $N > N_0(s)$ we have

$$(63)\qquad R_{2j_0} = \sum_{\substack{p_{x_1}^s + \dots + p_{x_{j_0}}^s - p_{z_1}^s - \dots - p_{z_{j_0}}^s = 0 \\ 1 \leqq p_{x_1}, \dots, p_{x_{j_0}}, p_{z_1}, \dots, p_{z_{j_0}} \leqq N^{1/s}}} 1 \leqq \sum_{\substack{u_1^s + \dots + u_{j_0}^s - v_1^s - \dots - v_{j_0}^s = 0 \\ 1 \leqq u_1, \dots, u_{j_0}, v_1, \dots, v_{j_0} \leqq N^{1/s}}} 1 =$$

$$= \int_0^1 \Big|\sum_{1 \leqq l \leqq N^{1/s}} e(l^s \alpha)\Big|^{2j_0} d\alpha < (N^{1/s})^{2j_0 - s} (\log N^{1/s})^{\varrho} \leqq N^{2j_0/s - 1} (\log N)^{\varrho}.$$

(40) implies that $k > 2j_0$ thus by (46), (47), (49) and (63),

$$(64)\quad \min_{1 \leqq j < k/2} \left(8\frac{HN}{PQ}\right)^{k-2j} \Delta^{2j} R_{2j} \leqq \left(8\frac{HN}{PQ}\right)^{k-2j_0} \Delta^{2j_0} R_{2j_0} < (9H)^{k-2j_0} \Delta^{2j_0} R_{2j_0} <$$

$$< c_{29}^k \big(N^{1/s} (\log N)^{-\sigma_2+1}\big)^{k-2j_0} (\log N)^{2j_0} N^{2j_0/s-1} (\log N)^{\varrho} =$$

$$= c_{30}(k) N^{k/s-1} (\log N)^{-(k-2j_0)\sigma_2 + k + \varrho} \leqq$$

$$\leqq c_{30} N^{k/s-1} (\log N)^{-\sigma_2 + k + \varrho} = c_{30} N^{k/s-1} (\log N)^{-1}.$$

In order to complete the proof of Corollary 1, we have to show that (17) holds. By (48) and (58), we have

$$(65)\quad |T| (F'(N))^k \binom{N-1}{k-1} > c_{16} \left(\frac{1}{s} N^{1/s-1}\right)^k \frac{(N-1)(N-2)\dots(N-k+1)}{(k-1)!} >$$

$$> c_{31}(s, k) N^{k/s-k} N^{k-1} = c_{31} N^{k/s-1}$$

for $N > N_0(s, k)$.
On the other hand, by (48) and (56) we have

$$(66)\qquad \int_{qQ}^{N} x^{-3} (F(x))^{k-1} dx < \int_1^N x^{-3+(k-1)/s} dx =$$

$$= \int_1^N x^{(k-3s-1)/s} dx < c_{32}(s, k) N^{(k-3s-1)/s+1} \log N = c_{32}(s, k) N^{(k-1)/s-2} \log N$$

(the factor $\log N$ is needed in the case $s=1$, $k=3s=3$) and

$$(67)\qquad \int_2^{qQ} x^{-2}(F(x))^k\,dx = \int_2^{qQ} x^{-2+k/s}\,dx = \int_2^{qQ} x^{(k-2s)/s}\,ds <$$
$$< (qQ)^{(k-2s)/s+1} = (qQ)^{k/s-1}.$$

Thus by (4), (5), (47), (48), (49), (50), (53), (55) (with $\varepsilon=1/2k$), (56), (64), (66) and (67), the right-hand side of (16) can be estimated in the following way:

$$(68)\qquad \min_{1\le j<k/2}\left(8\frac{HN}{PQ}\right)^{k-2j} \Delta^{2j}R_{2j} + \sum_{q=1}^{P} 2^{3k+1}k(A+5)^k q(L_q+T_qK)\left\{L_q^{k-1}N(qQ)^{-2}+\right.$$
$$+T_q^{k-1}\left((K+F(N))^{k-1}N^{-1}+K^{k-1}N^k(qQ)^{-(k+1)}+N\int_{qQ}^{N} x^{-3}(F(x))^{k-1}\,dx\right)\Big\}+$$
$$+\sum_{q=1}^{P} 2(2A+1)^k qT_q^k \int_2^{qQ} x^{-2}(F(x))^k\,dx < c_{33}(s,k)\{N^{k/s-1}(\log N)^{-1}+$$
$$+\sum_{q=1}^{P} qN^{1/s}e^{-c_{34}\sqrt{\log N}}\{N^{(k-1)/s}e^{-c_{35}\sqrt{\log N}}NQ^{-2}+$$
$$+N^{(k-1)/s}N^{-1}+N^{(k-1)/s}e^{-c_{11}\sqrt{\log N}}N^kQ^{-(k+1)}+N\cdot N^{(k-1)/s-2}\log N\}+$$
$$+\sum_{q=1}^{P} q(q^{-1/s+1/2k})^k(qQ)^{k/s-1}\} <$$
$$< c_{36}(s,k)\left\{N^{k/s-1}(\log N)^{-1}+N^{k/s-1}e^{-c_{37}\sqrt{\log N}}+Q^{k/s-1}\sum_{q=1}^{P} q^{1/2}\right\} <$$
$$< c_{38}(s,k)\{N^{k/s-1}(\log N)^{-1}+Q^{k/s-1}P^{3/2}\} =$$
$$= c_{38}(s,k)\{N^{k/s-1}(\log N)^{-1}+(PQ)^{k/s-1}P^{(5s-2k)/2s}\} \le$$
$$\le c_{38}(s,k)\{N^{k/s-1}(\log N)^{-1}+N^{k/s-1}P^{(5s-6s)/2s}\} \le$$
$$\le c_{39}(s,k)N^{k/s-1}\{(\log N)^{-1}+(\log N)^{-\sigma/2}\} < c_{40}(s,k)N^{k/s-1}(\log N)^{-1}$$

(since we have $\sigma=\sigma_1+s(\sigma_2-1)\ge\sigma_1=2^{6s}(\sigma_2+2)\ge 2^6\cdot 2>2$). For sufficiently large N, (65) and (68) yield (16), and this completes the proof of Corollary 1.

PROOF of Corollary 2.

Let p denote the least prime number satisfying $(p, V)=1$. Then for all N, there exists a non-negative integer t such that

$$N-tp^s \equiv k-V+1 \pmod V$$

and

$$0 \leqq t \leqq V-1. \tag{69}$$

If $N > N_1(k, s) = N_0(k-V+1, s)+(V-1)p^s$ then by Corollary 1 (with $k-V+1$ and $N-tp^s$ in place of k and N, respectively), $N-tp^s$ can be represented in the form

$$q_1^s+q_2^s+\ldots+q_{k-V+1}^s = N-tp^s$$

hence

$$N = q_1^s+q_2^s+\ldots+q_{k-V+1}^s+p^s+\ldots+p^s(= q_1^s+q_2^s+\ldots+q_{k-V+1}^s+tp^s).$$

By (9), the number of the terms on the right-hand side is

$$k-V+1+t \leqq k$$

which completes the proof of Corollary 2.

REFERENCES

[1] HUA, L. K.: *Additive Primzahltheorie,* Teubner, Leipzig, 1959.
[2] OSTMANN, H. H.: *Additive Zahlentheorie,* I, Springer-Verlag, Berlin—Göttingen—Heidelberg, 1956.
[3] PRACHAR, K.: *Primzahlverteilung,* Springer-Verlag, Berlin, 1957.
[4] VINOGRADOV, I. M.: *The Method of Trigonometrical Sums in the Theory of Numbers,* Interscience, New York, 1954.

(Received February 4, 1980)

MTA MATEMATIKAI KUTATÓ INTÉZETE
REÁLTANODA U. 13—15
H—1053 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 169—181

ISOMETRIES OF THE SPACE OF COMPACT SUBSETS OF E^d

by

PETER M. GRUBER and GÜNTER LETTL

Summary

The isometries of the space of compact subsets of E^d with respect to the Hausdorff-metric are the mappings generated by isometries of E^d.

1. Introduction

Let $\mathcal{K}$ denote the class of all (non-empty) compact subsets of d-dimensional Euclidean space E^d. For $A, B \subset E^d$ and $\lambda \in \mathbf{R}$ define $A+B := \{x+y | x \in A, y \in B\}$ and $\lambda A := \{\lambda x | x \in A\}$. Let U denote the unit ball and $\|\cdot\|$ the Euclidean norm of E^d. Then the *Hausdorff-metric* δ on $\mathcal{K}$ may be defined by

$$\delta(C, D) := \min\{\lambda \in \mathbf{R}^+ | C \subset D+\lambda U,\ D \subset C+\lambda U\}$$

for $C, D \in \mathcal{K}$ or, equivalently,

$$\delta(C, D) := \max\{\max_{x\in C}\min_{y\in D}\|x-y\|,\ \max_{y\in D}\min_{x\in C}\|x-y\|\}$$

for $C, D \in \mathcal{K}$ (see, e.g., HAUSDORFF [9], p. 145 or BLASCHKE [2], p. 60).

Let $\mathcal{K}$ be endowed with the topology induced by δ. Properties of $\mathcal{K}$ or, more generally, of spaces of closed or compact subsets of general metric and topological spaces have been investigated intensively during the last 50 years (see, e.g., HAUSDORFF [9], KURATOWSKI [11] and POPOV [12]). δ also plays an important role in convexity and approximation theory (see, e.g., BLASCHKE [2], HADWIGER [8], KELLY and WEISS [10] and SENDOV [15]). These investigations indicate that the particular metric space $\langle \mathcal{K}, \delta \rangle$ is of independent interest.

R. SCHNEIDER [14] conjectured that the isometries of $\langle \mathcal{K}, \delta \rangle$ into itself are precisely the mappings of the form $C \to i(C) (:= \{i(x) | x \in C\})$ for $C \in \mathcal{K}$ where i is a fixed isometry of E^d into itself, i.e., the isometries of $\langle \mathcal{K}, \delta \rangle$ are *generated* by isometries of the underlying space E^d. For the case of surjective isometries the conjecture has been confirmed by GRUBER [5] who has also proved some related results in [4]. In the present note we give a proof of the general case. This proof partly follows the proof of the corresponding result for the class $\mathcal{C}$ of compact convex subsets of E^d (see [7]).

AMS (MOS) subject classification scheme 1979/80. Primary 52A99, 54E40. Secondary 46C99, 54E35, 54E45.

Key words and phrases. Metric space, isometries, space of compact subsets, compact sets, Hausdorff metric, Euclidean space.

THEOREM. *A mapping I: $\mathcal{K}\to\mathcal{K}$ is an isometry of $\langle\mathcal{K},\delta\rangle$ into itself precisely when there is an isometry i of E^d into itself such that $I(C)=i(C)$ for all $C\in\mathcal{K}$.*

We conjecture that the Theorem holds also when the Euclidean norm is replaced by an arbitrary norm.

The above Theorem as well as the results of [6] and [7] belong to the following general problem: Let a class of functions $\mathcal{F}$ on, or a class of subsets $\mathcal{S}$ of some space X be given, such that one can metrisize $\mathcal{F}$ or $\mathcal{S}$ in a simple manner using a metric or a measure or some other property of X. Then one may expect that the isometries of $\mathcal{F}$ or $\mathcal{S}$ into itself are closely related to special mappings of X into itself. Well-known examples for this are the classical theorems of BANACH [1], p. 173, and STONE [17] (see [3], p. 115) and Banach [1], p. 174, and their modern descendants. (The paper of SOURMOUR [16] contains a review of some of them.)

Let $\langle\,,\rangle$ denote the inner product in E^d. There is no danger of confusing $\langle x,y\rangle$ and $\langle\mathcal{K},\delta\rangle$. For $x,y\in E^d$ let $[x,y]$ be the line segment with endpoints x, y. o denotes the origin, S the $(d-1)$-dimensional unit sphere of E^d. We will not distinguish between $x\in E^d$ and $\{x\}\in\mathcal{K}$. bd, diam, conv, lin and pos stand, respectively, for boundary, diameter and convex, linear and positive hull. $o(\cdot)$ is the Bachmann—Landau symbol.

2. Preliminaries

This section contains a collection of simple results which will be needed in the proof of the theorem in Section 3.

(1) *Let $I_1, I_2, \dots$ be a sequence of isometries of $\langle\mathcal{K},\delta\rangle$ (into itself) such that $I_1(o), I_2(o), \dots$ are contained in a bounded subset of E^d. Then there exist a subsequence $I_{k_1}, I_{k_2}, \dots$ and an isometry I_0 such that $I_0(C)=\lim\limits_{l\to+\infty} I_{k_l}(C)$ for all $C\in\mathcal{K}$.*

This has been proved for $\langle\mathcal{C},\delta\rangle$ in [7]. The same proof is valid in the present case if the more general version of the Blaschke selection theorem (as stated e.g. in [8], p. 154, or [13], p. 91) is used. The next proposition is due to GRUBER [4]:

(2) *Any isometry of $\langle\mathcal{K},\delta\rangle$ which maps some point onto a point is generated by an isometry of E^d.*

For $C, D\in\mathcal{K}$ the definition of $\delta(C,D)$ shows that there is a point $c\in C$ (or D) such that for each point $d\in D$ (or C, respectively) which is nearest to c we have $\|c-d\|=\delta(C,D)$. In general c and d will not be unique. $c-d$ will be called a *δ-vector* and $(c-d)/\|c-d\|$ a *δ-unit vector* from D to (the point c of) C. We show:

(3) *Let $C, C_0, C_l\in\mathcal{K}$ be given such that $\delta(C,C_0)+\delta(C_0,C_l)=\delta(C,C_l)$ and suppose, that there exists a δ-vector from C_l to $c\in C$. Then there exists a δ-vector from C_l to C_0.*

Choose $c_0\in C_0$ nearest to c, and $c_l\in C_l$ nearest to c_0. Then the definition of δ implies $\|c-c_0\|\leqq\delta(C,C_0)$, $\|c_0-c_l\|\leqq\delta(C_0,C_l)$ and thus $\|c-c_l\|\leqq\|c-c_0\|+\|c_0-c_l\|\leqq\delta(C,C_0)+\delta(C_0,C_l)=\delta(C,C_l)$. Our assumptions show that c has distance

$\geqq\delta(C, C_l)$ from each point of C_l. Therefore equality holds in each of the above inequalities. In particular $\|c_0-c_l\|=\delta(C_0, C_l)$. Since $c_l\in C_l$ is nearest to $c_0\in C_0$ we infer that c_0-c_l is a δ-vector from C_l to C_0. This proves (3). Essentially the same proof yields the following result:

(4) *Let $C_0, C_1, C_2, \ldots, C_l\in\mathcal{K}$ be such that $\delta(C_0, C_1)+\delta(C_1, C_2)+\ldots+\delta(C_{l-1}, C_l)=$ $=\delta(C_0, C_l)$ and suppose that there is a δ-vector from C_l to $c_0\in C_0$. Then there exists a vector u which is a δ-unit vector from each of $C_1, \ldots, C_l$ to $c_0\in C_0$.*

Choose $c_1\in C_1$ nearest to c_0, $c_2\in C_2$ nearest to $c_1, \ldots, c_l\in C_l$ nearest to c_{l-1}. As before

$$\|c_0-c_1\| \leqq \delta(C_0, C_1), \ \ldots, \|c_{l-1}-c_l\| \leqq \delta(C_{l-1}, C_l).$$

Hence

$$\|c_0-c_l\| \leqq \|c_0-c_1\|+\ldots+\|c_{l-1}-c_l\| \leqq \delta(C_0, C_1)+\ldots+\delta(C_{l-1}, C_l) = \delta(C_0, C_l).$$

Since the assumptions in (4) imply that c_0 has distance $\geqq\delta(C_0, C_l)$ from each point of C_l, equality holds throughout. This implies that $\|c_0-c_i\|=\delta(C_0, C_i)$ for $i\in\{1, \ldots, l\}$ and that $c_0-c_1, c_0-c_2, \ldots, c_0-c_l$ are positive multiples of each other, thus confirming (4).

The next two propositions are obvious:

(5) *Let $C, D\in\mathcal{K}$, $c\in C$, $d\in D$ be given. Then* $\big|\|c-d\|-\delta(C, D)\big|\leqq\operatorname{diam} C+\operatorname{diam} D$.

(6) *Let $C, D\in\mathcal{K}$. Then* $\delta(C, D)\geqq\delta(\operatorname{conv} C, \operatorname{conv} D)$.

The smoothness of $\|\cdot\|$ yields the following results:

(7) *Let R be a ray in E^d and $\alpha\in\mathbf{R}^+$. Denote by $'$ the orthogonal projection onto R. Suppose that for $\lambda\in\mathbf{R}^+$ compact sets $C, C(\lambda)\in\mathcal{K}$ are given such that $C', C(\lambda)'\neq\emptyset$, $C, C(\lambda)\subset R+\alpha U$ and* $\min\{\|y\| \mid y\in C(\lambda)\}\to+\infty$ *as $\lambda\to+\infty$. Then, as $\lambda\to+\infty$,*

$$\begin{aligned}\delta(C, C(\lambda)) &= \delta(\operatorname{conv} C', C(\lambda))+o(1)\\ &= \delta(\operatorname{conv} C', \operatorname{conv} C(\lambda)')+o(1).\end{aligned}$$

(8) *Let $C, D\in\mathcal{C}$. Then* $\bigcap_{x\in E^d}\{x+C+\delta(x, D)\,U\}=C+D$.

We will also need the following properties of convergent sequences in $\mathcal{K}$:

(9) *If $C_0, C_1, \ldots\in\mathcal{K}$ with $C_1, C_2, \ldots\to C_0$ then*

$$\operatorname{conv} C_1, \operatorname{conv} C_2, \ldots\to\operatorname{conv} C_0.$$

(10) *If $C_0, C_1, \ldots, D_0, D_1, \ldots\in\mathcal{C}$ with $C_1, C_2, \ldots\to C_0$, $D_1, D_2, \ldots\to D_0$ then*

$$C_1+D_1, C_2+D_2, \ldots\to C_0+D_0.$$

Let $i_1, \ldots, i_n$ be orthogonal transformations of E^d and $\lambda_1, \ldots, \lambda_n\in[0, 1]$ with $\lambda_1+\ldots+\lambda_n=1$. The mapping $m\colon \mathcal{C}\to\mathcal{C}$ defined by

$$m(C):=\{\lambda_1 i_1(x_1)+\ldots+\lambda_n i_n(x_n) \mid x_1, \ldots, x_n\in C\} \quad \text{for} \quad C\in\mathcal{C}$$

is called a *rotational mean* (see [8], p. 168). Then the sphericity theorem of Hadwiger may be formulated in the following way:

(11) *Let $D\in\mathcal{C}$ consist of more than one point. Then there exist a sequence $m_1, m_2, \ldots$ of rotational means and a number $\varrho\in\mathbf{R}^+$ such that $m_1(D), m_2(D), \ldots\to\varrho U$.*

If D has non-empty interior this is the theorem as stated in [8], p. 170. If D has empty interior choose some rotational mean l such that $l(D)$ has non-empty interior. Then there are rotational means $l_1, l_2, \dots$ such that $l_1(lD), l_2(lD), \dots$ converges to a ball with center o. Now let $m_1 := l_1 \circ l$, $m_2 := l_2 \circ l, \dots$. Taking into account that a composition of rotational means is again a rotational mean, we see that (11) holds also in the case when D has empty interior.

3. Proof of the Theorem

If i is an isometry of E^d, then the mapping $I\colon \mathcal{K} \to \mathcal{K}$ generated by i is obviously an isometry of $\langle \mathcal{K}, \delta \rangle$.

Conversely assume that $I\colon \mathcal{K} \to \mathcal{K}$ is an isometry of $\langle \mathcal{K}, \delta \rangle$ (into itself).

In the first part of the proof we construct an isometry J which is closely related to I and will permit us to construct an isometry i of E^d.

For each $k \in \mathbf{N}$ define an isometry I_k of $\langle \mathcal{K}, \delta \rangle$ by $I_k(C) := (1/k) I(kC)$ for $C \in \mathcal{K}$. Since $I_1(o), I_2(o), \dots \to \{o\}$, (1) yields the existence of a subsequence $I_{k_1}, I_{k_2}, \dots$ and of an isometry I_0 of $\langle \mathcal{K}, \delta \rangle$ such that

$$I_0(C) = \lim_{l \to \infty} I_{k_l}(C) \quad \textit{for all} \quad C \in \mathcal{K}. \tag{12}$$

In particular $I_0(o) = \{o\}$. Hence by (2) there exists an isometry i_0 of E^d such that

$$I_0(C) = i_0(C) \quad \textit{for all} \quad C \in \mathcal{K}. \tag{13}$$

Since $I_0(o) = \{o\}$, we have $i_0(o) = o$ and thus $i_0(U) = U$. Therefore $\lim\limits_{l \to \infty} I_{k_l}(o) = I_0(o) = \{o\}$ and $\lim\limits_{l \to \infty} I_{k_l}(U) = I_0(U) = i_0(U) = U$. Choose $\varepsilon \in]0, 1/2[$ and $m \in \mathbf{N}$ such that the mapping

$$J(:= I_{k_m}) \quad \textit{defined by} \quad J(C) := \frac{1}{k_m} I(k_m C) \quad \textit{for each} \quad C \in \mathcal{K} \tag{14}$$

is an isometry of $\langle \mathcal{K}, \delta \rangle$ satisfying the inequalities

$$\delta(o, J(o)) \left[= \delta(I_0(o), I_{k_m}(o)) \right] < \varepsilon, \tag{15}$$

$$\delta(U, J(U)) \left[= \delta(I_0(U), I_{k_m}(U)) \right] < \varepsilon. \tag{16}$$

We shall establish the inclusion

$$J(x) \subset \{y \mid \|x\| - \varepsilon \leqq \|y\| \leqq \|x\| + \varepsilon\} \quad \text{for each} \quad x \in E^d \quad \text{with} \quad \|x\| \geqq 1. \tag{17}$$

Choose x. Then $\delta(J(o), J(x)) = \delta(o, x) = \|x\|$ yields $J(x) \subset J(o) + \|x\| U$ and (15) implies $J(o) \subset \varepsilon U$. Thus

$$J(x) \subset [J(o) + \|x\| U \subset \varepsilon U + \|x\| U =] \{y \mid \|y\| \leqq \|x\| + \varepsilon\}. \tag{18}$$

This proves half of (17). Since by (16) $U \subset J(U) + \varepsilon U$, we deduce from (18), $\|x\| \geqq 1$ and $\varepsilon \in]0, 1/2[$ that

$$\begin{aligned} J(x) &\subset [(\varepsilon + \|x\|) U = U + (\varepsilon + \|x\| - 1) U \subset \\ &\subset J(U) + (\|x\| + 2\varepsilon - 1) U \subset] \, J(U) + \|x\| U. \end{aligned} \tag{19}$$

Because of (16) $J(U)\subset(1+\varepsilon)U$. Hence for each $y\in J(x)$,

$$(20)\qquad J(U)\subset[(1+\varepsilon)U\subset y+(\|y\|+1+\varepsilon)U\subset]\,J(x)+(\|y\|+1+\varepsilon)U.$$

The definition of δ together with (19) and (20) shows that

$$\max\{\|x\|,\|y\|+1+\varepsilon\}\geqq\delta\big(J(x),J(U)\big)=\delta(x,U)=\|x\|+1.$$

Thus $\|y\|+1+\varepsilon\geqq\|x\|+1$ for each $y\in J(x)$, i.e., $J(x)\subset\big\{y\big|\|x\|-\varepsilon\leqq\|y\|\big\}$. This proves the second half of (17).

Now we take the first step towards a description of the images of points and prove

$$(21)\qquad \operatorname{diam}J(k_l x)=o(k_l)\quad as\quad l\to\infty\quad for\ each\quad x\in E^d\setminus\{o\}.$$

If x is given, then (14), (12) and (11) imply that

$$\frac{1}{k_l}J(k_l x)=\frac{1}{k_m}\cdot\frac{1}{k_l}I\big(k_l(k_m x)\big)\to\frac{1}{k_m}I_0(k_m x)=\frac{1}{k_m}i_0(k_m x).$$

Consequently, diam $(1/k_l)J(k_l x)\to 0$ as $l\to\infty$. This proves (21).

We show:

(22) *For each $x\in S$ and $l\in\mathbf{N}$ the set T_l of pairs $(u,c)\in S\times J(o)$ where u is a δ-unit vector from $J(k_l x)$ to $c\in J(o)$ is non-empty and compact.*

Choose x, l. $J(k_l x)\subset(k_l+\varepsilon)U$ by (17). Since (16) implies $U\subset J(U)+\varepsilon U$ we conclude that $J(k_l x)\subset J(U)+(k_l+2\varepsilon-1)U\subset J(U)+k_l U$. Therefore each point of $J(k_l x)$ has distance $\leqq k_l$ from some point of $J(U)$. Taking into account that $\delta\big(J(k_l x),J(U)\big)=$ $=\delta(k_l x,U)=k_l+1>k_l$ the definition of δ-vectors implies that there is a δ-vector from $J(k_l x)$ to $J(U)$ and

$$\delta\big(J(U),J(o)\big)+\delta\big(J(o),J(k_l x)\big)=[\delta(U,o)+\delta(o,k_l x)=1+k_l]=\delta\big(J(U),J(k_l x)\big).$$

Thus by (3) there is a δ-vector from $J(k_l x)$ to some $c\in J(o)$. Hence $T_l\neq\emptyset$. The compactness of T_l is obvious. This concludes the proof of (22). The following is a refinement of (22):

(23) *For each $x\in S$ there is a unique vector $u\in S$ and some point $c\in J(o)$ such that u is a δ-unit vector from $J(k_l x)$ to $c\in J(o)$ for each $l\in\mathbf{N}$.*

Let $x\in S$ be given. For each $l\in\mathbf{N}$

$$\delta\big(J(o),J(k_1x)\big)+\delta\big(J(k_1x),J(k_2x)\big)+\ldots+\delta\big(J(k_{l-1}x),J(k_l x)\big)=$$
$$=k_1+(k_2-k_1)+\ldots+(k_l-k_{l-1})=\delta\big(J(o),J(k_l x)\big).$$

By (22) there exists a δ-unit vector from $J(k_l x)$ to some $c\in J(o)$. Now (4) implies existence of a vector $u\in S$ such that u is a δ-unit vector from each of $J(k_1x),\ldots,J(k_l x)$ to $c\in J(o)$. Using the notation of (22) we have $T_1\cap\ldots\cap T_l\neq\emptyset$ for each $l\in\mathbf{N}$. Since $T_1,T_2,\ldots$ are compact $T:=\bigcap\{T_l\,|\,l\in\mathbf{N}\}\neq\emptyset$. If $(u,c)\in T$ and $l\in\mathbf{N}$, the vector $k_l u$ is a δ-vector from $J(k_l x)$ to $c\in J(o)$. To prove uniqueness of u choose

$(u, c), (v, d)\in T$. Then $c-k_l u,\ d-k_l v\in J(k_l x)$ for all $l\in\mathbf{N}$ by the definition of δ-vectors, and thus

$$\big|k_l\|u-v\|-\|c-d\|\big|[\leqq \|(c-k_l u)-(d-k_l v)\| \leqq \operatorname{diam} J(k_l x)] = o(k_l) \quad as \quad l\to\infty$$

by (21). Hence $u=v$, concluding the proof of (23).

Let $-i$ denote the map $x\to u$ of S into itself appearing in (23). Then

(24) *i: $S\to S$ is an isometry.*

Let $x, y\in S$ and let $u:=-i(x)$, $v:=-i(y)$. Then for each $l\in\mathbf{N}$ the vectors $k_l u$, $k_l v$ are δ-vectors from $J(k_l x)$ and $J(k_l y)$ to points $c, d\in J(o)$. Hence $c-k_l u\in J(k_l x)$, $d-k_l v\in J(k_l y)$. From this, together with (5) and (21) we infer

$$\begin{aligned}&\big|\|(c-d)-k_l(u-v)\|-k_l\|x-y\|\big|\\ &\big[=\big|\|(c-k_l u)-(d-k_l v)\|-\delta\big(J(k_l x), J(k_l y)\big)\big|\\ &\leqq \operatorname{diam} J(k_l x)+\operatorname{diam} J(k_l y)\big]\\ &= o(k_l) \quad \text{as} \quad l\to\infty,\end{aligned}$$

thus proving (24).

i can be extended uniquely to an isometry of E^d which will also be denoted by i. Let

(25) $\qquad K\colon \mathcal{K}\to\mathcal{K}$ *be defined by* $K(C):=i^{-1}\big(J(C)\big)$ *for* $C\in\mathcal{K}$.

Obviously, K is an isometry of $\langle\mathcal{K}, \delta\rangle$. The propositions (17), (21) and (23) now take the following form:

(26) $\qquad K(x)\subset\{y\,\big|\,\|x\|-\varepsilon\leqq\|y\|\leqq\|x\|+\varepsilon\}$ *for* $x\in E^d$ *with* $\|x\|\geqq 1$,

(27) $\qquad \operatorname{diam} K(k_l x)=o(k_l)$ *as* $l\to\infty$ *for* $x\in E^d\setminus\{o\}$,

(28) *for all $x\in S$ and $l\in\mathbf{N}$ the vector $-k_l x$ is a δ-vector from $K(k_l x)$ to a $c\in K(o)$.*

(26)—(28) will be used to prove that

(29) $\qquad K(k_l x)\subset \operatorname{lin}\{x\}+2\varepsilon U$ *for* $x\in S$ *and* $l\in\mathbf{N}$.

Let x and l be given. Suppose there exists a $w\in K(k_l x)$ which is not contained in the cylinder on the right-hand side of the inclusion (29). Then one can choose a point

$$z\in S \quad with \quad \langle x, z\rangle=0, \quad w=\xi x+\zeta z, \quad \zeta<-2\varepsilon.$$

Let $\varphi\in]0, \pi/2[$ be so small that

(30) $$|\xi|\sin\varphi+\zeta\cos\varphi<-2\varepsilon.$$

Since z is orthogonal to x we have

(31) $\qquad \|k_l x-k_m z\|<k_m+\varepsilon$ *for all sufficiently large* $m\in\mathbf{N}$.

Since $\|z\|=1$, (26) yields

(32) $$K(k_m z) \subset \{y | k_m - \varepsilon \leqq \|y\| \leqq k_m + \varepsilon\} \quad \textit{for each} \quad m \in \mathbf{N}.$$

Applying (28) to z we see that there exists a point $d \in K(o)$ such that for all m the vector $-k_m z$ is a δ-vector from $K(k_m z)$ to $d \in K(o)$. (27) implies diam $K(k_m z) = o(k_m)$ as $m \to \infty$. Hence $K(k_m z) \subset \{y | \|y - k_m z\| \leqq o(k_m)\}$ as $m \to \infty$ and thus

(33) $$K(k_m z) \subset \left\{y \left| \left\langle \frac{y}{\|y\|}, z \right\rangle \geqq \cos \varphi \right.\right\} \quad \textit{for all sufficiently large} \quad m \in \mathbf{N}.$$

Fix an $m \in \mathbf{N}$ for which (31) and (33) hold. Then the definition of δ, (32), (33) and (31) imply

(34) $$\begin{aligned} w \in K(k_l x) &\subset K(k_m z) + \delta(K(k_l x), K(k_m z))U \\ &= K(k_m z) + \|k_l x - k_m z\| U \\ &\subset \left\{ y | k_m - \varepsilon \leqq \|y\| \leqq k_m + \varepsilon, \left\langle \frac{y}{\|y\|}, z \right\rangle \geqq \cos \varphi \right\} + (k_m + \varepsilon) U \\ &=: A + (k_m + \varepsilon) U. \end{aligned}$$

lin $\{z\}$ is the axis of rotation of the compact set A. Thus there exists a point $y \in A \cap \cap \operatorname{lin}\{x, z\}$ which is nearest to $w = \xi x + \zeta z \in \operatorname{lin}\{x, z\}$. An elementary argument shows that $y = (k_m - \varepsilon)(\sin \varphi \cdot \operatorname{sign} \xi \cdot x + \cos \varphi \cdot z)$. (Here sign $0 = 0$.) From this together with (34) and (30) we deduce that

$$\begin{aligned} (k_m + \varepsilon)^2 \geqq \|w - y\|^2 &= (\xi - (k_m - \varepsilon) \sin \varphi \operatorname{sign} \xi)^2 + (\zeta - (k_m - \varepsilon) \cos \varphi)^2 \\ &= (k_m - \varepsilon)^2 - 2|\xi|(k_m - \varepsilon) \sin \varphi - 2|\zeta|(k_m - \varepsilon) \cos \varphi + \xi^2 + \zeta^2 \\ &\geqq k_m^2 - 2k_m \varepsilon + \varepsilon^2 + 2(k_m - \varepsilon)(-|\xi| \sin \varphi - \zeta \cos \varphi) + \zeta^2 \\ &> k_m^2 - 2k_m \varepsilon + \varepsilon^2 + 2(k_m - \varepsilon) 2\varepsilon + 4\varepsilon^2 = (k_m + \varepsilon)^2. \end{aligned}$$

This contradiction shows that our assumption was false, thus confirming (29).

The next part of our proof contains a rough description of the images of points. This will be used to obtain a rough description of the images of arbitrary compact sets. It will turn out that the convex hull of the image (under K) of a compact set can be obtained in a simple way from the convex hull of the set itself.

At first we show the following proposition:

(35) *Let $x \in S$ and denote by $H(x)$ the supporting half space of $K(o)$ with exterior normal vector $-x$. Then*

$$K(k_l x) \subset H(x) + (k_l + o(1))x \quad \textit{as} \quad l \to \infty.$$

Because of (28) there is a point $c \in K(o)$ such that for each l the vector $-k_l x$ is a δ-vector from $K(k_l x)$ to $c \in K(o)$. Therefore the definition of δ-vectors shows that $K(k_l x) \subset \{y | \|y - c\| \geqq k_l (= \|-k_l x\|)\}$. Furthermore $c + k_l x \in K(k_l x)$, together with (27) and (29) shows that $K(k_l x) \subset \operatorname{pos}\{x\} + 2\varepsilon U$. It is obvious that $c \in \operatorname{bd} H(x)$. From these statements we conclude that $K(k_l x)\subset H(x)+(k_l+o(1))x$ as $l\to\infty$. This proves (35). Next we prove the following:

(36) *Let $z\in E^d$, $x\in S$ and let $H(x)$ denote the supporting half space of $K(o)$ with exterior normal vector $-x$. Then $K(z)\subset z+H(x)$.*

Let μx be the orthogonal projection of z onto lin $\{x\}$. Clearly,

$$\|z-k_l x\| = k_l-\mu+o(1) \quad \text{as} \quad l\to\infty. \tag{37}$$

Since $\delta(\operatorname{conv} K(z), \operatorname{conv} K(k_l x))\leqq\delta(K(z), K(k_l x))=\|z-k_l x\|$ by (6) it follows from (35) and (37) that

$$\begin{aligned} K(z) &\subset \operatorname{conv} K(z) \subset \operatorname{conv} K(k_l x)+\|z-k_l x\| U \\ &\subset H(x)+(k_l+o(1))x+(k_l-\mu+o(1))U \\ &= H(x)+\mu x+o(1)x = H(x)+z+o(1)x \quad \text{as} \quad l\to\infty \end{aligned}$$

and thus $K(z)\subset H(x)+z$ concluding the proof of (36). An immediate consequence of (36) is that

$$\operatorname{conv} K(z) \subset z+\operatorname{conv} K(o) \quad \text{for each} \quad z\in E^d. \tag{38}$$

The following refinement of (38) will be required later on

$$\operatorname{conv} K(z) = z+\operatorname{conv} K(o) \quad \text{for each} \quad z\in E^d. \tag{39}$$

Considering (38), (25), (14) and the definition of I_k we see that so far we have proved: For each isometry of $\langle\mathcal{K},\delta\rangle$ the convex hull of the image of an arbitrary point z is contained in a translate of the convex hull of the image of o. Applying this to the isometry K_z of $\langle\mathcal{K},\delta\rangle$ defined by $K_z(C):=K(C+z)$ for $C\in\mathcal{K}$ one sees that for each $z\in E^d$ the set conv $K(o)=\operatorname{conv} K_z(-z)$ is contained in a translate of conv $K(z)=$ $=\operatorname{conv} K_z(o)$. Together with (38) this proves (39).

Given $x\in S$ let $'$ denote the orthogonal projection onto lin $\{x\}$. Line segments of the form $[\alpha x, \beta x]$ will be written simply as $[\alpha, \beta]$. The following propositions (40) and (41) will be used in extending (39) from points to arbitrary compact sets C. (39) together with (7) implies:

(40) *Let $x\in S$ be given. Then*

$$\delta(C, K(\lambda x)) = \delta(\operatorname{conv} C', \operatorname{conv} K(\lambda x)')+o(1)$$

as $\lambda\to\pm\infty$ for each $C\in\mathcal{K}$.

We show:

(41) *Let $x\in S$ and* conv $K(o)'=[\alpha,\beta]$ $(\alpha\leqq\beta)$. *Then* conv $K([\mu,\nu])'=[\alpha+\mu,\beta+\nu]$ *for all $\mu,\nu\in\mathbf{R}$ with $\nu-\mu>\beta-\alpha$ $(\geqq 0)$.*

Choose μ, ν and let conv $K([\mu,\nu])'=:[\sigma,\tau]\,(\sigma\leqq\tau)$. (39) yields

$$\begin{aligned} \operatorname{conv} K(\lambda x)' &\left[= (\operatorname{conv} K(o)+\lambda x)' = \operatorname{conv} K(o)'+\lambda x\right] \\ &= [\alpha+\lambda, \beta+\lambda] \quad \text{for each} \quad \lambda\in\mathbf{R}. \end{aligned}$$

From this together with (40) we conclude that

$$\begin{aligned}
(42)\qquad \max\{\alpha+\lambda-\sigma, \beta+\lambda-\tau\} &= \delta([\sigma, \tau], [\alpha+\lambda, \beta+\lambda])\\
&= \delta(\operatorname{conv} K([\mu, \nu])', \operatorname{conv} K(\lambda x)')\\
&= \delta(K([\mu, \nu]), K(\lambda x))+o(1) = \delta([\mu, \nu], \lambda x)+o(1)\\
&= \lambda-\mu+o(1) \quad \textit{as} \quad \lambda \to +\infty,
\end{aligned}$$

$$\begin{aligned}
(43)\qquad \max\{\sigma-\alpha+\lambda, \tau-\beta+\lambda\} &= \delta([\sigma, \tau], [\alpha-\lambda, \beta-\lambda])\\
&= \delta(\operatorname{conv} K([\mu, \nu])', \operatorname{conv} K(-\lambda x)')\\
&= \delta(K([\mu, \nu]), K(-\lambda x))+o(1) = \delta([\mu, \nu], -\lambda x)+o(1)\\
&= \nu+\lambda+o(1) \quad \textit{as} \quad \lambda \to +\infty.
\end{aligned}$$

In particular, $\alpha+\lambda-\sigma \leqq \lambda-\mu+o(1)$, $\tau-\beta+\lambda \leqq \nu+\lambda+o(1)$ (as $\lambda\to+\infty$) and therefore $\alpha+\mu\leqq\sigma$, $\tau\leqq\beta+\nu$. Suppose $\alpha+\mu<\sigma$. Then $\alpha+\lambda-\sigma<\lambda-\mu+o(1)$ and thus $\beta+\lambda-\tau=\lambda-\mu+o(1)$ by (42). Hence $\beta=\tau-\mu$. Together with $\mu<\nu$ this implies $\tau-\beta+\lambda=\mu+\lambda<\nu+\lambda+o(1)$. Now (43) shows that $\sigma-\alpha+\lambda=\nu+\lambda+o(1)$, i.e., $\alpha=\sigma-\nu$. It follows from $\beta=\tau-\mu$ and $\alpha=\sigma-\nu$ that $\beta-\alpha=(\nu-\mu)+(\tau-\sigma)\geqq\nu-\mu$, a contradiction. Therefore $\alpha+\mu=\sigma$. Similarly, one can show that $\tau=\beta+\nu$. Thus $[\sigma, \tau]=[\alpha+\mu, \beta+\nu]$, confirming (41).

(44) *For all $x\in S$ and $\mu, \nu\in\mathbf{R}^+$ $(\mu<\nu)$ the inclusion* $K([\mu, \nu])\subset[\mu, \nu]+\operatorname{conv} K(o)$ *holds.*

Choose x, μ, ν. Then (44) is a consequence of the definition of δ, (6), (39) and (8), namely

$$\begin{aligned}
K([\mu, \nu]) &\subset \operatorname{conv} K([\mu, \nu])\\
&\subset \bigcap_{z\in E^d} \{\operatorname{conv} K(z)+\delta(\operatorname{conv} K(z), \operatorname{conv} K([\mu, \nu]))U\}\\
&\subset \bigcap_{z\in E^d} \{z+\operatorname{conv} K(o)+\delta(z, [\mu, \nu])U\}\\
&= [\mu, \nu]+\operatorname{conv} K(o).
\end{aligned}$$

It follows from (44) and (7) that

(45) *for all $x\in S$ and $C\in\mathcal{K}$ we have*

$$\begin{aligned}
&\delta(\operatorname{conv} C', \pm[\lambda, 2\lambda]) = \delta(C, \pm[\lambda, 2\lambda])+o(1)\\
&= \delta(K(C), K(\pm[\lambda, 2\lambda]))+o(1)\\
&\quad = \delta(\operatorname{conv} K(C)', \operatorname{conv} K(\pm[\lambda, 2\lambda])')+o(1) \quad \textit{as} \quad \lambda\to+\infty.
\end{aligned}$$

We are now in a position to prove that

(46) *for all $C\in\mathcal{K}$ and $x\in S$ the equality* $\operatorname{conv} K(C)'=\operatorname{conv} C'+\operatorname{conv} K(o)'$ *holds.*

For given C, x let $\operatorname{conv} K(o)'=[\alpha, \beta]$ $(\alpha \leqq \beta)$ be as before and suppose $\operatorname{conv} C'=$ $=[\xi, \eta]$ $(\xi \leqq \eta)$, $\operatorname{conv} K(C)'=[\varphi, \psi]$ $(\varphi \leqq \psi)$. Then (45) and (41) yield

$$\begin{aligned}
2\lambda-\eta &= \delta([\xi,\eta],\ [\lambda, 2\lambda]) = \delta(\operatorname{conv} C',\ [\lambda, 2\lambda]) \\
&= \delta\big(\operatorname{conv} K(C)',\ \operatorname{conv} K([\lambda, 2\lambda])'\big)+o(1) \\
&= \delta([\varphi,\psi],\ [\alpha+\lambda, \beta+2\lambda])+o(1) = 2\lambda+\beta-\psi+o(1) \quad \text{as} \quad \lambda \to +\infty, \\
2\lambda+\xi &= \delta([\xi,\eta],\ -[\lambda, 2\lambda]) = \delta(\operatorname{conv} C',\ -[\lambda, 2\lambda]) \\
&= \delta\big(\operatorname{conv} K(C)',\ \operatorname{conv} K(-[\lambda, 2\lambda])'\big)+o(1) \\
&= \delta([\varphi,\psi],\ [\alpha-2\lambda, \beta-\lambda])+o(1) = \varphi+2\lambda-\alpha+o(1) \quad \text{as} \quad \lambda \to +\infty.
\end{aligned}$$

Consequently, $2\lambda-\eta=2\lambda+\beta-\psi+o(1)$, $2\lambda+\xi=\varphi+2\lambda-\alpha+o(1)$ as $\lambda\to+\infty$ and thus $\varphi=\alpha+\xi$, $\psi=\beta+\eta$, i.e., $[\varphi, \psi]=[\xi, \eta]+[\alpha, \beta]$. This proves (46). An immediate consequence of (46) is:

(47) *For each* $C\in\mathscr{K}$ *the equality* $\operatorname{conv} K(C)=\operatorname{conv} C+\operatorname{conv} K(o)$ *holds.*

We now come to the final part of the proof. First we suppose that $K(o)$ consists of more than one point. Then it is possible to construct an isometry L for which $L(o)=U$. Now an investigation of images of sets consisting of two points only shows that L cannot be an isometry. This shows that we need to consider only the case when $K(o)$ consists of a single point. Then apply (2) to prove that K and thus I is generated by an isometry of E^d.

Suppose $K(o)$ consists of more than one point and let $D:=\operatorname{conv} K(o)$. By (11) there exists a sequence of rotational means $m_1, m_2, \ldots$ such that

$$(48) \qquad m_1(D), m_2(D), \ldots \to \varrho U \quad \textit{for suitable} \quad \varrho\in\mathbf{R}^+.$$

Given $k\in\mathbf{N}$ suppose $m_k(C)=\lambda_1 i_1(C)+\ldots+\lambda_n i_n(C)$ for $C\in\mathscr{C}$ and let $K_k\colon \mathscr{K}\to\mathscr{K}$ be defined for $C\in\mathscr{K}$ by

$$K_k(C):=\varrho^{-1}\lambda_1 i_1\big(K(\lambda_1^{-1} i_1^{-1}(\lambda_2 i_2(K(\lambda_2^{-1} i_2^{-1}\ldots(\lambda_n i_n(K(\lambda_n^{-1} i_n^{-1}(\varrho C))\ldots).$$

It is easy to see that K_k is an isometry of $\langle\mathscr{K}, \delta\rangle$ for each k. Furthermore (47) yields

$$\begin{aligned}
(49) \qquad \operatorname{conv} K_k(C)\big[&= \varrho^{-1}\lambda_1 i_1\big(\operatorname{conv}(K(\lambda_1^{-1} i_1^{-1}(\lambda_2 i_2(K(\lambda_2^{-1} i_2^{-1}\ldots))))))\\
&= \varrho^{-1}\lambda_1 i_1\big(\operatorname{conv}(\lambda_1^{-1} i_1^{-1}(\lambda_2 i_2(K(\lambda_2^{-1} i_2^{-1}\ldots))))+D\big)\\
&= \operatorname{conv}\big(\varrho^{-1}\lambda_2 i_2(K(\lambda_2^{-1} i_2^{-1}\ldots))\big)+\varrho^{-1}\lambda_1 i_1(D)\\
&= \varrho^{-1}\lambda_2 i_2\big(\operatorname{conv}(K(\lambda_2^{-1} i_2^{-1}\ldots))\big)+\varrho^{-1}\lambda_1 i_1(D)\\
&= \varrho^{-1}\lambda_2 i_2\big(\operatorname{conv}(\lambda_2^{-1} i_2^{-1}\ldots)+D\big)+\varrho^{-1}\lambda_1 i_1(D)\\
&= \operatorname{conv}(\ldots)+\varrho^{-1}\big(\lambda_1 i_1(D)+\lambda_2 i_2(D)\big)=\ldots\big]\\
&= \operatorname{conv} C+\varrho^{-1} m_k(D) \quad \textit{for each} \quad C\in\mathscr{K}.
\end{aligned}$$

This together with (48) implies that $\operatorname{conv} K_k(o)=\varrho^{-1}m_k(D)\to U$. Therefore the sequence $K_1(o), K_2(o), \ldots$ is contained in a bounded subset of E^d. Hence (1) shows that there are a subsequence $K_{k_1}, K_{k_2}, \ldots$ and an isometry L of $\langle \mathcal{K}, \delta \rangle$ such that

$$L(C) = \lim_{l\to\infty} K_{k_l}(C) \quad \text{for each} \quad C \in \mathcal{K}.$$

From this together with (9), (49), (48) and (10) we infer that

$$\text{(50)} \qquad L(C) \subset \operatorname{conv} L(C) [= \operatorname{conv} \lim_{l\to\infty} K_{k_l}(C) = \lim_{l\to\infty} \operatorname{conv} K_{k_l}(C)$$

$$= \lim_{l\to\infty} (\operatorname{conv} C + \varrho^{-1} m_k(D))] = \operatorname{conv} C + U.$$

Let $x, y \in S$ be chosen so that $\delta(x, y)(=\|x-y\|)=2\sin(\pi/8)$. Then

$$\text{(51)} \qquad \delta(\{-\lambda x, \lambda x\}, \{-\lambda y, \lambda y\}) = 2\lambda \sin\frac{\pi}{8} \quad \textit{for each} \quad \lambda \in \mathbf{R}^+.$$

Let $'$ and $''$ be the orthogonal projections of E^d onto $\operatorname{lin}\{x\}$ and $\operatorname{lin}\{y\}$, respectively. For line segments $[\alpha x, \beta x] \subset \operatorname{lin}\{x\}$ we simply write $[\alpha, \beta]$. Let $S(x)$ denote the hemisphere $\{z \in S \mid \langle x, z\rangle \geqq 0\}$. Similar notations will be used for y. We have

$$\text{(52)} \quad (-S(x)-\lambda x) \cup (S(x)+\lambda x) \subset L(\{-\lambda x, \lambda x\}) \subset [-\lambda, \lambda] + U \quad \textit{for each} \quad \lambda \in \mathbf{R}^+,$$

$$L(\{-\lambda x, \lambda x\})' \subset [-\lambda-1, -\lambda+o(1)] \cup [\lambda - o(1), \lambda+1] \quad \textit{as} \quad \lambda \to +\infty.$$

To prove this choose $\lambda \in \mathbf{R}^+$. By (50) $\operatorname{conv}\big(L(\{-\lambda x, \lambda x\})\big) = [-\lambda, \lambda] + U$. Since each extreme point of the convex hull of a compact set belongs to the set, $\big(-S(x)-\lambda x\big) \cup \big(S(x)+\lambda x\big) \subset L(\{-\lambda x, \lambda x\})$. This proves half of (52). Now assume $\lambda > 2$. As before

$$\left(-S(x)-\frac{\lambda}{2}x\right) \cup \left(S(x)+\frac{\lambda}{2}x\right) \subset L\left(\left[-\frac{\lambda}{2}, \frac{\lambda}{2}\right]\right).$$

This together with (48) and the first part of (52) implies

$$L\left(\left[-\frac{\lambda}{2}, \frac{\lambda}{2}\right]\right) + \left(\frac{\lambda}{2}+1\right)U \supset \left[\left(\left(-S(x)-\frac{\lambda}{2}x\right) \cup \left(S(x)+\frac{\lambda}{2}x\right)\right) + \left(\frac{\lambda}{2}+1\right)U\right.$$

$$\left. \supset [-\lambda, \lambda] + U \supset \operatorname{conv} L(\{-\lambda x, \lambda x\})\right] \supset L(\{-\lambda x, \lambda x\}).$$

Thus $\delta([-\lambda/2, \lambda/2], \{-\lambda x, \lambda x\}) = \lambda > (\lambda/2)+1$ and the definition of δ imply that

$$\text{(53)} \qquad L(\{-\lambda x, \lambda x\}) + \mu U \supset L\left(\left[-\frac{\lambda}{2}, \frac{\lambda}{2}\right]\right)$$

holds for $\mu = \lambda$ *but not for* $\mu \in [0, \lambda[$.

In order to prove

$$(54) \qquad L(\{-\lambda x, \lambda x\})' \cap]-\lambda+\varepsilon(\lambda), \lambda-\varepsilon(\lambda)[=\emptyset$$

$$\textit{where} \quad \varepsilon(\lambda) = \max\{\lambda-\sqrt{\lambda^2-2},\ 2\lambda-2\sqrt{\lambda^2-1}\}$$

assume that there is a point $z \in L(\{-\lambda x, \lambda x\})$ with $z' \in]-\lambda+\varepsilon(\lambda), \lambda-\varepsilon(\lambda)[$. Then by the first half of (52) and (47)

$$L(\{-\lambda x, \lambda x\})+\mu U \supset [((-S(x)-\lambda x) \cup (S(x)+\lambda x) \cup \{z\})+\mu U$$

$$\supset \left[-\frac{\lambda}{2}, \frac{\lambda}{2}\right]+U = \operatorname{conv}\left(\left[-\frac{\lambda}{2}, \frac{\lambda}{2}\right]\right)+U \supset \Big] L\left(\left[-\frac{\lambda}{2}, \frac{\lambda}{2}\right]\right)$$

for some $\mu \in [0, \lambda[$, which contradicts (53). Thus (54) holds. The first half of (52) together with (54) proves the second half of (52). The same reasoning yields:

$$(55) \quad (-S(y)-\lambda y) \cup (S(y)+\lambda y) \subset L(\{-\lambda y, \lambda y\}) \subset [-\lambda y, \lambda y]+U \quad \textit{for} \quad \lambda \in \mathbf{R}^+,$$

$$L(\{-\lambda y, \lambda y\})'' \subset [(-\lambda-1)y, (-\lambda+o(1))y] \cup [(\lambda-o(1))y, (\lambda+1)y] \quad \textit{as} \quad \lambda \to +\infty.$$

Choose a coordinate system in lin $\{x, y\}$ such that $x=(1, 0)$, $y=(1/\sqrt{2}, 1/\sqrt{2})$. Then

$$z := \lambda y + \left(-\sin\frac{\pi}{8}, \cos\frac{\pi}{8}\right) \in S(y)+\lambda y \subset L(\{-\lambda y, \lambda y\})$$

by (55). The point of $L(\{-\lambda x, \lambda x\})$ which is nearest to z is of the form $\lambda x+(0, 1)+w$ with $\|w\|=o(1)$ as $\lambda \to +\infty$ by (52). Hence

$$(56) \qquad \|z-(\lambda x+(0, 1)+w)\| \geqq 2\lambda \sin\frac{\pi}{8}+\left(1-\cos\frac{\pi}{8}\right)+o(1) \quad \textit{as} \quad \lambda \to +\infty.$$

On the other hand the definition of δ and (51) imply that

$$\|z-(\lambda x+(0, 1)+w)\| \leqq \delta(L(\{-\lambda x, \lambda x\}), L(\{-\lambda y, \lambda y\})) = 2\lambda \sin\frac{\pi}{8}$$

which contradicts (56). Hence $K(o)$ cannot consist of more than one point.

Now assume that $K(o)$ consists of one point only. Then (2) implies that the isometry K of $\langle \mathscr{K}, \delta \rangle$ is generated by an isometry of E^d. Because of (25) and (14) this shows that I is generated by an isometry of E^d, concluding the proof of the theorem.

4. Acknowledgement

This paper was written during a stay of the first author at the University of Bologna. For this opportunity we should like to express our gratitude to Prof. P. L. Papini and the Consiglio Nazionale delle Ricerche.

REFERENCES

[1] BANACH, ST.: *Théorie des opérations linéaires*, Monografie Matematyczne, Vol. 1, Państwowe Wydawnictwo Naukowe, Warszawa, 1932; Chelsea, New York, 1955.

[2] BLASCHKE, W.: *Kreis und Kugel*, G. Göschen'sche Verlagshandlung, Leipzig, 1916; Chelsea, New York, 1949; Walter de Gruyter & Co., Berlin, 1956.

[3] DAY, M. M.: *Normed linear spaces*, 3rd edition, Ergebnisse der Mathematik und ihrer Grenzgebiete, Vol. 21, Springer-Verlag, Berlin—New York, 1973.

[4] GRUBER, P. M.: Isometrien des Konvexringes, *Colloquium Math.* **43** (1980), 99—109.

[5] GRUBER, P. M.: The space of compact subsets of E^d, *Geometriae Dedicata* **8** (1979), 87—90.

[6] GRUBER, P. M.: Isometries of the space of convex bodies of E^d, *Mathematika* **25** (1978), 270—278.

[7] GRUBER, P. M. and LETTL, G.: Isometries of the space of convex bodies of E^d, *Bull. London Math. Soc.* **12** (1980), 455—462.

[8] HADWIGER, H.: *Vorlesungen über Inhalt, Oberfläche und Isoperimetrie*, Die Grundlehren der Math. Wissenschaften, Bd. 93, Springer-Verlag, Berlin—Göttingen—Heidelberg, 1957.

[9] HAUSDORFF, F.: *Mengenlehre*, 2nd edition, Walter de Gruyter & Co., Berlin—Leipzig, 1927.

[10] KELLY, P. J. and WEISS, M. L.: *Geometry and convexity*, Interscience Publishers, John Wiley & Sons, Ltd., New York, 1979.

[11] KURATOWSKI, K.: *Topology*, Vol. II, Państwowe Wydawnictwo Naukowe, Warszawa, Academic Press, New York—London, 1968.

[12] POPOV, V. V.: On the space of closed subsets, *Dokl. Akad. Nauk. SSSR* **229** (1976), 1051—1054 (in Russian) = *Soviet Math. Dokl.* **17** (1976), 1145—1149.

[13] ROGERS, C. A.: *Hausdorff measures*, Cambridge University Press, Cambridge, 1970.

[14] SCHNEIDER, R.: Isometrien des Raumes der konvexen Körper, *Colloquium Math.* **33** (1975), 219—224.

[15] SENDOV, BL.: *Hausdorff approximations*, Bulg. Akad. Nauk, Sofia, 1979 (in Russian).

[16] SOURMOUR, A. R.: The isometries of $L^p(\Omega, X)$, *J. Functional Analysis* **30** (1978), 276—285.

[17] STONE, M. H.: Applications of the theory of Boolean rings to general topology, *Trans. Amer. Math. Soc.* **41** (1937), 375—381.

(Received February 6, 1980)

INSTITUT FÜR ANALYSIS
TECHNISCHE UNIVERSITÄT WIEN
GUSSHAUSSTRASSE 27
A—1040 WIEN
AUSTRIA

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 183—188

SOME APPLICATIONS OF BOOLEAN SKEW-LATTICES

by

LADISLAV BERAN

Abstract

The lower and upper symmetric difference in an ortholattice L are defined. If L is orthomodular, conditions are given such that L admits the Boolean structure. We use a theorem of Marsden on the distributive identity in orthomodular lattices to obtain a similar result for Boolean skew-lattices.

§ 1. Introduction

Let $\mathscr{L}=(L, \vee, \wedge, {}^{\perp}, 0, 1)$ be an ortholattice. Recall that the *Boolean skew-lattice* $\mathscr{L}^{\cdot}$ corresponding to an orthomodular lattice $\mathscr{L}$ can be introduced as the algebra $\mathscr{L}^{\cdot}=(L, \dot{\vee}, \dot{\wedge}, {}^{\perp}, 0, 1)$ where the derived operations $\dot{\vee}$, $\dot{\wedge}$ are defined by $a\dot{\vee} b=(a\wedge b^{\perp})\vee b$ and $a\dot{\wedge} b=(a\vee b^{\perp})\wedge b$. For basic facts about Boolean skew-lattices see [4], and for basic ortholattice definitions see [1].

§ 2. Properties of symmetric differences

Let $\mathscr{L}$ denote an ortholattice and let $a, b\in L$. By definition, $aDb=(a^{\perp}\vee b^{\perp})\wedge \wedge(a\vee b)$, $a\Delta b=(a^{\perp}\wedge b)\vee(a\wedge b^{\perp})$. The operation D (respectively Δ) is called the *upper* (respectively *lower*) *symmetric difference.*

The next lemma follows easily from the definition and is given without proof.

LEMMA 1. *Let $\mathscr{L}$ be an ortholattice and let $a, b\in L$. Then*

(i) $a\Delta b \leqq aDb$;

(ii) $(aDb)^{\perp} = a\Delta b^{\perp} = a^{\perp}\Delta b$;

(iii) $(a\Delta b)^{\perp} = aDb^{\perp} = a^{\perp}Db$.

If $\mathscr{L}$ is an ortholattice and $a, b\in L$, write aCb if and only if $a=(a\wedge b)\vee(a\wedge b^{\perp})$. In this case we say that a, b *commute.*

AMS (MOS) subject classifications (1970). Primary 06A30.

Key words and phrases. Boolean skew-lattice, orthomodular lattice.

THEOREM 2. *Let $\mathcal{L}^{\cdot}$ be the Boolean skew-lattice corresponding to an orthomodular lattice $\mathcal{L}$ and let $a, b \in L$. Then*

$$aDb = (a^{\perp} \dot{\wedge} b) \dot{\vee} (a \dot{\wedge} b^{\perp}) = (b \dot{\wedge} a^{\perp}) \dot{\vee} (b^{\perp} \dot{\wedge} a);$$

$$a\Delta b = (a^{\perp} \dot{\vee} b^{\perp}) \dot{\wedge} (a \dot{\vee} b) = (a \dot{\vee} b) \dot{\wedge} (a^{\perp} \dot{\vee} b^{\perp}).$$

PROOF. Let

$$V = (a^{\perp} \dot{\wedge} b) \dot{\vee} (a \dot{\wedge} b^{\perp}) = [(a^{\perp} \vee b^{\perp}) \wedge b] \dot{\vee} [(a \vee b) \wedge b^{\perp}].$$

Since $b^{\perp} Cb$, $b^{\perp} Ca^{\perp} \vee b^{\perp}$, $a \vee bCa$ and $a \vee bCa^{\perp} \vee b^{\perp}$, we have

$$V = [(a^{\perp} \vee b^{\perp}) \wedge b] \vee [(a \vee b) \wedge b^{\perp}] =$$

$$= \{a^{\perp} \vee b^{\perp} \vee [(a \vee b) \wedge b^{\perp}]\} \wedge$$

$$\wedge \{b \vee [(a \vee b) \wedge b^{\perp}]\} = (a^{\perp} \vee b^{\perp}) \wedge (a \vee b).$$

By a similar reasoning we also have the second assertion.

THEOREM 3. *If $\mathcal{L}$ is an orthomodular lattice and $a, b \in L$, then the following conditions are equivalent:*

(i) aCb;

(ii) $a\Delta b = aDb$;

(iii) $a\Delta b \geqq aDb$.

PROOF. That (i) implies (ii) and (ii) implies (iii) is obvious.
Now, suppose

$$(a^{\perp} \wedge b) \vee (a \wedge b^{\perp}) \geqq (a^{\perp} \vee b^{\perp}) \wedge (a \vee b).$$

Then also

$$a^{\perp} \vee (a \wedge b^{\perp}) \geqq a^{\perp} \vee [(a^{\perp} \vee b^{\perp}) \wedge (a \vee b)].$$

But $a^{\perp} Ca \vee b$ and $a^{\perp} Ca^{\perp} \vee b^{\perp}$. Thus $a^{\perp} \vee (a \wedge b^{\perp}) \geqq a^{\perp} \vee b^{\perp}$ and, consequently, $a^{\perp} \vee (a \wedge b^{\perp}) = a^{\perp} \vee b^{\perp}$. Hence $a \wedge (a^{\perp} \vee b) = a \wedge b$ and, by [1, p. 53], it follows aCb.

COROLLARY 4. Let $\mathcal{L}^{\cdot}$ be the Boolean skew-lattice corresponding to an orthomodular lattice $\mathcal{L}$. Then the following are equivalent:

(i) $\forall a, b \in L \;\; (a^{\perp} \dot{\vee} b^{\perp}) \dot{\wedge} (a \dot{\vee} b) = (a^{\perp} \vee b^{\perp}) \wedge (a \vee b)$;

(ii) $\forall a, b \in L \;\; (a^{\perp} \dot{\wedge} b) \dot{\vee} (a \dot{\wedge} b^{\perp}) = (a^{\perp} \wedge b) \vee (a \wedge b^{\perp})$.

(iii) $\mathcal{L}$ is a Boolean lattice.

THEOREM 5. *Let $\mathcal{L}^{\cdot}$ be the Boolean skew-lattice corresponding to an orthomodular lattice $\mathcal{L}$. The following are equivalent:*

(i) $\mathcal{L}^{\cdot}$ *is a Boolean lattice;*

(ii) $\mathcal{L}$ *is a Boolean lattice;*

(iii) $\forall a, b, c \in L \quad a \wedge (bDc) = (a \wedge b) D (a \wedge c)$;

(iv) $\forall a, b, c \in L \quad a \wedge (b \Delta c) = (a \wedge b) \Delta (a \wedge c)$;

(v) $\forall a, b, c \in L \quad a \dot{\wedge} (bDc) = (a \dot{\wedge} b) D (a \dot{\wedge} c)$;

(vi) $\forall a, b, c \in L \quad a \dot{\wedge} (b \Delta c) = (a \dot{\wedge} b) \Delta (a \dot{\wedge} c)$.

PROOF. It is easy to see that in a Boolean lattice the conditions (iii)—(vi) hold and that (i) and (ii) are equivalent conditions.

(v)$\Rightarrow$(ii). Substitution $b = a^{\perp}$ yields

$$l = a \dot{\wedge} (a^{\perp} Dc) = a \dot{\wedge} [(a \vee c^{\perp}) \wedge (a^{\perp} \vee c)].$$

However, $aCa \vee c^{\perp}$, $aCa^{\perp} \vee c$ and so

$$l = a \wedge (a \vee c^{\perp}) \wedge (a^{\perp} \vee c) = a \wedge (a^{\perp} \vee c) = c \dot{\wedge} a.$$

Obviously,

$$r = (a \dot{\wedge} a^{\perp}) D (a \dot{\wedge} c) = 0 \; D (a \dot{\wedge} c) = a \dot{\wedge} c.$$

Using [4, p. 54], we get aCc for every $a, c \in L$.

(vi)$\Rightarrow$(ii). Put $b = a^{\perp}$. Since $aCa \wedge c$ and $aCa^{\perp} \wedge c^{\perp}$, we then have

$$\lambda = a \dot{\wedge} (a^{\perp} \Delta c) = a \wedge (a^{\perp} \Delta c) =$$

$$= a \wedge [(a \wedge c) \vee (a^{\perp} \wedge c^{\perp})] = a \wedge c.$$

Again, we get

$$\varrho = (a \dot{\wedge} a^{\perp}) \Delta (a \dot{\wedge} c) = 0 \; \Delta (a \dot{\wedge} c) = a \dot{\wedge} c.$$

This together with [1, p. 53] gives aCc.
The proof of the rest of the theorem is now routine.

COROLLARY 6. An ortholattice $\mathcal{L}$ is orthomodular if and only if $(a \vee c) \dot{\wedge} \dot{\wedge} [(a^{\perp} \wedge c^{\perp}) Dc] = c$ for every $a, c \in L$.

THEOREM 7. *If $\mathcal{L}$ is an orthomodular lattice, then the following statements are equivalent:*

(i) $\mathcal{L}$ *is a Boolean lattice;*
(ii) *the operation Δ is associative;*
(iii) *the operation D is associative.*

PROOF. 1) Suppose Δ is associative. Since $b^{\perp} Ca^{\perp} \wedge b$ and $b^{\perp} Ca \wedge b^{\perp}$, we have

$$l = (a \Delta b) \Delta b = \{[(a^{\perp} \wedge b) \vee (a \wedge b^{\perp})]^{\perp} \wedge b\} \vee \{[(a^{\perp} \wedge b) \vee (a \wedge b^{\perp})] \wedge b^{\perp}\} =$$

$$= [(a \vee b^{\perp}) \wedge (a^{\perp} \vee b) \wedge b] \vee (a \wedge b^{\perp}) = [(a \vee b^{\perp}) \wedge b] \vee (a \wedge b^{\perp}).$$

On the other hand, $r=a\Delta(b\Delta b)=a$. By hypothesis, $l=r$ and it follows that $(a\vee b^{\perp})\wedge\wedge(a\vee b)=a\vee[(a\vee b^{\perp})\wedge b]=a$. In view of [1, p. 53] we can assert that aCb for every $a, b\in L$.

2) In order to investigate D, we observe that, by Lemma 1,

$$[(aDb)Dc]^{\perp}=(aDb)^{\perp}\Delta c=(a\Delta b^{\perp})\Delta c.$$

Similarly, $a\Delta(b^{\perp}\Delta c)=[aD(bDc)]^{\perp}$. Hence D is associative if and only if Δ is associative.

THEOREM 8. *Let x, y be elements of an orthomodular lattice. Then x and y do not commute if and only if the element xDy is the greatest and the element $x\Delta y$ the least element of a sublattice having the diagram of Figure 1:*

PROOF. 1) If xDy and $x\Delta y$ are elements with the indicated property, then x and y do not commute by Theorem 3.

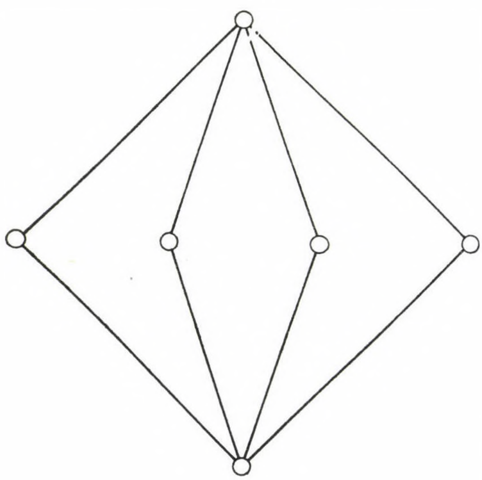


Fig. 1

2) Conversely, if x, y do not commute, then, again by Theorem 3, $xDy\neq x\Delta y$. Write, as usual, $a/b\searrow c/d$ (or $c/d\nearrow a/b$) if and only if $b\wedge c=d$ and $b\vee c=a$. Put

$$x_0=x\Delta y,\quad x_1=(x^{\perp}\vee y^{\perp})\wedge[x\vee(x^{\perp}\wedge y)],$$
$$x_2=(x^{\perp}\vee y^{\perp})\wedge[y\vee(y^{\perp}\wedge x)],$$
$$x_3=(x\vee y)\wedge[y^{\perp}\vee(y\wedge x^{\perp})],$$
$$x_4=(x\vee y)\wedge[x^{\perp}\vee(x\wedge y^{\perp})],\quad x_5=xDy.$$

By orthomodularity, we can easily show that

$$x_5/x_1\searrow x_4/x_0\nearrow x_5/x_3\searrow x_2/x_0\nearrow$$
$$\nearrow x_5/x_4\searrow x_3/x_0\nearrow x_5/x_2\searrow x_1/x_0.$$

Then since $x_0\neq x_5$, we have $x_i\neq x_j$ for every $0\leqq i\neq j\leqq 5$.

§ 3. A distributive identity

In [2] we established the basic cases in which holds the distributive identity in Boolean skew-lattices. For orthomodular lattices, a similar investigation has been studied in [3]. Here we give an analogue for a result of [3] in Boolean skew-lattices.

THEOREM 9. *Let $\mathcal{L}^{\cdot}$ be the Boolean skew-lattice corresponding to an orthomodular lattice $\mathcal{L}$. If $a, b, c, d\in L$ are such that aCb, bCc, cCd and aCd, then*

$$(a\dot\vee b)\dot\wedge(c\dot\vee d)=[(a\dot\wedge d)\dot\vee(b\dot\wedge c)]\dot\vee[(a\dot\wedge c)\dot\vee(b\dot\wedge d)]$$

if and only if

$$a\vee b\vee c^{\perp}Cd\quad\textit{and}\quad a\vee b\vee d^{\perp}Cc.$$

PROOF. 1) By assumption, $a\dot\vee b=a\vee b$ and $c\dot\vee d=c\vee d$. Therefore,

$$l=(a\dot\vee b)\dot\wedge(c\dot\vee d)=(a\vee b)\dot\wedge(c\vee d)=$$
$$=[(a\vee b\vee(c^{\perp}\wedge d^{\perp})]\wedge(c\vee d).$$

Notice that $l=\{a\vee[(b\vee c^{\perp})\wedge(b\vee d^{\perp})]\}\wedge(c\vee d)$ since $bCc^{\perp}$ and $c^{\perp}Cd^{\perp}$. Now, the element $b\vee d^{\perp}$ commutes with all the elements of the subalgebra $\langle a, b, c, d\rangle$ of $\mathscr{L}$ generated by a, b, c, d and so

$$l=(a\vee b\vee c^{\perp})\wedge(a\vee b\vee d^{\perp})\wedge(c\vee d).$$

Further,

$$r=[(a\dot{\wedge} d)\dot{\vee}(b\dot{\wedge} c)]\dot{\vee}[(a\dot{\wedge} c)\dot{\vee}(b\dot{\wedge} d)]=$$

$$=[(a\wedge d)\dot{\vee}(b\wedge c)]\dot{\vee}\{[(a\vee c^{\perp})\wedge c]\dot{\vee}[(b\vee d^{\perp})\wedge d]\}.$$

In order to simplify this expression, observe that $(a\vee c^{\perp})\wedge c$ and $(b\vee d^{\perp})\wedge d$ commute. Hence,

$$r=\{[a\wedge d\wedge(b^{\perp}\vee c^{\perp})]\vee(b\wedge c)\}\dot{\vee}\{[(a\vee c^{\perp})\wedge c]\vee[(b\vee d^{\perp})\wedge d]\}=$$

$$=\{[(a\wedge d\wedge(b^{\perp}\vee c^{\perp}))\vee(b\wedge c)]\wedge[c^{\perp}\vee(a^{\perp}\wedge c)]\wedge[(b^{\perp}\wedge d)\vee d^{\perp}]\}\vee$$

$$\vee\{[(a\vee c^{\perp})\wedge c]\vee[(b\vee d^{\perp})\wedge d]\}.$$

Here we have $a\wedge d\wedge(b^{\perp}\vee c^{\perp})Cb\wedge c$, $b\wedge cCc^{\perp}$, $c^{\perp}Ca^{\perp}\wedge c$ and $a^{\perp}\wedge cCa\wedge d\wedge(b^{\perp}\vee c^{\perp})$. By [3, Corollary 5], it follows

$$r=[(a^{\perp}\wedge b\wedge c)\vee(a\wedge c^{\perp}\wedge d)]\wedge[d^{\perp}\vee(b^{\perp}\wedge d)]\vee\{\dots\}.$$

It is easy to check that $a^{\perp}\wedge b\wedge cCa\wedge c^{\perp}\wedge d$, $a\wedge c^{\perp}\wedge dCd^{\perp}$, $d^{\perp}Cb^{\perp}\wedge d$ and $b^{\perp}\wedge dCa^{\perp}\wedge\wedge b\wedge c$. By the same result we get

$$r=(a\wedge b^{\perp}\wedge c^{\perp}\wedge d)\vee(a^{\perp}\wedge b\wedge c\wedge d^{\perp})\vee$$

$$\vee[(a\vee c^{\perp})\wedge c]\vee[(b\vee d^{\perp})\wedge d].$$

2) Suppose $l=r$. This implies

$$(b\vee d^{\perp})\wedge d\leqq a\vee b\vee c^{\perp} \tag{1}$$

and also

$$(a\vee c^{\perp})\wedge c\leqq a\vee b\vee d^{\perp}. \tag{2}$$

From (1) we conclude that

$$b\vee[(b\vee d^{\perp})\wedge d]\leqq a\vee b\vee c^{\perp}.$$

Using the fact that $bCb\vee d^{\perp}$ and $b\vee d^{\perp}Cd$, we get $(b\vee d^{\perp})\wedge(b\vee d)\leqq a\vee b\vee c^{\perp}$. Since the element $a\vee c^{\perp}$ commutes with all the elements of the subalgebra $\langle a, b, c, d\rangle$, we can see that

$$\begin{cases}a\vee b\vee c^{\perp}\geqq[(b\vee d^{\perp})\wedge(b\vee d)]\vee(a\vee c^{\perp})=\\=(a\vee b\vee c^{\perp}\vee d^{\perp})\wedge(a\vee b\vee c^{\perp}\vee d).\end{cases} \tag{1'}$$

This yields $a\vee b\vee c^{\perp}=(a\vee b\vee c^{\perp}\vee d^{\perp})\wedge(a\vee b\vee c^{\perp}\vee d)$, i.e.,

$$a\vee b\vee c^{\perp}Cd. \tag{1''}$$

The same reasoning applies on (2) and so we have also

$$a\vee b\vee d^{\perp}Cc. \tag{2''}$$

3) Actually, the conditions (1), (1″) ((2), (2″)) are equivalent. In fact, if (1″) is valid, then from (1′) we can deduce that

$$(b\vee d^{\perp})\wedge d \leqq (a\vee b\vee c^{\perp}\vee d^{\perp})\wedge(a\vee b\vee c^{\perp}\vee d) = a\vee b\vee c^{\perp}.$$

4) Suppose now (1″) and (2″) are satisfied. Then (1) and (2) hold and it is clear that

$$r = (a\wedge b^{\perp}\wedge c^{\perp}\wedge d)\vee(a^{\perp}\wedge b\wedge c\wedge d^{\perp})\vee[(a\vee c^{\perp})\wedge c]\vee[(b\vee d^{\perp})\wedge d] \leqq$$
$$\leqq (a\vee b\vee c^{\perp})\wedge(a\vee b\vee d^{\perp})\wedge(c\vee d) = l.$$

In order to prove the converse inequality, first note that the elements $a^{\perp}\wedge b\wedge \wedge c\wedge d^{\perp}$, $a\wedge b^{\perp}\wedge c^{\perp}\wedge d$ commute with all the elements of $\langle a, b, c, d\rangle$. Combining this remark with the fact that

$$[(a\vee c^{\perp})\wedge c]\vee[d\wedge(b\vee d^{\perp})] =$$
$$= (a\vee c^{\perp}\vee d)\wedge(c\vee d)\wedge(a\vee b\vee c^{\perp}\vee d^{\perp})\wedge(b\vee c\vee d^{\perp}),$$

by $a\vee c^{\perp}Cc$, cCd, $dCb\vee d^{\perp}$, $a\vee c^{\perp}Cb\vee d^{\perp}$ and [3, Corollary 5], we conclude

$$r = [a\vee c^{\perp}\vee d\vee(a^{\perp}\wedge b\wedge c\wedge d^{\perp})]\wedge(c\vee d)\wedge$$
$$\wedge(a\vee b\vee c^{\perp}\vee d^{\perp})\wedge[b\vee c\vee d^{\perp}\vee(a\wedge b^{\perp}\wedge c^{\perp}\wedge d)].$$

Since $a\vee c^{\perp}$ commutes with all the elements of $\langle a, b, c, d\rangle$, we get for the first expression here

$$[\ldots] = d\vee(a\vee c^{\perp})\vee[(a^{\perp}\wedge c)\wedge(b\wedge d^{\perp})] =$$
$$= d\vee a\vee c^{\perp}\vee(b\wedge d^{\perp}) = d\vee(a\vee b\vee c^{\perp})\wedge(a\vee c^{\perp}\vee d^{\perp}).$$

Since (1″) holds, the element $a\vee b\vee c^{\perp}$ commutes with all the elements of $\langle a, b, c, d\rangle$. Therefore, $[\ldots]=a\vee b\vee c^{\perp}\vee d$. Analogously, the second expression $[\ldots]$ gives $[\ldots]= =a\vee b\vee c\vee d^{\perp}$. This means that

$$r = (a\vee b\vee c^{\perp}\vee d)\wedge(c\vee d)\wedge(a\vee b\vee c^{\perp}\vee d^{\perp})\wedge(a\vee b\vee c\vee d^{\perp}) \geqq$$
$$\geqq (a\vee b\vee c^{\perp})\wedge(c\vee d)\wedge(a\vee b\vee d^{\perp}) = l$$

and, thus, $r=l$, as claimed by the theorem.

REFERENCES

[1] Birkhoff, G.: *Lattice Theory*, Amer. Math. Soc. Colloq. Publ., Vol. XXV, Providence, R. I. 1967.
[2] Beran, L.: Über gewisse Sätze vom Foulis—Holland-Typ in Boole'schen Zwerchverbänden, *J. Reine Angew. Math.* **297** (1978), 214—220.
[3] Beran, L.: Formulas for orthomodular lattices, *Studia Sci. Math. Hungar.* **11** (1976), 451—455.
[4] Kröger, H.: Das Assoziativgesetz als Kommutativitätsaxiom in Boole'schen Zwerchverbänden, *J. Reine Angew. Math.* **285** (1976), 53—58.

(Received February 14, 1980)

DEPARTMENT OF ALGEBRA OF THE
CHARLES UNIVERSITY
SOKOLOVSKÁ 83
ČS — 186 00 PRAHA 8 — KARLIN
CZECHOSLOVAKIA

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 189—191

ON A PROBLEM OF P. ERDŐS AND E. G. STRAUS

by

J. DÉNES and K. H. KIM

In [6] P. ERDŐS and E. G. STRAUS suggested the following problem:

Given a group G with finite commutator subgroup G', then there exist n (not necessarily distinct) elements $x_1, x_2, \ldots, x_n$ of G such that

$$G' = \{x_{i_1} x_{i_2} \ldots x_{i_n} \colon (i_1, \ldots, i_n) \in S_l\}.$$

That is, G' consists of the product of all the x_i in all possible order. We know that we can choose $n \leqq 3|G'|$. What better bounds for n in terms of $|G'|$ can be obtained? Is it possible to choose distinct $x_1, x_2, \ldots, x_n$?

The aim of this paper is to obtain better bounds for n.

One of the present authors defined P- and C-groups in the following way:

DEFINITION. A finite group G of order l $(a_i \in G,\ i=1, 2, \ldots, l)$ is called a P-group if every element b of its commutator subgroup G' or mG' (where m is an element of order two of G such that $m \notin G'$) can be represented by the following product $b = a_{i_1} a_{i_2} \ldots a_{i_l}$, $(i_1, i_2, \ldots, i_l) \in S_l$. G is called a C-group if every element of G' is a commutator itself (see [3]).

P- and C-groups and their interconnections were studied in [1], [2], [3]. In [10] A. R. RHEMTULLA proved that every finite solvable group is a P-group. (Rhemtulla's result implies that if G is a finite solvable group of order l then $n \leqq l+1$ holds.)

THEOREM 1. *Let G be a P-group of order l whose commutator subgroup is of order r and of index 2. Then $n \leqq 2r+1$ holds.*

PROOF. Let $a_1, a_2, \ldots, a_l$ be the elements of G then clearly all the products $a_{i_1} a_{i_2} \ldots a_{i_l}$ $(i_1, i_2, \ldots, i_l) \in S_l$ are included in $G \setminus G'$ or in G'. Since G is a P-group, every element of $G \setminus G'$ (or G') can be represented as a product $a_{i_1} a_{i_2} \ldots a_{i_l}$. If $\prod_{j=1}^{l} a_{i_j}$ is in G' then the number of factors is l and, obviously, $l < 2r+1$. In the case when $\prod_{j=1}^{l} a_{i_j}$ is in $G \setminus G'$ then there exists at least one element b of G such that $b \prod_{j=1}^{l} a_{i_j} \in G'$ and $l+1 = 2r+1$ holds and this completes our proof.

Theorem 2 will show that even for groups in Theorem 1 there exists better bound than the one in Theorem 1.

THEOREM 2. *Let G be a finite group and let G' be a perfect (i.e., $G' = G''$) P-group of order r then $n \leqq r$ holds.*

PROOF. Since G' is a perfect P-group of order r every element of G' can be represented as a product whose factors are distinct elements of G'. (By a quite recent joint result of P. HERMANN and one of the present authors: every finite group is a P-group (see [5]), so Theorem 2 is valid without supposing that G' is a P-group.)
With the aid of the above theorem one can obtain a further result which can be considered as a generalization of Theorem 2, see Remark below.)

COROLLARY. *For* S_l $(l \geqq 5)$, $n \leqq \frac{l!}{2}$ *holds.*

This Corollary is implied by Theorem 2.
An alternative proof is as follows:
In [7], [9] O. ORE and N. ITO independently proved that A_l $(l \geqq 5)$, i.e., the alternating group of degree l, is a C-group. (For a generalization of Ore—Ito Theorem see [8].) It is known (see [1]) that if a group G of order s is a C-group and the number of its elements of order 2 is less than $\frac{s}{2}$ then it is a P-group. This implies that A_l $(l \geqq 5)$ is a P-group. Since the alternating group is the commutator subgroup of the symmetric group the Corollary is proved.

REMARK. In [3] one of the present authors published his conjecture that every finite group is a P-group. A quite recent joint result of P. HERMANN and one of the present authors gave an affirmative solution of this conjecture. (The proof of that result will be published elsewhere, see [5].) With the aid of the above result one can easily prove the following theorem:

THEOREM 3. *Let G_0 be a finite non-solvable group if G_k denotes the commutator subgroup of G_{k-1} $(k=1, 2, \ldots, t)$ and G_t is perfect of order s and furthermore the order of the factor group G_1/G_t is u, then every element of G_1 can be represented as a product*

$$a_{i_1} a_{i_2} \ldots a_{i_{s+u-1}} \quad (i_1 i_2 \ldots i_{s+u-1}) \in S_l$$

such that all the factors are distinct.

Until now (in Theorems 1, 2 and 3) we used products with distinct factors to represent the elements of the commutator subgroups. In the next theorem we shall show that for elements which are not necessarily distinct we might have better bounds.

THEOREM 4. *For* A_l $(l \geqq 5)$ $n \leqq \frac{l(l-1)(l-2)(l-3)}{2}$ *holds.*

PROOF. Clearly, the set H of all double transpositions (i.e., the permutations of the form $(ab)(cd)$) form a generating system of A_l $(l \geqq 5)$. H is a conjugate class of S_l which does not split into two in A_l, see [4]. The cardinality of H is equal to $\frac{l(l-1)(l-2)(l-3)}{8}$. It is easy to prove that every element of A_l $(l \geqq 5)$ can be represented as a product whose factors are distinct elements of H. Let us suppose that $\alpha \in A_l$ and $\alpha = \beta_1 \beta_2 \ldots \beta_i \ldots \beta_k$ where $\beta_j \in H$ $(j=1, 2, \ldots, k)$ and $\beta_1 = \beta_i$ since

$$\beta_1 = \beta_1^{-1}, \; \alpha = \beta_2' \beta_3' \ldots \beta_{i-1}' \beta_{i+1} \ldots \beta_k$$

where

$$\beta_1 \beta_2 \beta_1 = \beta_2', \quad \beta_1 \beta_3 \beta_1 = \beta_3', \ldots, \beta_1 \beta_{i-1} \beta_1 = \beta_{i-1}',$$

and clearly β'-s are elements of H. This procedure can be iterated until we obtain a product with distinct factors which form a subset H_α of H. Since every element of A_l ($l \geqq 5$) is a commutator (see [7], [9]) and $H_\alpha = H_{\alpha^{-1}}$, every element α of A_l ($l \geqq 5$) can be represented in the following way:

$$\alpha = \beta\gamma\beta^{-1}\gamma^{-1} = \beta_1\beta_2 \dots \beta_k\gamma_1\gamma_2 \dots \gamma_{k'}\beta_k \dots \beta_2\beta_1\gamma_{k'} \dots \gamma_2\gamma_1;$$

where β_i ($i=1, 2, \dots, k$) are all distinct elements of H_β and γ_j ($j=1, 2, \dots, k'$) are all distinct elements of H_γ.

If the elements of $H \setminus H_\beta$ and $H \setminus H_\gamma$ are $\beta'_i \left(i=k+1, \dots, \frac{l(l-1)(l-2)(l-3)}{8}\right)$, $\gamma'_j \left(j=k'+1, \dots, \frac{l(l-1)(l-2)(l-3)}{8}\right)$, respectively, then obviously the identity element of A_l can be represented as $\prod_{i=k+1}^{v} \beta_i'^2 \prod_{j=k'+1}^{v} \gamma_i'^2$ where $v = \frac{l(l-1)(l-2)(l-3)}{8}$. In this way we used every element of H four times exactly we have a product with $4v$ factors. This proves the Theorem.

The results of this paper were published as a preprint of Department of Information and Computer Sciences, Faculty of Engineering Science, Osaka University Toyonaka, Osaka, Japan, February 1979. The authors are indebted for the referee of this paper who called their attention to the validity of Theorem 4.

REFERENCES

[1] DÉNES, J.: On some properties of commutator subgroups, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **7** (1964), 123—127.

[2] DÉNES, J.: On commutator subgroups of finite groups, *Comment. Math. Univ. Sancti Pauli Tokyo* **15** (1967), 61—65.

[3] DÉNES, J.: On some properties of commutator subgroups, Lecture given at the International Mathematical Congress held in Moscow 1966. Abstract, Section 2, p. 86.

[4] DÉNES, J.—ERDŐS, P. and TURÁN, P.: On some statistical properties of the alternating group of degree *n*, *L'Enseignement mathématique* **15** (1969), 89—99.

[5] DÉNES, J. and HERMANN, P.: On the product of all elements in a finite group (to appear).

[6] ERDŐS, P. and STRAUS, E. G.: Groups. Research Problem. Combinatorics Proc., Canberra, 1977; Lecture Notes in Math. Vol. 686, Springer-Verlag, Berlin—Heidelberg—New York, 1978, p. 347.

[7] ITO, N.: A theorem on alternating group, *Math Japonicae* **2** (1951), 59—60.

[8] JACQUES, A.—LENORMAND, C.—LENTIN, A. and PERROT, J. F.: Un résultat extrémal en théorie des permutations, *C. R. Acad. Sci. Paris* **266** (1968), 446—448.

[9] ORE, O.: Some remarks on commutators, *Proc. Amer. Math. Soc.* **2** (1951), 307—314.

[10] RHEMTULLA, A. R.: On a problem of L. Fuchs, *Studia Sci. Math. Hungar.* **4** (1969), 195—200.

(Received February 26, 1980; in revised form July 1, 1980)

SZÁMÍTÁSTECHNIKAI KOORDINÁCIÓS INTÉZET
POSTAFIÓK 224
H—1368 BUDAPEST
AKADÉMIA U. 17
HUNGARY

MATHEMATICS RESEARCH GROUP
ALABAMA STATE UNIVERSITY
MONTGOMERY, AL 36101
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 193—202

SOME PROBLEMS OF ANALYTIC NUMBER THEORY II

by

R. BALASUBRAMANIAN and K. RAMACHANDRA

§ 1. Introduction

In the present note we report on some progress made by us towards the following

PROBLEM. *Let $A(x)$ be a complex valued integrable (in the sense of Riemann) function defined in $0<x<\infty$ by $A(x)=0$ in $0<x<1$, and subject to $|A(x)|\leqq(x+2)^C$ where C is a positive constant. Consider the function*

$$F(s)=\int_0^\infty \frac{dA(x)}{x}=s\int_0^\infty \frac{A(x)\,dx}{x^{s+1}},\quad (s=\sigma+it),$$

which is plainly analytic in $\sigma>C$. Suppose that $F(s)$ can be continued analytically in $\sigma\geqq\theta_0-\delta$ (where θ_0 and δ are constants satisfying $\theta_0\geqq 2\delta>0$) except possibly along certain lines $t=t_j$ $(j=1, 2, 3, \ldots)$. We assume that t_j are distinct and do not have a limit point except possibly at ∞. We also assume that there exists a singularity of $F(s)$ in $\sigma>\theta_0-\delta$ and that θ_0 is the least upper bound of the singularities of $F(s)$. The singularities have to exist on the lines $t=t_j$. They can be poles, algebraic singularities, logarithmic singularities, or essential singularities of any kind. We assume that they do not have a limit point in the finite part of the half plane $\sigma\geqq\theta_0-\delta$ and that no singularity lies on $\sigma=\theta_0-\delta$. Let ϱ_0 be a singularity in $\sigma>\theta_0-\delta$ and denote by $D(\varrho_0)$ a disk with centre ϱ_0 within which the only singularity is ϱ_0. Let L_0 denote the anticlockwise boundary of $D(\varrho_0)$, with the following point removed. Namely we remove that point of intersection of the boundary with the line $t=\operatorname{Im}\varrho_0$ which lies to the left of ϱ_0. We now put

$$B(x)=\frac{A(x+0)+A(x-0)}{2}x^{-\varrho_0},$$

and

$$R(x)=\left|\int_{L_0}(F(s))x^{s-\varrho_0}\frac{ds}{s}\right|.$$

Then the problem is to find effective positive constants Λ, C_1, C_2, C_3, such that if $m=[C_2 y]$ then for all $y\geqq C_3$ there holds

$$\frac{1}{\Lambda^m}\int_0^\Lambda\cdots\int_0^\Lambda |B(\exp(y+u_1+\ldots+u_m))|\,du_1\,du_2\ldots du_m>$$
$$>C_1\left|\int_{L_0}(F(s)s)\exp(y(s-\varrho_0))\frac{ds}{s}\right|.$$

Further if $F(s)$ is regular on the line $\sigma \geqq \mathrm{Re}\, \varrho_0$, $t=0$, and $A(x)$ is real, then to determine effective positive constants C_4, C_5, C_6 such that for all $y \geqq C_6$ there hold the inequalities,

$$M\big((e^y, e^{C_4 y})\big) > C_5 R(e^y)$$

and

$$m\big((e^y, e^{C_4 y})\big) > -C_5 R(e^y),$$

where $M(I)$ and $m(I)$ denote the maximum (resp. minimum) of $B(x)x^{(\mathrm{Im}\,\varrho_0)}$ as x varies over the interval I.

REMARK 1. We give a solution of the problem where the only singularities of $F(s)$ in $\sigma \geqq \theta_0 - \delta$ are poles. Next we make some comments about the general problem.

REMARK 2. This note is an attempt to give an effective version of a well-known theorem of E. Landau.

REMARK 3. A number of special cases follow. For example if $A_2(x) = \sum_{n \leqq x} |\mu(n)| - \frac{6x}{\pi^2}$ we can take $2\varrho_0$ = the first zero of $\zeta(s)$ on the critical line. It follows that the maximum M and the minimum m of $A_2(x)x^{-\frac{1}{4}}$ in the interval $e^y \leqq x \leqq e^{100y}$ satisfy $M \geqq 10^{-1000}$ and $m \leqq -10^{-1000}$, respectively, for all $y \geqq 10$. A similar result is also true for instance if

$$A_k(x) = \sum_{n \leqq x} \Big|\sum_{d^k/n} \mu(d)\Big| - \frac{x}{\zeta(k)}$$

where $k \geqq 2$ is an integer. Here $\frac{1}{4}$ has to be replaced by $\frac{1}{2k}$. The other constants are effective and will depend on k. These results give the solution of the conjecture which we made at the end of § 1 of our earlier paper [1], in fact in a sharper form.

§ 2. Solution of the problem when the singularities are poles

We begin with

LEMMA 1. *If $x > 0$, we have*

$$\frac{A(x+0)+A(x-0)}{2} = \frac{1}{2\pi i} \int_{C+3-i\infty}^{C+3+i\infty} F(s)\frac{x^s}{s}\, ds.$$

PROOF. It follows from the formula

$$\frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \frac{y^s}{s}\, ds = 0,\ \frac{1}{2} \quad \text{or} \quad 1,$$

according as $0<y<1$, or $y=1$ or $y>1$.

We next introduce a definition.

DEFINITION. With any pair of real numbers λ, g with $1<\lambda\leqq 2$, $g\geqq 1$ we introduce a killing operator $T_{\varrho_0, g, \lambda}$ which destroys the singularities $\varrho\neq\varrho_0$, of $F(s)$ in the rectangle $R(\varrho_0, g, \lambda)$, defined by $\sigma\geqq\operatorname{Re}\varrho_0$, $|\operatorname{Im}\varrho_0-t|\leqq g$, as follows. We write

$$T_{\varrho_0, g, \lambda}(F(s)) = F(s)\{\prod_{\substack{\varrho\neq\varrho_0 \\ \varrho \text{ in } R(\varrho_0, g, \lambda)}} (1-\lambda^{s-\varrho})\}^a$$

where a is the maximum of the orders of the poles in the rectangle $R(\varrho_0, g, \lambda)$. The product may destroy the singularity ϱ_0 as well. To avoid this we choose λ such that the distance $\dfrac{2\pi}{\log\lambda}$ between the zeros of $1-\lambda^{it}$ is big enough depending on g.

LEMMA 2. *Let the complex numbers a_n be defined by*

$$T_{\varrho_0, g, \lambda}(F(s)) = F(s)\sum_{n=0}^{N} a_n\lambda_n^s$$

where $\lambda_n=\lambda^n$ and N is a suitable integer depending only on ϱ_0, g, λ. Then we have

$$\sum_{n=0}^{N} a_n\lambda_n^{\varrho_0}B(\lambda_n x) = \frac{1}{2\pi i}\int_{C+3-i\infty}^{C+3+i\infty} T_{\varrho_0, g, \lambda}(F(s))x^{s-\varrho_0}\frac{ds}{s}.$$

PROOF. It follows by Lemma 1.

LEMMA 3. *We have,*

$$\sum_{n=0}^{N} a_n\lambda_n^{\varrho_0}B(\lambda_n x) = \frac{1}{2\pi i}\int_{L_0} T_{\varrho_0, g, \lambda}(F(s))\frac{x^{s-\varrho_0}}{s}ds + \frac{1}{2\pi i}\int_{L_1} T_{\varrho_0, g, \lambda}(F(s))\frac{x^{s-\varrho_0}}{s}ds,$$

where L_0 is already defined and L_1 is defined as follows. Now $T_{\varrho_0, g, \lambda}(F(s))$ is regular in $\sigma\geqq\operatorname{Re}\varrho_0$, $|\operatorname{Im}\varrho_0-t|<g$ except at ϱ_0. We may assume that the point omitted in the boundary of $D(\varrho_0)$ to form L_0 is $\varrho_0-\delta_1$, and that $T_{\varrho_0, g, \lambda}(F(s))$ is regular in $\sigma\geqq\operatorname{Re}(\varrho_0-\delta_1)$, $|\operatorname{Im}\varrho_0-t|\leqq g$, except at ϱ_0. Let $P_1=C+3-i\infty$, $P_2=C+3-ig+$ $+i\operatorname{Im}\varrho_0$, $P_3=\varrho_0-\delta_1-ig$, $P_4=\varrho_0-\delta_1-io$, $P_5=\varrho_0-\delta_1+io$, $P_6=\varrho_0-\delta_1+ig$, $P_7=$ $=C+3+ig+i\operatorname{Im}\varrho_0$ and $P_8=C+3+i\infty$. L_1 denotes the shortest path connecting $P_1P_2P_3P_4$ in this order and then the shortest path connecting $P_5P_6P_7P_8$ in this order.

PROOF. It follows by Cauchy's theorem.

LEMMA 4. *Let $A\geqq 1$ be a real constant and $m\geqq 1$ an integer. Put $dv=du_1du_2\ldots du_m$ and $f(x)=\sum a_n\lambda_n^{\varrho_0}B(\lambda_n x)$. Then for $y\geqq 1$, we have,*

$$A^{-m}\int_0^A\ldots\int_0^A f(e^{y+u_1+\ldots+u_m})\,dv = I_0+I_1,$$

where

$$I_0 = A^{-m}\int_0^A\ldots\int_0^A\left(\frac{1}{2\pi i}\int_{L_0} T_{\varrho_0, g, \lambda}(F(s))\exp\left((s-\varrho_0)(y+u_1+\ldots+u_m)\right)\frac{ds}{s}\right)dv$$

and

$$I_1 = A^{-m}\int_0^A\ldots\int_0^A\left(\frac{1}{2\pi i}\int_{L_1} T_{\varrho_0, g, \lambda}(F(s))\exp\left((s-\varrho_0)(y+u_1+\ldots+u_m)\right)\frac{ds}{s}\right)dv.$$

PROOF. It follows from Lemma 3.

LEMMA 5. *We have*

$$I_1 = O\left(\exp(-y\delta_1) + \left(\frac{2}{gA}\right)^m \exp((y+mA)(C+3))\right)$$

where the O-constant depends only on ϱ_0, g, λ and is effective.

PROOF. The paths P_3P_4 and P_5P_6 contribute $O(\exp(-y\delta_1))$, and the rest of the path L_1 contributes the remaining part of the O-term.

LEMMA 6. *Let $g=14A\exp(A(C+3))$ and $m=[(C+6)y]$. Further let $0<\delta_1<\log 2$. Then, we have,*

$$I_1 = O(\exp(-y\delta_1)),$$

where the O-constant is effective.

PROOF. It follows from Lemma 5.

LEMMA 7. *We have*

$$|I_0| \geqq C_7 \left| \int_{L_0} F(s)\exp((s-\varrho_0)y)\frac{ds}{s}\right|$$

where $C_7>0$ is an effective constant.

PROOF. It follows since ϱ_0 is a pole of $T_{\varrho_0, g, \lambda}(F(s))$ and also of $F(s)$, of the same order.

LEMMA 8. *If $dv=du_1\dots du_m$ then for some n in $0<n<N$, we have,*

$$A^{-m}\int_0^A \dots \int_0^A |B(\lambda_n \exp(y+u_1+\dots+u_m))|\,dv \geqq$$

$$\geqq C_7\left(\sum_{n=0}^{N} |a_n|\lambda_n^{\mathrm{Re}\,\varrho_0}\right)^{-1} \left| \int_{L_0} F(s)\exp((s-\varrho_0)y)\frac{ds}{s}\right|.$$

PROOF. It follows from Lemma 7 and the definition of $f(x)$.

LEMMA 9. *Let us write* Re $\varrho_0=\theta_1$. *Let $F(s)$ be regular in $\sigma\geqq\theta_1+\varepsilon$, $|t|\leqq h$. Let $F(s)$ be regular in $\sigma\geqq\theta_1$, $t=0$ and so in $\sigma\geqq\theta_1-\delta_2$, $|t|\leqq\tau$. Let $A(x)$ be real for all $x>0$, and let*

$$B_1(x) = \left(\frac{1}{2}(A(x+0)+A(x-0))\right)x^{-\theta_1}.$$

Then we have, with $dv=du_1du_2\dots du_m$,

$$A^{-m}\int_0^A \dots \int_0^A B_1(\lambda_n \exp(y+u_1+\dots+u_m))\,dv =$$

$$= O\left(e^{-y\delta_2} + \left(\frac{2}{\tau A}\right)^m \exp(\varepsilon(y+mA)) + \left(\frac{2}{Ah}\right)^m \exp((y+mA)(C+3))\right),$$

where the O-constant is effective and $m\geqq 1$ is an integer.

PROOF. We start from the formula

$$B_1(x) = \frac{1}{2\pi i} \int_{C+3-i\infty}^{C+3+i\infty} F(s) \frac{x^{s-\theta_1}}{s} \, ds, \quad (x > 0),$$

and obtain by multiple integration (with the notation $dv = du_1 \dots du_m$)

$$\Lambda^{-m} \int_0^{\Lambda} \dots \int_0^{\Lambda} B(\lambda_n \exp(y+u_1+\dots+u_m)) \, dv =$$

$$= \Lambda^{-m} \sum_{j=1}^{2^m} \pm \left(\frac{1}{2\pi i} \int_{C+3-i\infty}^{C+3+i\infty} (F(s)\lambda_n^s) \frac{e^{y_j(s-\theta_1)}}{(s-\theta_1)^m s} \, ds \right),$$

where $y_j = y + j'\Lambda$ (j' being an integer satisfying $0 \leqq j' \leqq m$) and the combinations of signs is the obvious one. Consider a typical summand in the last expression. We deform the path to be the shortest distance joining the points $P_9, P_{10}, P_{11}, P_{12}, \dots, P_{18}$ in that order. Here $P_9 = C+3-i\infty$, $P_{10} = C+3-ih$, $P_{11} = \theta_1 + \delta - ih$, $P_{12} = \theta_1 + \varepsilon - i\tau$, $P_{13} = \theta_1 - \delta_2 - i\tau$, $P_{14} = \theta_1 - \delta_2 + i\tau$, $P_{15} = \theta_1 + \varepsilon + i\tau$, $P_{16} = \theta_1 + \varepsilon + ih$, $P_{17} = C+3+ih$, $P_{18} = C+3+i\infty$. The paths $P_9 P_{10} P_{11}$ and $P_{16} P_{17} P_{18}$ contribute

$$O\left(\left(\frac{2}{\Lambda h} \right)^m \exp((y+m\Lambda)(C+3)) \right),$$

and the paths $P_{11} P_{12} P_{13}$ and $P_{14} P_{15} P_{16}$ contribute $O\left(\left(\frac{2}{\Lambda \tau} \right)^m \exp(\varepsilon(y+m\Lambda)) \right)$ while the path $P_{13} P_{14}$ contributes $O(\exp(-y\delta_2))$ to the last expression involving the sum of 2^m terms. This completes the proof of Lemma 9.

LEMMA 10. *We may choose $\tau = \delta_2 < \log 2 < h$ without loss of generality. Next we put $\varepsilon \leqq \frac{1}{\Lambda}$, $\Lambda = 100\tau^{-1}$, $h = 14\Lambda \exp(\Lambda(C+3))$ and $m = [(C+6)y]$. Then we have,*

$$\Lambda^{-m} \int_0^{\Lambda} \dots \int_0^{\Lambda} B_1(\lambda_n \exp(y+u_1+\dots+u_m)) \, dv = O(\exp(-y\delta_2)),$$

where $dv = du_1 \dots du_m$ and the O-constant is effective.

PROOF. It follows from Lemma 9.

We can now complete the solution of the problem which we proposed to solve in this section. Lemma 8 solves the first part. Observe that $B(x)$ and $B_1(x)$ have the same absolute value. Lemmas 8 and 10 give, for $y \geqq C_9$,

$$\Lambda^{-m} \int_0^{\Lambda} \dots \int_0^{\Lambda} \left| B_1(\lambda_n \exp(y+u_1+\dots+u_m)) \right| dv > C_8 R \tag{1}$$

and

$$\left| \Lambda^{-m} \int_0^{\Lambda} \dots \int_0^{\Lambda} B_1(\lambda_n \exp(y+u_1+\dots+u_m)) \, dv \right| < (100)^{-1} C_8 R, \tag{2}$$

respectively, where $dv=du_1\dots du_m$ and

$$R=\left|\int_{L_0} F(s)\exp\big((s-\varrho_0)y\big)\frac{ds}{s}\right|.$$

Here C_8 and C_9 are effective. Plainly, ϱ_0 can be chosen to satisfy the requirements of Lemma 8 and 10. Let us consider the error in the LHS of (1) by excluding the integration over the subregion V where $|B_1(\lambda_n\exp(\dots))|$ does not exceed $(100)^{-1}C_8R$. If $B_1(\lambda_n\exp(\dots))$ does not change sign in the complement of V we first limit the region of integration in LHS of (2) to this complement and then take the absolute value sign from the integral to the integrand. We see that

$$C_8R(1-(100)^{-1})<((100)^{-1}+2(100)^{-1})C_8R$$

which is a contradiction which proves the result required.

§ 3. Some comments on the general problem

The method of § 2 shows that it is possible to work with the case when θ_0 is attained and the singularities on the line $\sigma=\theta_0$ are of the form $\left(\frac{1}{s-\varrho}\right)^{\alpha}\left(\log\frac{1}{s-\varrho}\right)^{\beta}$ where α and β are real constants. Because the killing operator, though it does not annihilate the singularities, reduces very much the contributions $\int_{L_\varrho} T(F(s))\frac{x^s}{s}\,ds$ and so on, where L_ϱ is the latch associated with ϱ. For an asymptotic evaluation of the contributions from these latches see Remark 2 on page 316 of [3]. It is clearly possible to solve the problem in slightly greater generality. But we have not been able to solve it in the generality stated in the introduction.

§ 4. Concluding remarks

The results of this paper deal with the oscillatory properties of kernels of Dirichlet integrals due to its singularities. We have also some results on effective oscillation coming from the growth properties. For example divisor problem, circle problem, Abelian group problem (in the last case we prove $\Omega(x^{1/6}(\log x)^{1/2})$, $\Omega_{\pm}(x^{1/14})$. In the two other cases we prove $\Omega(x^{1/4})$, $\Omega_{\pm}(x^{1/8})$) and further generalities. We had decided to publish these results in this paper itself. But we publish the oscillations resulting from the growth in another paper [2].

§ 5. Appendix

After writing the first draft of the paper we came across a paper of Kátai, where* he has proved some results very similar to the ones proved in this paper. The methods of proof are also almost similar. We show, in this appendix, that

* Kátai, I., On oscillations of number theoretic functions, *Acta Arith.* **13** (1967/68), 107—123.

our results are slightly stronger than those of Kátai, by improving his theorems. For the sake of simplicity, we restrict only to the Ω version of $M(x)=\sum_{n\le x}\mu(n)$ and the general result can also be proved by exactly same method.

Let $\frac{1}{2}+i\tau$ be the first zero of $\zeta(s)$ above the real axis and b the absolute value of the residue of $(\zeta(s+i\tau))^{-1}$ at $s=\frac{1}{2}$. Further we assume that there exists a constant B ($\geqq 100$) such that the region ($\sigma>\frac{1}{2}$; $|t|\leqq B$) is free of zeros of $\zeta(s)$. Let

$$m \geqq \frac{y}{2\log B}+10$$

and define

$$M_1(x) = \sum_{n\le x}\frac{\mu(n)}{n^{i\tau}}.$$

Then we have the

THEOREM. *There exists an absolute constant C_1, such that for all*

$$y \geqq C_1 \log B \log\log B,$$

we have

$$\max_{e^y\le x\le e^{y+m}}\left|\frac{M_1(x)}{x^{1/2}}\right| \geqq \frac{b}{20}.$$

COROLLARY 1. *There exists an absolute constant C_2, such that*

$$\max_{C_2 e^{y/2}\le u\le e^{y+m}}\left|\frac{M(u)}{u^{1/2}}\right| \geqq \frac{b}{40}.$$

COROLLARY 2 *(Under Riemann Hypothesis). There exist absolute constants C_2 and C_3 such that*

$$\max_{C_2 e^{y/2}\le u\le e^y y^{C_3}}\left|\frac{M(u)}{u^{1/2}}\right| \geqq \frac{b}{40}.$$

REMARK. Corollary 1 improves Kátai's result, which reads

$$\max_{X\le u\le X^2\exp\left(2C\frac{\log\log B\log X}{\log B}\right)}\left|\frac{M(u)}{u^{1/2}}\right| \gg 1 \quad \text{for some} \quad C>0.$$

PROOF. Consider the integral

$$J=\frac{1}{2\pi i}\int_{1+\frac{1}{y}-i\infty}^{1+\frac{1}{y}+i\infty}\frac{1}{\zeta(s+i\tau)}\frac{e^{(s-(1/2))y}}{s}\left(\frac{e^{s-(1/2)}-1}{s-(1/2)}\right)^m ds.$$

Let L be the contour defined by joining the points

$$1+\frac{1}{y}-i\infty,\ 1+\frac{1}{y}-iB,\ \frac{1}{2}+\frac{1}{y}-iB,\ \frac{1}{2}+\frac{1}{y}-i,\ \frac{1}{2}-\varepsilon-i,$$

$$\frac{1}{2}-\varepsilon+i,\ \frac{1}{2}+\frac{1}{y}+i,\ \frac{1}{2}+\frac{1}{y}+iB,\ 1+\frac{1}{y}+iB,\ 1+\frac{1}{y}+i\infty$$

by straight lines in this order and call them L_1, L_2, L_3, L_4, L_5, L_6, L_7, L_8 and L_9 in this order. Move the line of integration of J to L. Define $J_j = \frac{1}{2\pi i}\int\limits_{L_j} \dots$. Then $J = R + \sum_{j=1}^{9} J_j$ where R is the residue at the point $s=\frac{1}{2}$.

LEMMA 11. *We have* $|R| \geqq 2|b|$.

PROOF. This is clear.

LEMMA 12. *We have* $|J_k| \leqq \frac{b}{100}$, *for* $k=1, 2, 8$ *and* 9.

PROOF. *In* J_1, $|\zeta(s+i\tau)|^{-1} \ll |s|$ $(|t| \geqq B)$,

$$\left|\left(\frac{e^{s-(1/2)}-1}{s-(1/2)}\right)^m\right| \leqq \left(\frac{1}{B}\right)^{m-5}\frac{1}{|t|^5},$$

$\left|e^{(s-1/2)y}\right| \ll e^{y/2}$ and this proves the result about J_1. The other results are similar.

LEMMA 13. *There exists an absolute constant* $C_4 < 1$ *such that*

$$\left|\frac{e^{(s-(1/2))}-1}{s-(1/2)}\right| \leqq C_4, \quad \sigma = \frac{1}{2}+\frac{1}{y}, \quad |t| > 1.$$

PROOF. We have

$$\left|\frac{e^{(s-(1/2))}-1}{s-(1/2)}\right| = \left|\int_0^1 e^{(s-(1/2))u}\,du\right| =$$

$$= \left|\int_0^1 e^{iut}\,du + O\left(\frac{1}{y}\right)\right| = \left|\frac{\sin(t/2)}{t/2} + O\left(\frac{1}{y}\right)\right| \leqq C_4.$$

LEMMA 14. *We have* $J_k \leqq \frac{b}{100}$, $k=3, 4, 6$ *and* 7.

PROOF. The result follows easily from Lemma 13 and the fact that

$$|\zeta(s+i\tau)|^{-1} \ll |t|^{C\log y} \quad \text{on} \quad \sigma = \frac{1}{2}+\frac{1}{y}.$$

LEMMA 15. *We have* $J_5 \leqq \frac{b}{100}$.

PROOF. Since, in J_5,

$$|\zeta(s+i\tau)|^{-1} \ll 1; \quad \left|\left(\frac{e^{(s-(1/2))}-1}{s-(1/2)}\right)^m\right| \ll 1;$$

$$|e^{(s-(1/2))y}| < e^{-\varepsilon y},$$

the result follows.

From Lemmas 11 to 15, there follows

LEMMA 16. *We have* $J \geqq \frac{b}{2}$.

PROOF OF THE THEOREM.

$$J = \int_0^1 du_1 \int_0^1 du_2 \dots \int_0^1 du_m \int_{C-i\infty}^{C+i\infty} \frac{1}{\zeta(s+i\tau)} e^{(y+u_1+u_2+\dots+u_m)(s-(1/2))} \frac{ds}{2\pi i s} \geqq \frac{b}{2}$$

(here $\int_{C-i\infty}^{C+i\infty}$ is to be taken as $\lim_{T\to\infty} \int_{C-iT}^{C+iT}$). Consequently, there exist $u_1, u_2, \dots, u_m$ such that

$$\left| \int_{C-i\infty}^{C+i\infty} \frac{1}{\zeta(s+i\tau)} e^{(y+u_1+\dots+u_m)(s-(1/2))} \frac{ds}{2\pi i s} \right| \geqq \frac{b}{2}$$

and $0 \leqq u_1, u_2, \dots, u_m \leqq 1$.

Putting $x = e^{y+u_1+u_2+\dots+u_m}$, we have

$$\left| x^{-(1/2)} \sum_{n \leqq x} \frac{\mu(n)}{n^{i\tau}} \right| \geqq \frac{b}{2}$$

and this proves the theorem.

PROOF OF THE COROLLARIES. Fix x as in the theorem. Then

$$|M(x)| = \left| \int_1^x \frac{1}{u^{i\tau}} d(M(u)) \right| \leqq |M(x)| + \int_1^x \tau \left| \frac{M(u)}{u} \right| du.$$

Hence

$$|M(x)| + \int_1^x \tau \left| \frac{M(u)}{u} \right| du \geqq \frac{b}{2} x^{1/2}.$$

Since

$$\int_1^{(b/4)x^{1/2}} \tau \left| \frac{M(u)}{u} \right| du \leqq \frac{b}{4} x^{1/2},$$

it follows that

$$M(x) + \int_{(b/4)x^{1/2}}^{x} \left| \frac{M(u)}{u} \right| du \geqq \frac{b}{4} x^{1/2}.$$

Hence

$$\max_{(b/4)x^{1/2} \leqq u \leqq x} \left| \frac{M(u)}{u^{1/2}} \right| \geqq \frac{b}{20}$$

and this proves Corollary 1.

In Corollary 1, take B such that $y = C_1 \log B \log\log B$. This proves Corollary 2.

REFERENCES

[1] BALASUBRAMANIAN, R. and RAMACHANDRA, K.: Effective and non-effective results on certain arithmetical functions, *J. Number Theory* **12** (1980), 10—19.
[2] BALASUBRAMANIAN, R. and RAMACHANDRA, K.: Some problems of analytic number theory III, *Hardy—Ramanujan Journal* **4** (1981), 1—12.
[3] RAMACHANDRA, K.: Some problems of analytic number theory I, *Acta Arith.* **31** (1976), 313—324.

(Received March 5, 1980)

SCHOOL OF MATHEMATICS
TATA INSTITUTE OF FUNDAMENTAL RESEARCH
HOMI BHABHA ROAD
IND—400 005 BOMBAY
INDIA

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 203—207

SOLID PACKING OF CIRCLES IN THE HYPERBOLIC PLANE

by

A. BEZDEK

Let us recall the fact [1] that the face-incircles of the regular trihedral tilings $\{2, 3\}$, $\{3, 3\}$, $\{4, 3\}$, $\{5, 3\}$ and $\{6, 3\}$ form a densest packing of 3, 4, 6, 12 and infinitely many equal circles, respectively. Can these results be extended to the hyperbolic tilings $\{7, 3\}$, $\{8, 3\}$, ...? This question started some investigations [2, 3] which led to the perception that in the hyperbolic plane it is impossible to define any "reasonable" notion of density. In order to avoid this difficulty L. FEJES TÓTH [4] introduced the notion of the solidity of a packing: A packing of convex discs is said to be *solid* if no finite subset of the discs can be rearranged (by a set of rigid motions) so as to form, together with the rest of the discs, a packing not congruent with the original one. With this notion a nice result [4, 5] can be phrased: For any integer $p>1$ the face-incircles of $\{p, 3\}$ constitute a solid packing. The conjecture that also the face-incircles of any spherical, Euclidean or hyperbolic trihedral Archimedean tiling form a solid packing, has been confirmed in many cases [3, 6].

In this paper we want to deal with another conjecture of L. Fejes Tóth according to which the face-incircles of a tiling $\{p, 3\}$ with $p\geqq 6$ form a *strongly solid packing* in the sense that removing one of the circles the remaining circles continue to be solidly packed. This conjecture was suggested by the corresponding considerations on the sphere. Removing one of the face-incircles of $\{2, 3\}$ or $\{3, 3\}$ the remaining three circles will have a "free" play. On the other hand, removing one of the six face-incircles of $\{4, 3\}$ the remaining five circles will have only a rather restricted play: two of the circles must be centered, say, in the north- and south-pole, while the rest of the centres must lie on the equator. Finally, the packing of eleven circles arising by removing one of the twelve face-incircles of $\{5, 3\}$ is, apart from rotations of the sphere, uniquely determined. Thus the face-incircles of $\{5, 3\}$ form a strongly solid packing. Even more can this be expected for $\{6, 3\}$, $\{7, 3\}$, ...

We shall prove the following

THEOREM. *The face-incircles of a tiling $\{p, 3\}$ with $p\geqq 8$ form a strongly solid packing.*

The problem for $p=6$ and 7 is still open.

First we consider the case when $p=8$.

Let $S=\{c_1, c_2, ...\}$ be the set of circles which arises from the face-incircles $c_0, c_1, ...$ of $\{8, 3\}$ by removing c_0. We say that a subset $s=\{c_1, ..., c_n\}$ of S is *good*

AMS (MOS) subject classifications (1979). Primary 52A45; Secondary 51M10.
Key words and phrases. Solid packing, hyperbolic plane.

if there is no other rearrangement of the circles $c_1, \ldots, c_n$ than the trivial one which results in a packing congruent to S. If the property of being good is transmitted from s to the union of s and c_{n+1}, we say that c_{n+1} can be joined to s. Starting with a set s_0 consisting of a single circle of S, which is obviously good, we shall successively join to s_0 new circles of S so as to obtain an infinite set of enlarged sets $s_0 \subset s_1 \subset s_2 \subset \ldots$ so that $s_0 \cup s_1 \cup s_2 \cup \ldots = S$.

We start with the following

REMARK. If the circle c_n cannot be joined to $\{c_1, \ldots, c_{n-1}\}$ then in a non-trivial rearrangement of $c_1, \ldots, c_n$ at least two circles overlap the original place of c_n.

To see this we observe that in a non-trivial rearrangement at least two circles are not inscribed in faces of $\{8, 3\}$. Thus, supposing that only one (or none) of the rearranged circles overlap the original place of c_n, we put this circle (or any of the rearranged circles) to this place obtaining a non-trivial rearrangement of $c_1, \ldots, c_{n-1}$. This contradicts the tacit assumption that $\{c_1, \ldots, c_{n-1}\}$ is good.

LEMMA 1. *If from among the circles touching c_n at most two belong to $\{c_0, c_1, \ldots, c_{n-1}\}$ then c_n can be joined to $\{c_1, \ldots, c_{n-1}\}$.*

We suppose that c_n cannot be joined to $\{c_1, \ldots, c_{n-1}\}$ and consider the locus of the centres of those circles which overlap c_n without overlapping $c_{n+1}, c_{n+2}, \ldots$. If from the circles $c_0, \ldots, c_{n-1}$ only one or only two not adjacent circles touch c_n then this locus is empty. On the other hand, if among $c_0, \ldots, c_{n-1}$ there are only two adjacent circles which touch c_n then the locus is an open point-set whose diameter is equal to the diameter $2r$ of a circle c_i (Fig. 1). Thus, in contradiction to the above remark, we have at most one circle overlapping c_n.

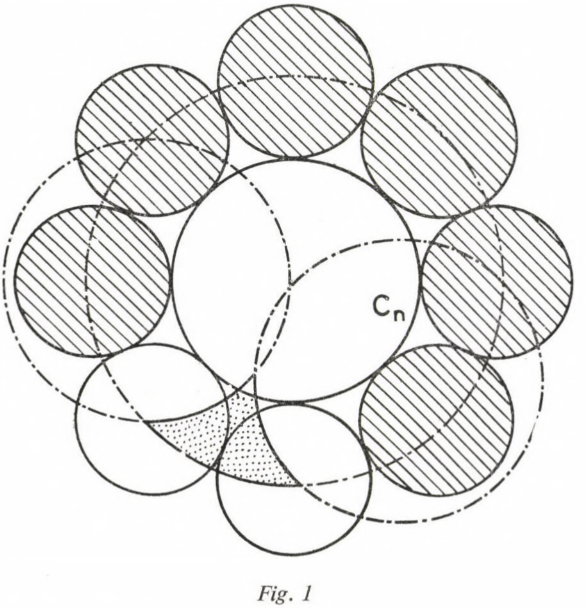


Fig. 1

LEMMA 2. *Let the circles in the following triples mutually touch one another: (c_1, c_2, c_n), (c_2, c_3, c_n), (c_1, c_2, c_4), (c_2, c_3, c_5). If in the set $\{c_0, c_1, \ldots, c_5, \ldots, c_{n-1}\}$ there is, apart from c_4, c_2 and c_5, no circle touching c_1 or c_n or c_3 then c_n can be joined to $\{c_1, \ldots, c_n\}$.*

We suppose that c_n cannot be joined to $\{c_1, \ldots, c_{n-1}\}$ and consider the circles $c_1, \ldots, c_n$ in a not trivially rearranged position X. Since $\{c_1, \ldots, c_{n-1}\}$ is good, $\{c_2, c_3, c_4, c_5, \ldots, c_{n-1}\}$ is also good. By Lemma 1 c_n can be joined to $\{c_2, \ldots, c_{n-1}\}$, but because of our supposition c_1 cannot be joined to $\{c_2, \ldots, c_n\}$. Thus at least two circles of X overlap the original place of c_1.

Let O_i be the centre of c_i. The locus of the centres of circles which overlap c_1 without overlapping $c_{n+1}, c_{n+2}, \ldots$ is the open "triangle" $T = O_i O_4 O_n$ bounded by arcs of circles of radius $2r$ concentric with c_1 and the circles other than c_2 touching on the one hand c_1 and c_4, on the other hand c_1 and c_n (Fig. 2). Let T_1 and T_2 be $C \cap T$ resp. $T \setminus C$ where C denotes the circle of radius $2r$ concentric with c_4. Since the diameter of T_1 is equal to $2r$ only one of the centres of the circles overlapping c_1 lies in T_1. The centre of another circle of X lies in T_2.

We claim that this circle contains the point of contact P of c_2 and c_n. Let O_k be the centre of the circle containing the side $O_1 O_n$ of T. Since $\sphericalangle O_4 O_1 O_k = 3\pi/4 < \pi$, the circle C intersects the side $O_1 O_n$ in a point, say, Q. By some computation we

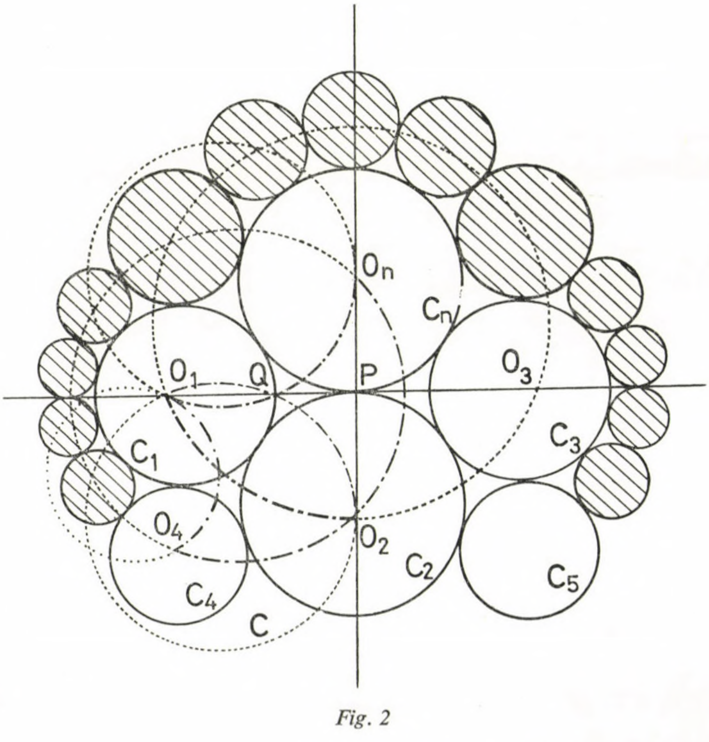


Fig. 2

obtain

$$\operatorname{sh} PQ = \frac{\sqrt{1+\sqrt{2}}}{3} < \frac{1}{\sqrt[4]{2}} = \operatorname{sh} r,$$

i.e., $PQ<r$, showing that the circle of radius r centred at P contains T_2. This proves the assertion.

The same argument shows that in X there is a circle overlapping c_3 and containing P. But the circles containing P and overlapping on the one hand c_1, on the other hand c_3 must be identical. Besides this circle no other circle of X can overlap c_n because then either the circular-arc-triangle $O_1O_2O_n$ or its image $O_3O_2O_n$ reflected in the line O_2O_n would contain two points at distance $>2r$. This is impossible since the diameter of $O_1O_2O_n$ is equal to $2r$. Thus, in view of our Remark, we have a contradiction to the assumption that c_n cannot be joined to $\{c_1, ..., c_5, ..., c_{n-1}\}$. This proves the lemma.

We write $Z_0=\{c_0\}$ and define the mth zone Z_m of c_0 as the set of those circles of S which touch at least one circle of Z_{n-1} but do not belong to a zone Z_n with $k<n$. We consider the union $U_n=Z_1\cup Z_2\cup\ldots\cup Z_n$. Since to any finite subset s of S there is an index N such that $s\subset U_N$, the theorem will be proved by showing that for any integer $n>0$ U_n is good.

In order to obtain a better insight into the structure of a zone, it is convenient to consider besides $\{8, 3\}$ the dual tiling $\{3, 8\}$. Let V_1 be the union of the triangles of $\{3, 8\}$ meeting at O_0. Add to V_1 all triangles of $\{3, 8\}$ which have a boundary-point in common with a triangle of V_1 obtaining V_2, and so on. The triangles of V_1 fill a regular octagon without overlapping and without interstices. Since the triangles added to V_1 fill a ring without overlapping and without interstices, the triangles of V_2 form a simple polygon. Since the angles of this polygon are equal either to $2\pi/4$ or to $3\pi/4$, the polygon is convex. The same holds for the nth polygon P_n formed by the triangles of V_n.

Fig. 3 shows a part of $\{3, 8\}$. The horizontal lines represent the boundaries of P_1, P_2, P_3 and P_4. The vertical edges issuing downwards from the vertices of P_1 meet at O_0. The construction of the figure can easily be continued. Since the vertices of P_n are the centres of the circles of Z_n, and the centres of adjacent circles are connected with edges of $\{3, 8\}$, we see that each circle of Z_n has either one or two neigh-

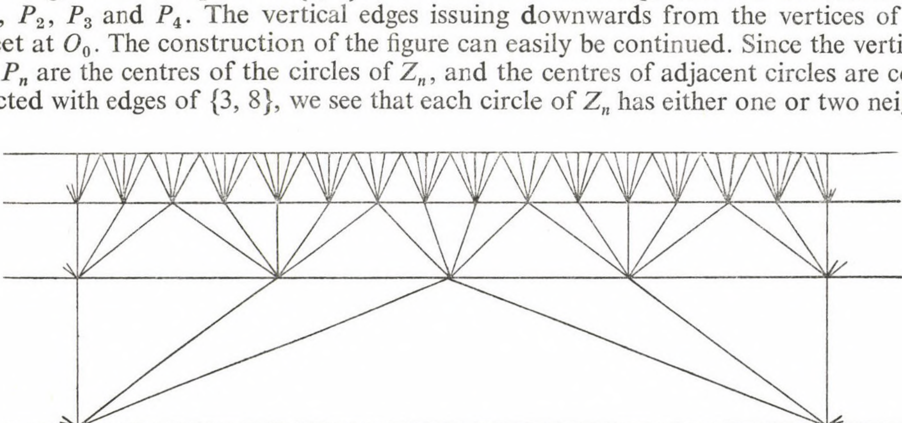


Fig. 3

bours in Z_{n-1}. Calling the respective circles of type 1 and type 2, we observe that in any zone Z_n with $n>1$ between two consecutive circles of type 2 there are either two or three circles of type 1.

Now we start with a circle of Z_1. By Lemma 1 we can join to this circle the other circles of Z_1 successively in their cyclic order except the last one. The last circle can be joined to the previous ones by Lemma 2, showing that Z_1 is good. Referring to Lemma 1 we consecutively join to Z_1 all circles of Z_2 of type 2, then all circles of Z_2 of type 1 adjacent to those of type 2. The remaining circles of type 1 which lie between two circles of type 1 can be joined to the previous ones by Lemma 2. Thus U_2 is also good.

In Z_n with $n>2$ there are, besides consecutive circles of type 21112, also circles of type 2112. Now we proceed as follows. Using Lemma 1, again we first join to U_{n-1} all circles of Z_n of type 2. Then in all quadruples of type 2112 we join one of the circles of type 1. The other circle of type 1 can be joined by Lemma 2. The rest of the circles of type 1 can be joined similarly as in Z_2.

This completes the proof of the theorem for $p=8$.

The above proof can be applied also in the case when $p>8$. Now the inequality $PQ<r$ can easily be proved without computation. Let O be the centre of the triangle $O_1O_2O_n$. Since $\sphericalangle O_4O_2O=\frac{3}{2}\frac{2\pi}{p}\leqq\frac{\pi}{3}=\sphericalangle O_4OO_2$, we have $O_4O<O_2O_4=2r$, so that O lies within the circle with centre O_4 and radius $2r$. It follows that $QP<OP$. But because of $\sphericalangle OO_2P=\frac{\pi}{p}\leqq\frac{\pi}{9}<\frac{\pi}{3}=\sphericalangle O_2OP$, we have $OP<PO_2=r$, and thus $QP<r$.

Finally, let us observe that in a zone Z_n with $n>1$ generated by $\{p, 3\}$ there are between two circles of type 2 either $p-5$ or $p-6$ circles of type 1. Therefore there are for $p>8$ no quadruples of type 2112 which slightly simplifies the proof.

REFERENCES

[1] FEJES TÓTH, L.: *Lagerungen in der Ebene, auf der Kugel und im Raum*, Springer-Verlag, Berlin—Heidelberg—New York, second edition, 1972.

[2] FEJES TÓTH, L.: Kreisausfüllungen der hyperbolischen Ebene, *Acta Math. Acad. Sci. Hungar.* **4** (1953), 103—110.

[3] BÖRÖCZKY, K.: Packing of spheres in spaces of constant curvature I, *Mat. Lapok* **25** (1974), 265—306.

[4] FEJES TÓTH, L.: Solid circle-packings and circle-coverings, *Studia Sci. Math. Hungar.* **3** (1968), 401—409.

[5] IMRE, M.: Kreislagerungen auf Flächen konstanter Krümmung, *Acta Math. Acad. Sci. Hungar.* **15** (1964), 115—121.

[6] FEJES TÓTH, G.: Solid sets of circles, *Studia Sci. Math. Hungar.* **9** (1974), 101—109.

(Received March 14, 1980)

DEZSŐ U. 11., I. EM. 1
H—1016 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 209—226

ON A GENERALIZATION OF THE GAME GO-MOKU I

by

M. AJTAI, L. CSIRMAZ and ZS. NAGY

Abstract

We investigate the winning and drawing strategies in the generalized Go-moku games which are defined in 1.1. It is proved that every open game is equivalent, in a certain sense, to some generalized Go-moku. We give an example of a recursive game in which player I has a winning strategy but has no recursively enumerable one. Examples are constructed for non-determined games of length $\omega+20$ and $\omega\cdot 2$ under certain set-theoretical assumptions. Finally, we determine the possible minimal lengths of winning strategies.

0. Introduction

The well-known game Go-moku [5] is played on an infinite chessboard. Two players, I and II, occupy the squares alternately. The winner is the player who has first 5 or more adjacent squares in a row, horizontally, vertically or along either diagonal. An easy argument shows (see, e.g., Proposition 2.1 below) that II cannot have a winning strategy (WS in the sequel) and either I has a WS or II has a drawing strategy (DS), i.e., a strategy which allows II to play indefinitely.

As far as our knowledge goes, it is not known which is the case in Go-moku. It is not even known whether a WS of I, if any, can be bounded in time. In other words, assume that I has a WS. Is there a natural number n such that I can win before his n-th move? The best available result is in [6, pp. 257—258]:

THEOREM 0.1. *If the game Go-moku is played on countably many boards (i.e., at every step the player chooses a board and occupies a square on it) and* I *has a WS then* II *does not have arbitrarily long counterplay. (A WS can always be bounded in time.)* ∎

The game n-Go-moku is a slight modification of the original game. The winner is required to have at least n adjacent squares. It is an easy exercise to prove the

PROPOSITION 0.2. I *wins* 4-*Go-moku in at most* 6 *moves.* ∎

On the other hand we succeeded in proving (cf. [1])

THEOREM 0.3. *For* $n>8$, I *has no WS in n-Go-moku.*

PROOF (Sketch). We describe a strategy of II which prevents I from occupying 9 adjacent squares. Divide the board into pieces of size 5×5 along the border-lines of the squares. Let II play independently in each of these pieces. (If there is no more room for his move, occupy a square in any other piece.) The goal of II in each of these pieces is to prevent I from occupying either a full row, a full column, or one of the ten diagonals shown in Fig. 1. It can be checked that II can achieve his goal. I cannot have 9 or more adjacent squares because their intersection with some piece would be a part which I was prevented from occupying. ∎

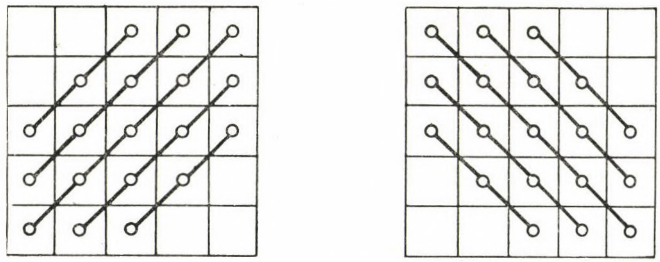


Fig. 1

We know nothing about the missing cases $5 \leqq n \leqq 8$[1].

These results led to some generalizations of the game Go-moku, one of which is discussed in detail here. For other generalizations, see our forthcoming paper.

1. Definitions

Our set-theoretical notation will be standard. Ordinal numbers are denoted by α, β, etc., ω denotes the first infinite ordinal as well as the cardinality of countable sets. The cardinality of the continuum is denoted by 2^{ω}. Sequences are enclosed between angular brackets $\langle$ and $\rangle$, for example the empty sequence is denoted by $\langle\rangle$.

The family of all finite 0—1 sequences is denoted by ${}^{\underset{\smile}{\omega}}2$, and ${}^{\omega}2$ is the family of 0—1 sequences of length ω. If $\sigma \in {}^{\omega}2$ then $\sigma | n \in {}^{\underset{\smile}{\omega}}2$ is the unique initial segment of σ of length n. Let s and t be sequences; the relation $s \prec t$ holds if s is a proper initial segment of t.

1.1. $\mathfrak{A}$-*games*

The $\mathfrak{A}$-game $[A, F]^{\alpha}$ consists of the board A, the family F of nonempty finite subsets of A which are the winning sets, and the ordinal number α which is an upper bound for the length of the game.

$\mathfrak{A}$-games are played by two players, I and II. They occupy elements of A alternately. Every element can be chosen at most once. I begins and every limit step (if any) is I's turn. The game ends if either I or II occupies all elements of some $X \in F$ (covers X), the winner is the one who does it. The game ends if all elements of A have been chosen or if α moves have been made. In these cases the game is a draw. If $\alpha = \omega$, the game is said to be finite.

[1] R. K. Guy kindly informed us that Andreas Brouwer and others can prove Theorem 0.3 for $n=7$ and $n=8$.

1.2. $\mathfrak{B}$-*games*

The $\mathfrak{B}$-game $[A, F^0, F^1]^\alpha$ consists of A and α as above, and of the families F^0 and F^1 of finite subsets of A. $\mathfrak{B}$-games are played by I and II as follows. First I chooses an element of $\{0, 1\}$ which we denote by k. After this they occupy elements of A alternately, starting with I if $k=0$, and with II if $k=1$. The title of the player who starts picking is W, the other's is B (from white and black). Every limit step (if any) is I's turn. The game ends if either I covers some element of F^k, or II covers some element of F^{1-k}; the winner is the one who does it. The game ends if all elements of A have been chosen or if α moves have been made. In these cases the game is a draw. If $\alpha=\omega$, the game is said to be finite.

1.3. *Positional games*

Positional games contain (at least an evidently equivalent form of) every infinite two-person game of perfect information of length ω where the families of the winning positions of both players are open sets [4]. The components of a positional game are a sequence of sets $\langle A_i \colon i\in\omega\rangle$, and two disjoint sets F_W and F_B of finite sequences such that $\langle a_0, a_1, ..., a_k\rangle \in F_W \cup F_B$ implies $a_i \in A_i$ $(i \leqq k)$. We assume that no sequence in $F_W \cup F_B$ is a proper initial segment of any other element.

Two players, W and B take turns alternately. First W picks $a_0 \in A_0$ then B picks $a_1 \in A_1$, again W picks $a_2 \in A_2$, etc. W wins if $\langle a_0, a_1, ..., a_k\rangle \in F_W$ for some k, B wins if $\langle a_0, a_1, ..., a_k\rangle \in F_B$ for some $k\in\omega$, otherwise the game is a draw.

Evidently, every finite $\mathfrak{A}$ and $\mathfrak{B}$-game can be transformed into a positional game. This transition can be done by a recursive function.

1.4. *Snub-games*

Let G be any positional game. The snub-G snub-game is played by I and II as follows. First I chooses who he wants be: W or B. If he chooses W, he begins the game G as W and II plays as B. Otherwise II begins the game as W and I plays as B. The other moves go by the rules of G. The winner of the game G is the winner of the game snub-G.

1.5. *Strategies, equivalent games*

The notion of strategy and that of play according to a strategy is discussed in [8]. The strategy S is a I-winning strategy, I-WS in short, if every play according to S is a win for I. The strategy S is a I-drawing strategy, I-DS, if every play according to S is either a win for I or is a draw. Similarly, for II-WS, II-DS, etc.

The winning strategy S is *α-bounded*, if there is a $\gamma<\alpha$ such that every play according to S ends before the γ-th move. In case of $\mathfrak{A}$- and $\mathfrak{B}$-games, the I-winning strategy S is *bounded in space*, if there is a finite subset of the board such that in every play according to S, I occupies elements of this subset only. Hence boundedness in space implies ω-boundedness.

A game is *determined* if either both players have DS or one of them has WS.

Let the games G_0, and G_1 be played, say, by I and II. We say that G_0 is *equivalent* to G_1, if there are recursive functions φ_0 and φ_1 (in the sense of [3]) such that

(i) $\varphi_i(G_i)=G_{1-i}$, $(i=0, 1)$;

(ii) if S is I-WS (I-DS, II-WS, II-DS) in G_i then $\varphi_i(S)$ is a I-WS (I-DS, II-WS, II-DS) in G_{1-i} $(i=0, 1)$.

Obviously, every finite $\mathfrak{A}$- and $\mathfrak{B}$-game is equivalent to some positional game. Equivalent games are equidetermined.

2. Basic results

PROPOSITION 2.1. *In $\mathfrak{A}$-games* II *has no WS.*

PROOF. Suppose the contrary and let S be a II-WS. Let I play by S as follows. Choose any $t_0 \in A$ (an "imagined" element) and let I's first move be the answer by S to this imagined move. If II's answer is not the imagined element, answer I simply by S. If it is, drop t_0 as an imagined element, choose an entirely new one and regard it as the move of the other. At limit steps I always has to choose a new imagined element. Now just the strategy ensures I to cover a winning set before II could do it, which is a contradiction. ∎

The same argument as above shows that

PROPOSITION 2.2. *In $\mathfrak{A}$-games, $\mathfrak{B}$-games, and snub-games* II *has no WS. If* II *has DS then* I *has DS, too.* ∎

REMARK. This proposition remains true even if we allow the winning sets to be infinite.

PROPOSITION 2.3. *Positional games are determined.*

PROOF. This is the GALE—STEWART result for open games [4]. If I has no WS, II make a move such that I still has no WS. Since if I wins he wins after finitely many moves, this strategy is a DS for II. ∎

COROLLARY 2.4. *Finite $\mathfrak{A}$ and $\mathfrak{B}$-games as well as snub-games are determined. In these games* I *always has DS.* ∎

PROPOSITION 2.5. *Even in the case of finite $\mathfrak{A}$-games, the existence of a* I-*WS does not imply the existence of an ω-bounded WS; the existence of an ω-bounded* I-*WS does not imply the existence of a space-bounded WS.*

PROOF. We give two examples which witness the assertions. Let first B_1 and B_n for $n \geqq 2$ be the set of nodes of the trees shown in Fig. 2.

Let $F_i=\{$the full branches of $B_i\}$, for example, every element of F_1 consists of four nodes. I has WS in $[\bigcup_{i \geqq 1} B_i, \bigcup_{i \geqq 1} F_i]^\omega$ because I can win in B_1, but II may postpone his defeat for n moves by threatening in B_n.

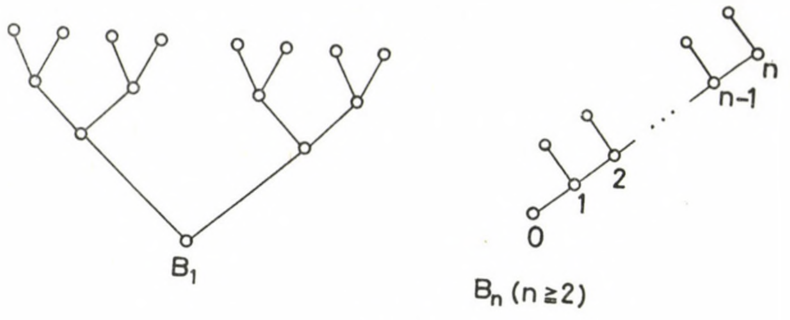


Fig. 2

In the second example let the board be $A=\{R\}\cup\{P_j^i\colon\ i\in\omega,\ j\leqq 7\}$, and let $A_k=\{R\}\cup\{P_j^i\colon\ i<k,\ j\leqq 7\}$. The winning sets are

$$\{R, P_0^i, P_1^i\},\ \{R, P_0^i, P_2^i, P_3^i\},\ \{R, P_0^i, P_2^i, P_4^i, P_5^i\},\ \{R, P_0^i, P_2^i, P_4^i, P_6^i\}$$

and

$$\{P_1^i, P_3^i, P_5^i, P_7^j\} \quad \text{for all} \quad i\leqq j\in\omega$$

(see Fig. 3). I wins this game playing as follows. Start with occupying R. The response of II is an element of some A_k, then pick $P_0^k, P_2^k, P_4^k, P_6^k$ in succession. On the other hand II has counterplay in every A_k picking first either R or P_7^{k-1}. ∎

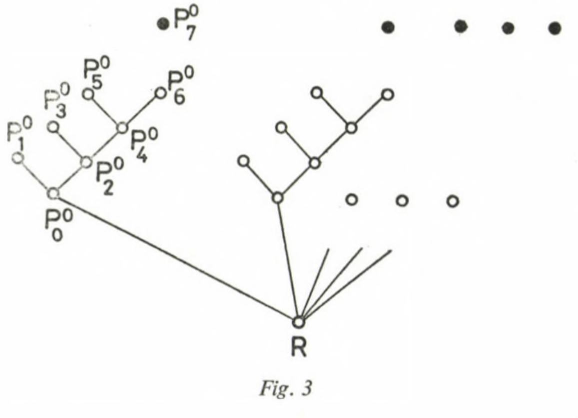


Fig. 3

We define the *rank* of I-WS of finite $\mathfrak{A}$-games as follows. Winning strategies can be regarded as trees, the root is I's first move, the edges starting from the root are labelled by the possible moves of II, the nodes at the other end are I's responses by the strategy, etc. To each node v assign the least ordinal which is greater than the ordinals assigned to the successors of v. This definition is sound because these trees are well-founded. The ordinal assigned to the root is the rank of the strategy. For example a I-WS is ω-bounded if and only if its rank is less than ω.

The rank of a finite $\mathfrak{A}$-game is the infimum of the ranks of I-winning strategies. So the rank of the first example in Proposition 2.5 is ω. The construction described in the proof of Theorem 4.6 gives

PROPOSITION 2.6. *For every ordinal α there is a finite $\mathfrak{A}$-game of rank* $\geqq \alpha$. ∎

3. Equivalence of finite games

In Section 1 we have remarked that every finite $\mathfrak{A}$-game is equivalent to some positional game. The converse cannot be true because there are positional games where II has WS. But in view of 2.4 we may hope for

THEOREM 3.1. *Every snub-game is equivalent to some finite $\mathfrak{A}$-game.*

PROOF. The theorem is an immediate consequence of Lemmas 3.7 and 3.8 below. ∎

To describe the construction we shall need the following structure.

DEFINITION 3.2. Let Γ be an index set. The *broom* associated with Γ is the 7-tuple $\langle B, B^0, B^1, C^0, C^1, D^0, D^1 \rangle$ where B is a set, C^0 and C^1 are functions from Γ to finite subsets of B, and the others are families of finite subsets of B.

The set B consists of the points of the handle $H=\{U, V_0, V_1, V_2\}$ and of the points of the broomcorn $J=\{X_j^a: j\leqq 8,\ a\in\Gamma\}$, see Fig. 4. The elements of $B^0 \cup B^1$ are the *winning* sets, B^0 consists of the subsets $\{X_0^a, X_1^a\}$, $\{X_1^a, X_2^a, X_3^a\}$, $\{X_1^a, X_3^a, X_4^a, X_5^a\}$, $\{X_1^a, X_3^a, X_5^a, X_6^a, X_7^a\}$ and $\{X_1^a, X_3^a, X_5^a, X_7^a, X_8^a, U\}$ for all $a\in\Gamma$, and B^1 contains all two-element subsets of H and all 8-element subsets of J. The values of functions C^0 and C^1 are the *choosing* sets $C^0(a)=\{X_1^a, X_3^a, X_5^a, X_7^a, U\}$ and $C^1(a)=\{X_0^a, X_2^a, X_4^a, X_6^a, X_8^a\}$ for all $a\in\Gamma$. Finally, the elements of D^0 and D^1 are the *validating* sets, D^0 contains the two-element subsets $\{U, V_0\}$, $\{U, V_1\}$, $\{U, V_2\}$, and D^1 contains the one-element subsets $\{V_0\}$, $\{V_1\}$, $\{V_2\}$. ∎

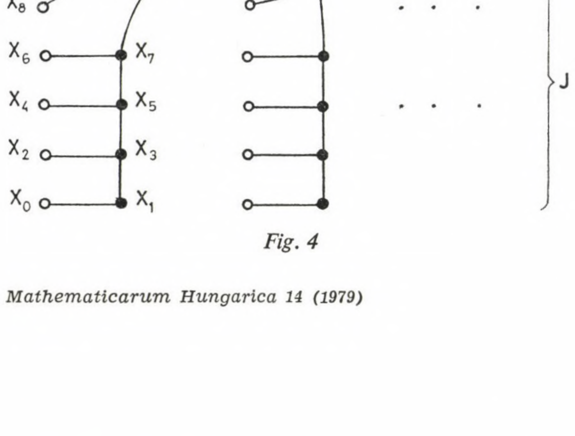


Fig. 4

LEMMA 3.3. *Let $\langle B, B^0, B^1, C^0, C^1, D^0, D^1\rangle$ be the broom associated with Γ. The $\mathfrak{B}$-game $[B, B^0, B^1]^\omega$ has the following property. Either one of the players can win, or they cover the choosing sets $C^0(a)$, $C^1(a)$ for exactly one $a\in\Gamma$ and W covers a validating set from D^0 in the first 11 steps. In the latter case B has a two-step threat playing which he can cover a validating set from D^1. Except for this, no player has a realizable two-step threat.*

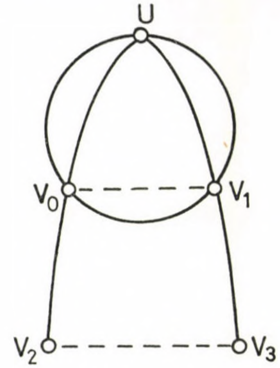


Fig. 5

PROOF. B threatens a two-step victory in the handle so W always has to counterthreaten. Therefore W must pick X_1^a, X_3^a, X_5^a, X_7^a, U and one of V_i's in succession for some $a\in\Gamma$, because W cannot counterthreaten more than 8 times. Then B picks a free element of the handle which forces W to pick the remaining point.

After these steps B has no more two-step threat whilst W has a lot. But W cannot cover more choosing set and cannot threaten more than three times because then W loses the game. ∎

A simplified version of brooms for the case $\Gamma=\{0, 1\}$ is the brush as follows.

DEFINITION 3.4. A *brush* is the 7-tuple $\langle B, B^0, B^1, C^0, C^1, D^0, D^1\rangle$. The set B consists of 5 points, U and V_j $(j\leqq 3)$, see Fig. 5. The family B^0 consists of $\{U, V_0, V_1\}$, $\{U, V_0, V_2\}$, and $\{U, V_1, V_3\}$. The family B^1 contains $\{U\}$, $\{V_0, V_1\}$, and $\{V_2, V_3\}$. The functions C^0 and C^1 are defined as $\mathrm{Dom}\,(C^0)=\mathrm{Dom}\,(C^1)=\{0, 1\}$ and

$$C^0(0) = C^1(1) = \{V_0, V_3\} \qquad C^0(1) = C^1(0) = \{V_1, V_2\}.$$

Finally, $D^0=D^1$ and they contain the subsets $\{V_0, V_3\}$ and $\{V_1, V_2\}$. ∎

The properties of a brush are the same as of a broom, furthermore

LEMMA 3.5. *After playing off a brush, all the points are occupied and no player has any threat.* ∎

We interpret playing off a broom as W's choosing exactly one $a\in\Gamma$, and playing off a brush as B's choosing either 0 or 1.

Brooms (and brushes as well) can be fitted together to form an ω-long sequence. The winning sets of the i-th member of the sequence are validated by the corresponding validating sets of the $(i-1)$st member. So the players are forced to play off these brooms in sequence, exchanging the role of W and B at every new broom. A bit more formally,

DEFINITION 3.6. Let $\langle B_i, B_i^0, B_i^1, C_i^0, C_i^1, D_i^0, D_i^1\rangle$ be brooms (brushes) associated with $\Gamma_i(=\{0, 1\})$ for $i\in\omega$ and let $D_{-1}^0=D_{-1}^1=\emptyset$. The ω-broom (ω-brush) is the triplet $\langle B, B^0, B^1\rangle$ such that for $k=0, 1$

$$B = \bigcup\{B_i\colon\ i\in\omega\}$$

$$B^k = \{X\cup Y\colon\ X\in D_{2i}^k \text{ and } Y\in B_{2i+1}^{1-k} \text{ or } X\in D_{2i-1}^{1-k} \text{ and } Y\in B_{2i}^k,\ i\in\omega\}.$$ ∎

After these preliminaries we turn our attention to the proof of Theorem 3.1. We start by

LEMMA 3.7. *Every snub-game is equivalent to some finite* $\mathfrak{B}$*-game.*

PROOF. Let the positional game G be given by the sets A_i $(i\in\omega)$ and by the winning sets F_W and F_B. Let $\langle B, B^0, B^1\rangle$ be the ω-broom built from the brooms $\mathcal{B}_i=\langle B_i, B_i^0, B_i^1, C_i^0, C_i^1, D_i^0, D_i^1\rangle$ associated with the sets A_i. We are going to define a $\mathfrak{B}$-game G^* which is equivalent to the game snub-G. The board of G^* is B. For every finite sequence $s=\langle a_0, a_1, \dots, a_n\rangle$ such that $a_i\in A_i$ let

$$S_s^0 = C_0^0(a_0)\cup C_1^1(a_1)\cup\dots\cup C_n^k(a_n)$$

$$S_s^1 = C_0^1(a_0)\cup C_1^0(a_1)\cup\dots\cup C_n^{1-k}(a_n)$$

with $k=0$ if n is even and $k=1$ if n is odd. Now let

$$F^0 = B^0\cup\{S_s^0:\ s\in F_W\}$$

$$F^1 = B^1\cup\{S_s^1:\ s\in F_B\}.$$

We claim that the games snub-G and $G^*=[B, F^0, F^1]^\omega$ are equivalent. (i) of the definition is evidently true. (ii) follows from the fact that the players in G^* are forced to simulate the game G. I's choosing to be W in G corresponds to I's choosing the winning sets F^0 in G^* (i.e., being W in G^*, too). In G^* W is forced to start to play off the broom $\mathcal{B}_0$ (otherwise he loses the game). So first the players play off the broom $\mathcal{B}_0$ and player W covers some $C_0^0(a_0)$ with $a_0\in A_0$, i.e., W chooses this a_0. Moreover they validate the winning sets of the broom $\mathcal{B}_1$ by some validators in D_0^0 and D_0^1 so next they have to play off $\mathcal{B}_1$. Here player B covers some $C_1^0(a_1)$, i.e., B chooses $a_1\in A_1$, and so on. If any player does not follow this simulation, he loses.

Last, the winning sets of the form S_s^k ensure in G^* the victory to that player who wins in the snub-G game. ∎

LEMMA 3.8. *Every* $\mathfrak{B}$*-game of limit length is equivalent to some* $\mathfrak{A}$*-game of the same length. If* β *is a limit ordinal then* β*-bounded strategies are preserved.*

PROOF. Let the $\mathfrak{B}$-game be $G=[B_0, F_0^0, F_0^1]^\alpha$ and let $[B_1, F_1^0, F_1^1]^\alpha$ be a new instance of G. Our aim is to construct an $\mathfrak{A}$-game $G^*=[A, F]^\alpha$ in which the players are forced to simulate G. B_0 and B_1 will be subsets of A. The elements of F_k^0 (F_k^1) amplified with some validating sets are among the winning sets of G^*. These validating sets are subsets of the finite set $A-B_0-B_1$. We demand these sets to have the following property. Both players must be able either to win or to cover some validating set within finitely many moves. Moreover, if a player validates (the elements of) F_k^0, the other must not be the first occupying any element of B_k. If he validates F_k^1, the other must have no more than one occupied element in B_k.

This property ensures the portability of WS and DS from G^* to G and back. We describe here a structure satisfying this. It consists of 10 points, U_k and $V_{i,j,k}$ for $i, j, k=0, 1$. There are two F_k^0 validating sets (Fig. 6)

$$D_k^0 = \{U_0, V_{k,0,0}, V_{k,1,0}, V_{1-k,0,1}\}$$

and

$$D_k^1 = \{U_0, V_{k,0,0}, V_{k,1,0}, V_{1-k,1,1}\}$$

and there is one F_k^1 validating set

$$D_k^2 = \{U_1, V_{k,0,1}, V_{k,1,1}\}$$

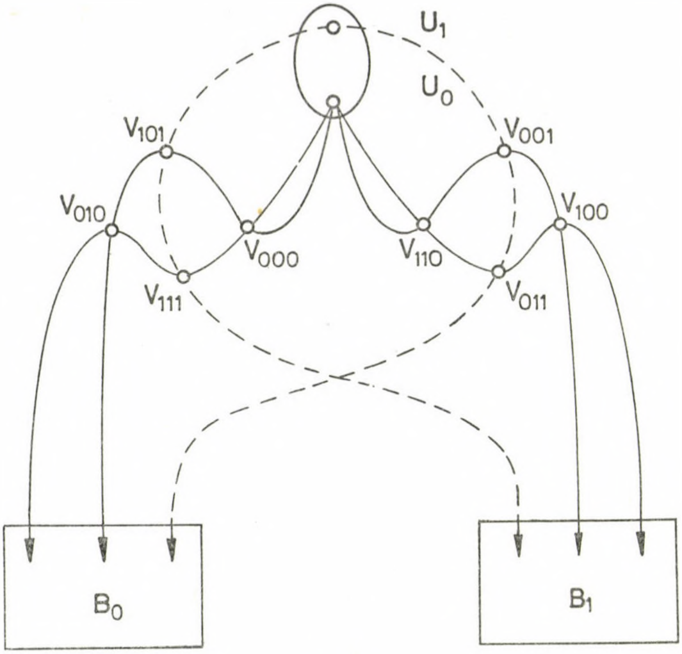


Fig. 6

for $k=0, 1$. The only two-element winning set in F is $\{U_0, U_1\}$. If I's first move is U_0 we interpret it as I's choice $k=0$ in the game G. Otherwise II's first move is U_0, therefore we may assume that in the first two moves they occupy U_0 and U_1. Other winning sets in F are

$$\{U_0, V_{i,j,0}, V_{i,j,1}\} \quad \text{for} \quad i, j = 0, 1$$

and those among the sets

$$\{U_1, V_{0,0,k_1}, V_{0,1,k_2}, V_{1,0,k_3}, V_{1,1,k_4}\}$$

which do not contain D_0^2 or D_1^2.

Let $A=\{U_k, V_{i,j,k}\}\cup B_0\cup B_1$ and let F consist of the 14 winning sets defined above and of the sets $D_k^0\cup X$, $D_k^1\cup X$, $D_k^2\cup Y$ with $X\in F_k^0$ and $Y\in F_k^1$ for $k=0, 1$.

It is easy to check that for $G^*=[A, F]^\alpha$ the (ii) of the definition of the equivalence holds and (i) is trivial. ∎

This proves Theorem 3.1, too.

The following lemma is about the converse of 3.8.

LEMMA 3.9. *Suppose that there is a recursively definable choice function on the board of the $\mathfrak{A}$-game $G=[A, F]^\alpha$. (This is the case if A is an ordinal number.) Then G is equivalent to the $\mathfrak{B}$-game $G^*=[A, F, F]^\alpha$. Moreover, β-bounded strategies are preserved for every ordinal β.*

PROOF. Both I and II can carry strategies recursively by the method of "imagined elements" as was described in 2.1 because there is a recursive choice function on A which gives imagined elements whenever they are necessary. ∎

We mention here a rather surprising consequence of these lemmas.

COROLLARY 3.10. *There is a finite* $\mathfrak{A}$*-game* $[\omega, F]^\omega$ *such that*

(i) *every element of* F *has fewer than* 100 *elements;*

(ii) F *is recursive, i.e., there is a recursive procedure which decides whether a given finite subset of* ω *is an element of* F *or not;*

(iii) *every play ends before the* 100*th move, no matter how the players play;*

(iv) *(as a consequence of* (iii)*) player* I *has WS, but*

(v) I *has no recursively enumerable WS.*

PROOF. We define a positional game G as follows. Let Φ be the set of all formulas of the ZF set-theory and let $\Psi\subset\Phi$ be the ZFC-provable formulas. Obviously, both Φ and Ψ are countable and, by Gödel's theorem, Ψ is not recursive. The game starts by W's saying 1 (if he does not say 1, he loses). Then B says a formula $\varphi\in\Phi$, W says a proof from ZFC, and finally B says a proof from ZFC, too. W wins if either he has proved φ or neither he nor B proved φ. Otherwise the winner is B.

Obviously, W has WS in G because if a formula is provable, he can prove it. On the other side I has no recursively enumerable WS. Supposing it were so, there would be a recursive function which assigns a proof to every provable formula, i.e., Φ would be recursive, a contradiction.

Because W has WS in G, W has WS but has no recursively enumerable WS in snub-G. Making the transformations as was described in 3.8 and 3.7 we get the desired game. ∎

4. Infinite games

In this section we study the determinacy and boundedness of $\mathfrak{A}$-games $[A, F]^\alpha$ with limit $\alpha>\omega$. If $|A|<\omega$ or $|F|<\omega$ then the game is equivalent to $[A, F]^n$ for some $n<\omega$ so we can assume $|A|\geqq\omega$ and $|F|\geqq\omega$. In these latter cases, however, the study of $\mathfrak{B}$-games gives all the information as was shown in Lemmas 3.8 and 3.9 because the cardinality of the boards and that of the families of winning sets are preserved.

First we recall here the basic properties of ω-brushes (see Definition 3.6).

LEMMA 4.1. *Suppose that the* $\mathfrak{B}$*-game* $[A, F^0, F^1]^\alpha$ *with* $\alpha\geqq\omega$ *contains the* ω*-brush* $\langle B, B^0, B^1\rangle$ *(i.e.,* $B\subset A$, $B^0\subset F^0$, $B^1\subset F^1$*) and there is no two-step threat but in the* ω*-brush. Then the players are forced to play off the elements of the* ω*-brush in the first* ω *moves so that*

(i) *after these moves every point of the* ω*-brush is occupied;*

(ii) *the choosing sets* $C_i^0(d_i)$, $C_i^1(d_i)$ $(d_i\in\{0, 1\})$ *are covered in succession;*

(iii) *the value of the digit* d_i *is chosen by* B *if* i *is even and by* W *if* i *is odd in possession of full information about the previous digits and no information about the rest of the digits.*

Let $s=\langle d_0, d_1, \ldots, d_n\rangle$ *be any finite* 0—1 *sequence (i.e.,* $s\in {}^{\underset{\smile}{\omega}}2$*) and let*

$$(s)_W=\bigcup\{C_i^0(d_i)\colon\ i\leqq n\}\qquad (s)_B=\bigcup\{C_i^1(d_i)\colon\ i\leqq n\}.$$

In the ω-brush either one of the players may win before the ω-th step or there is exactly one $\sigma\in{}^{\omega}2$ such that every $(\sigma|n)_W$ is covered by W and every $(\sigma|n)_B$ is covered by B. ▮

This lemma says that an ω-brush forces the players to play an infinite 0--1 game [4].

THEOREM 4.2. *There is a non-determined $\mathfrak{A}$-game the board of which has cardinality at most 2^{ω} such that some player may win before the $(\omega+20)$th step.*

PROOF. By the remarks at the beginning of this section, it is enough to give a $\mathfrak{B}$-game with these properties. The non-determinacy means, by Proposition 2.2, that I has no DS, i.e., if either player W or player B plays by a strategy, he loses.

Now it is well-known that there are non-determined 0—1 games [4], let the family of sequences $C\subset{}^{\omega}2$ witness it. Of course, the cardinality of C is $\leqq 2^{\omega}$. We build a $\mathfrak{B}$-game G as follows. We start with an ω-brush. If the players in the first ω moves encode a sequence $\sigma\in C$ then player W may win within 9 moves (no matter whether I or II is acting as W) and if they encode a sequence $\sigma\notin C$ then player B may win. This property ensures the non-determinacy of G, otherwise some player would be able to win the 0—1 game with C by a strategy.

We call the reader's attention to the problem of the ω-th move. It can be taken by either W or B and we may assume the worst, i.e., it is the turn of the one who is going to lose.

Summarizing, let $\langle B, B^0, B^1\rangle$ be an ω-brush, $(s)_W$ and $(s)_B$ for $s\in\underset{\smile}{\omega}2$ be the subsets of B as defined in 4.1. For every $\sigma\in C$ we take two instances of a game similar to the second example in Proposition 2.5 with boards A^0_σ and A^1_σ. Let the elements of A^l_σ be R and P^i_j $(i\in\omega,\ j\leqq 7)$ and let the elements of the family $F^{0,l}_\sigma$ be $\{R, P^i_0, P^i_1\}\cup(\sigma|i)_W$, $\{R, P^i_0, P^i_2, P^i_3\}\cup(\sigma|i)_W$, $\{R, P^i_0, P^i_2, P^i_4, P^i_5\}\cup(\sigma|i)_W$ and $\{R, P^i_0, P^i_2, P^i_4, P^i_6\}\cup(\sigma|i)_W$ for all $i\in\omega$. (The branches of the i-th tree over the sequence σ are validated by the i-th cut of σ.) Let moreover the elements of the family $F^{1,l}_\sigma$ be $\{P^i_1, P^i_3, P^i_5, P^j_7\}$ for all $i\leqq j\in\omega$. Let $A=\bigcup\{A^l_\sigma\colon \sigma\in C,\ l=0, 1\}$.

The board of G is $A\cup B$, the family of W-winning sets is

$$F^0 = B^0\cup\bigcup\{F^{0,l}_\sigma\colon \sigma\in C,\ l=0, 1\}$$

and the family of B-winning sets is

$$F^1 = B^1\cup\bigcup\{F^{1,l}_\sigma\colon \sigma\in C,\ l=0, 1\}\cup\{\text{six-element subsets of } A\}.$$

We leave to the reader to check that the $\mathfrak{B}$-game $[A\cup B, F^0, F^1]^{\omega+20}$ has all the described properties. ▮

This construction cannot give a nondetermined $\mathfrak{A}$-game with countable board because every 0—1 game with countable winning set is determined. The following theorems deal with the case of countable boards.

THEOREM 4.3. *If there is a Ramsey cardinal then every $\mathfrak{A}$-game $[\omega, F]^{\omega\cdot 2}$ is determined.*

PROOF. The existence of a Ramsey cardinal implies that every Σ^1_1 game is determined [7]. It is easy to check that the family of ω-long plays after which I can win within finitely many moves forms a Σ^1_1 set. By the assumption this set is determined, i.e., either I has a strategy to remain in it, in which case I has a WS, or II has a strategy to avoid it, which means a DS for II. ∎

The axiom of constructibility $V=L$ implies the existence of a non-determined Σ_1^1 game [9]. The existence of a non-determined Σ_1^1 game, however, seems not to imply the existence of a non-determined $[\omega, F]^{\omega\cdot 2}$ game. But the construction can be carried over.

We proceed with two lemmas.

LEMMA 4.4. *Let W and B play two instances G and $\bar{G}$ of an unsymmetric game* [2] *as follows. First W chooses a finite* 0—1 *sequence $s_0 \in {}^{\underset{\smile}{\omega}}2$ then B chooses a digit $t_0 \in \{0, 1\}$ and a finite* 0—1 *sequence $\bar{s}_0$. W responds by choosing a digit $\bar{t}_0 \in \{0, 1\}$ and a finite* 0—1 *sequence s_1, etc. Thus the players form two infinite* 0—1 *sequences, $\sigma = s_0{}^\frown t_0{}^\frown s_1{}^\frown t_1{}^\frown \dots$ and $\bar{\sigma} = \bar{s}_0{}^\frown \bar{t}_0{}^\frown \bar{s}_1{}^\frown \bar{t}_1{}^\frown \dots$ where ${}^\frown$ denotes concatenation.*

Let $C \subset {}^\omega 2$ be uncountable and containing no perfect subset. If W plays by a strategy, then B has a counterplay such that $\sigma \notin C$ and $\bar{\sigma} \in C$. If B plays by a strategy then W has a counterplay such that $\sigma \in C$ and $\bar{\sigma} \notin C$.

PROOF. Assume that W plays by a strategy S, the other case is similar. The moves of W in $\bar{G}$ depend not only on the moves of B in $\bar{G}$ but on the moves in G, too. In view of this, we define countable sets X_n, Y_n for $n \in \omega$ by induction which satisfy the following conditions. The elements of X_n are sequences of the form

$$\langle t_0, \bar{s}_0, \bar{t}_1, t_1, \bar{s}_1, \bar{t}_2, \dots, t_{n-1}, \bar{s}_{n-1}, \bar{t}_n \rangle$$

where $t_i, \bar{t}_{i+1} \in \{0, 1\}$ and $\bar{s}_i \in {}^{\underset{\smile}{\omega}}2$ for $i<n$ such that $\bar{t}_{i+1}$ is the response of W by S to the sequence of moves $t_0, \bar{s}_0, t_1, \bar{s}_1, \dots, t_i, \bar{s}_i$ of B. Y_n is a subset of ${}^\omega 2$ and for every $\tau \in {}^\omega 2 - Y_n$ and for every $t_i \in \{0, 1\}$ $(i<n)$ there is exactly one sequence

$$\langle t_0, \bar{s}_0, \bar{t}_1, \dots, t_{n-1}, \bar{s}_{n-1}, \bar{t}_n \rangle \in X_n$$

with the given digits t_i such that $\bar{s}_0{}^\frown \bar{t}_1{}^\frown \dots {}^\frown \bar{s}_{n-1}{}^\frown \bar{t}_n \prec \tau$.

Let $X_0 = \{\langle\rangle\}$, $Y_0 = \emptyset$ and suppose that X_n, Y_n are defined and have the described properties. Then for every

$$s = \langle t_0, \bar{s}_0, \bar{t}_1, \dots, t_{n-1}, \bar{s}_{n-1}, \bar{t}_n \rangle \in X_n$$

and $t_n \in \{0, 1\}$ let X_{n+1} contain the sequences

$$s^\frown t_n^\frown \langle\rangle^\frown \bar{t}_{n+1}^0$$
$$s^\frown t_n^\frown \langle 1-\bar{t}_{n+1}^0 \rangle^\frown \bar{t}_{n+1}^1$$
$$s^\frown t_n^\frown \langle 1-\bar{t}_{n+1}^0, 1-\bar{t}_{n+1}^1 \rangle^\frown \bar{t}_{n+1}^2$$
$$\vdots$$

where $\bar{t}_{n+1}^0, \bar{t}_{n+1}^1, \dots$ are the corresponding responses of W by the strategy S. Let Y_{n+1} contain the elements of Y_n and the infinite sequences

$$\bar{s}_0^\frown \bar{t}_1^\frown \dots {}^\frown \bar{s}_{n-1}{}^\frown \bar{t}_n{}^\frown \langle 1-\bar{t}_{n+1}^0, 1-\bar{t}_{n+1}^1, \dots \rangle$$

with the above notations for every $s \in X_n$, $t_n \in \{0, 1\}$.

Obviously, X_{n+1} and Y_{n+1} are countable if X_n and Y_n were, and they have the described properties. Therefore $\bigcup\{Y_n: n\in\omega\}$ is countable, C is not, so there is a $\tau\in C$ for which $\tau\notin Y_n$ for every $n\in\omega$.

Because X_{n+1} is an end-extension of X_n, we can define a "strategy" $\bar{S}$ of B as follows. Given $t_i\in\{0, 1\}$ for $i<n$, let $\bar{S}(t_0, \dots, t_{n-1})$ be the only sequence for which

$$\langle t_0, \bar{S}(t_0), \bar{t}_1, t_1, \bar{S}(t_0, t_1), \bar{t}_2, \dots, t_{n-1}, \bar{S}(t_0, \dots, t_{n-1}), \bar{t}_n\rangle\in X_n$$

and

$$\bar{S}(t_0)^\frown \bar{t}_1^\frown \dots {}^\frown\bar{S}(t_0, \dots, t_{n-1})^\frown \bar{t}_n \prec \tau.$$

Playing by this "strategy" $\bar{S}$ in $\bar{G}$, B forces $\bar{\sigma}\in C$ independently of his moves in G. What is more, B forces to form always the same sequence $\bar{\sigma}=\tau$.

Now we turn our attention to G. Every move t_i of B in G determines uniquely $\bar{S}$ in $\bar{G}$ (via $\bar{S}$) so it determines uniquely W's response s_{i+1} in G. This means that we may forget about $\bar{G}$ totally and the strategy S reduces to a strategy S' for W in G only. Now the outcomes of plays according to S' form a perfect subset of ${}^\omega 2$. C contains no perfect subset so B can choose a counterplay which lead out of C. Combining this counterplay with $\bar{S}$ we are done. ∎

LEMMA 4.5. *Assume $V=L$. There are finite trees T_s^0 and T_s^1 (trees in the set-theoretical sense, see* [3]*) for every $s\in{}^{\underset{\smile}{\omega}}2$ such that* (i) *and* (ii) *below are satisfied.*

(i) *If $s, t\in{}^{\underset{\smile}{\omega}}2$ and $s\prec t$ then T_t^l is an end-extension of T_s^l.*

By (i), *$T_\sigma^l=\bigcup\{T_s^l: s\prec\sigma\}$ is a tree of height $\leqq\omega$ for every $\sigma\in{}^\omega 2$.*

(ii) *Let W and B play an infinite* 0—1 *game and let $\sigma\in{}^\omega 2$ denote the resulting sequence. If W plays by a strategy then B has a counterplay such that either W picks only* 1 *and B picks only* 0 *after finitely many moves, or T_σ^0 is well-founded and T_σ^1 is not. If B plays by a strategy then W has a counterplay such that either B picks only* 1 *and W picks only* 0 *after finitely many moves, or T_σ^1 is well-founded and T_σ^0 is not.*

PROOF. $V=L$ implies that there is a Π_1^1 subset $C\subset{}^\omega 2$ of cardinality 2^ω without a perfect subset [9]. Given any Π_1^1 subset $C\subset{}^\omega 2$ one can assign finite trees T_s to every $s\in{}^{\underset{\smile}{\omega}}2$ such that $\bigcup\{T_s: s\prec\sigma\}$ is well-founded if and only if $\sigma\in C$, see [10]. By these facts and by the previous lemma, we are done if we can code the twofold unsymmetric game G of Lemma 4.4 in a single 0—1 game G^*.

But this latter task is easy. Enumerate all finite 0—1 sequences. Say W (or B) wants to choose in G some finite sequence s. Suppose s is the n-th in the enumeration then W (or B) chooses n consecutive 1's followed by a 0 in G^*. The single moves in G correspond to single moves in G^*. Clearly, every position of G^* determines uniquely the status of the simulation and every play in G^* corresponds to some play in G except for those where W or B picks only 1 after finitely many turns. ∎

THEOREM 4.6. *Assume $V=L$. There is a non-determined $\mathfrak{A}$-game $[\omega, F]^{\omega\cdot 2}$.*

PROOF. As in the proof of Theorem 4.2, we shall make a $\mathfrak{B}$-game G with these properties. We start with an ω-brush $\langle B, B^0, B^1\rangle$ and let $(s)_W$ and $(s)_B$ for $s\in{}^{\underset{\smile}{\omega}}2$ be the subsets of B as defined in 4.1. We define the remaining part of G with the help of the previous lemma. There are two essentially different cases.

Case A. The players in the first ω moves encode an exceptional sequence τ. For $k=0, 1$ let $C^k\subset{}^\omega 2$ be defined by

$$\langle d_0, d_1, \dots\rangle\in C^k \quad \text{iff} \quad d_{2i}=k, \quad d_{2i+1}=1-k \quad \text{for every} \quad i\geqq i_0.$$

Obviously, $\tau \in C^0 \cup C^1$, $C^0 \cup C^1$ is countable, and $\tau \in C^0$ if B plays by a strategy, $\tau \in C^1$ if W plays by a strategy. Let A^l_σ, $F^{0,l}_\sigma$, $F^{1,l}_\sigma$ ($l=0, 1$) be as in the proof of 4.2, let

$$A^k_E = \bigcup \{A^l_\sigma :\ \sigma \in C^k,\ l = 0, 1\} \quad (k = 0, 1)$$

be the boards for the exceptional cases, and let

$$E^k = \bigcup \{F^{1-k,l}_\sigma :\ \sigma \in C^k,\ l = 0, 1\} \cup \bigcup \{F^{k,l}_\sigma :\ \sigma \in C^{1-k},\ l = 0, 1\} \cup$$
$$\cup \{21\text{-element subsets of } A^{1-k}_E\}$$

be the W winning sets for $k=0$, and the B winning sets for $k=1$. The boards are countable, and this part of the game has the following properties.

Suppose W plays by a strategy (the other case is quite similar). If they encode a sequence $\tau \in C^1$ then B can win at his fifth move in A^1_E. Moreover, B has a three-step threat in A^1_E so either W loses or W has to threaten to win within three steps. (Remember, the ω-th step is W's turn.) Well, W cannot threaten in A^1_E and cannot threaten in the remaining parts of the game. So W must play in A^0_E. But B can fend off every threat in A^0_E and after 21 pairs of moves B wins eventually.

If the encoded sequence $\tau \notin C^0 \cup C^1$ (observe, $\tau \in C^0$ cannot occur if B plays properly) then W's moves in A^0_E, as before, do not count and there can be at most 21 of them. The W's moves in A^1_E, however, cause a little problem. Here W wins after his 21-st move but he gives B ten free moves.

Indeed, W can threaten at his every other move only, and B can fend these threats off by one move. So B must be able to win at the other parts of the game with 10 free, but not necessarily consecutive moves.

Case B. Otherwise. Let T^l_s be the trees for $s \in {}^{\underline{\omega}}2$, $l=0, 1$ the existence of which was shown in Lemma 4.5. Let T^0 and T^1 be disjoint trees of height ω such that every node has countably many immediate successors. If v is a node in T^l then $p(v)$ denotes its immediate predecessor and $h(v) \in \omega$ denotes its height: if v is the root then $h(v)=0$, otherwise $h(v)=h(p(v))+1$.

We may assume that T^l_s are embedded in T^l so that T^l is an end-extension of T^l_s and $T^l_s \cap T^l_t = T_r$ where r is the longest common initial segment of s and t. Therefore for each node v of T^l we can define a sequence $s^l_v \in {}^{\underline{\omega}}2$ such that $v \in T^l_s$ if and only if $s^l_v = s$ or $s^l_v \prec s$.

Replace each node v in T^l by a broom associated with the set of immediate successors of v. These brooms will be fitted together in such a way that the players must climb on branches of the trees (Fig. 7). First B has a threat in T^0 and W can fend this off only by occupying a "validated" edge starting from the root of T^0. Doing so W has a threat in T^1 and B fends it off occupying a "validated" edge of T^1, etc. The validatings are done at the first ω steps. If they encode in the ω-brush the sequence $\sigma \in {}^\omega 2$ then the valid edges are just that of the subtrees $T^l_\sigma = \bigcup \{T^l_s :\ s \prec \sigma\}$. Now it is clear that if T^0_σ is well-founded and T^1_σ is not then W cannot fend all the threats of B off because W "runs out" of his tree eventually. It means that B wins within finitely many moves. Similarly, if T^1_σ is well-founded and T^0_σ is not, then W can win within finitely many moves. The exact definition of this part goes as follows.

Let T^l be the trees and s^l_v the sequences as discussed above. Let $\mathscr{B}_{l,v} = \langle B_{l,v}, B^0_{l,v}, B^1_{l,v}, C^0_{l,v}, C^1_{l,v}, D^0_{l,v}, D^1_{l,v} \rangle$ be the broom associated with the immediate

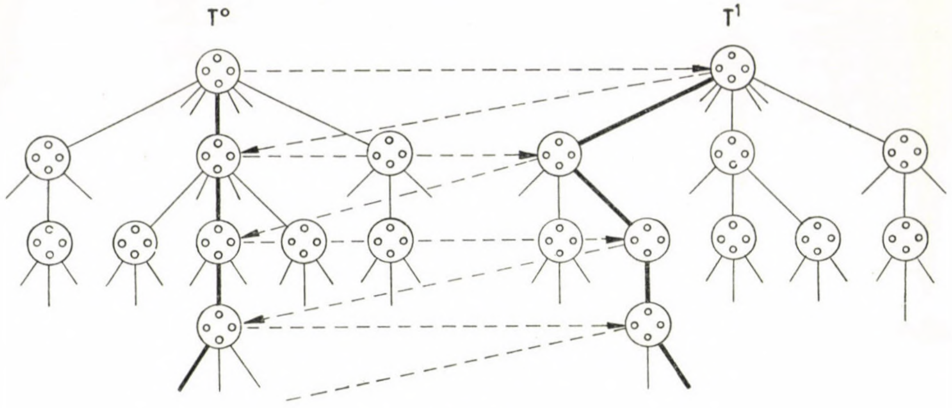


Fig. 7

successors of the node v of T^l. Let

$$A_M = \bigcup \{B_{l,v}\colon \ v \text{ is a node of } T^l,\ l = 0, 1\}$$

the board for the main case. We recall that for any node v, $p(v)$ denotes its only immediate predecessor and $h(v)$ denotes its height. Let, by definition, $C^k_{l,p(v)}(v)=\emptyset$ if v is the root of T (i.e., if $p(v)$ does not exist) and let

$$H^k_{l,v} = \begin{cases} \{\emptyset\} \text{ if } l=1 \text{ and } v \text{ is the root of } T^0 \\ \bigcup \{D^k_{l,w}\colon \ w \text{ is a node of } T^l \text{ and } h(w)=h(v)-l\} \text{ otherwise.} \end{cases}$$

We define the families $M^k_{l,v}$ for every node v of T^l and $k, l=0, 1$ as follows.

$$M^0_{0,v} = \{X\cup Y\cup C^0_{0,p(v)}(v)\cup (s^0_v)_W\colon \ X\in B^0_{0,v},\ Y\in H^1_{1,v}\}$$

$$M^0_{1,v} = \{X\cup Y\cup C^1_{1,p(v)}(v)\colon \ X\in B^1_{1,v},\ Y\in H^0_{0,v}\}$$

$$M^1_{0,v} = \{X\cup Y\cup C^1_{0,p(v)}(v)\colon \ X\in B^1_{0,v},\ Y\in H^0_{1,v}\}$$

$$M^1_{1,v} = \{X\cup Y\cup C^0_{1,p(v)}(v)\cup (s^1_v)_B\colon \ X\in B^0_{1,v},\ Y\in H^1_{0,v}\}.$$

Finally, let

$$M^k = \bigcup \{M^k_{l,v}\colon \ v \text{ is a node of } T^l,\ l = 0, 1\}$$

be the W winning sets for $k=0$ and the B winning sets for $k=1$.

The main part $\langle A_M, M^0, M^1\rangle$ has the following properties. The board, A_M is countable. We assume that always W picks the first element of A_M. At any moment, both W and B have one or two-step threats so any free move means victory immediately or at the next move. If W plays by strategy then B may win within finitely many moves, and if B plays by a strategy then W may win within finitely many moves (but only if W starts picking the elements of A_M).

This main part cannot be put over an ω-brush because there are two-step threats in it. By Case A, it must not contain even a four-step threat. And, what is the worst, it cannot be assumed that W starts picking the points of A_M. To solve these problems, we double the main part and put a little "prelude" before them as it was done in Lemma 3.8. Doing so we increase the size of the winning sets and assure that W starts picking.

Let $\langle A_M, M^0, M^1\rangle$ and $\langle A_N, N^0, N^1\rangle$ be two disjoint instances of the main part, U_M and U_N be disjoint 8-element sets. The family of i-element subsets of U is denoted by $[U]^i$. Let

$$F = A_M \cup A_N \cup U_M \cup U_N$$

be the board, and let

$$F^0 = [U_N]^5 \cup [U_M]^5 \cup \{X \cup Y\colon\ X \in M^0 \text{ and } Y \in [U_M]^4 \text{ or } X \in N^0 \text{ and } Y \in [U_N]^4\}$$

$$F^1 = \{X \cup Y\colon\ X \in [U_M]^4,\ Y \in [U_N]^4\} \cup$$
$$\cup \{X \cup Y\colon\ X \in M^1 \text{ and } Y \in [U_M]^4 \text{ or } X \in N^1 \text{ and } Y \in [U_N]^4\}$$

be the W winning and B winning sets. Every winning set in $F^0 \cup F^1$ has at least five elements. Now suppose W plays by a strategy, i.e., B is going to win. If W picks an element of $A_M \cup U_M$ $(A_N \cup U_N)$ then B picks simply an unoccupied element of U_M (U_N). Then B either wins or validates just one of the parts A_M or A_N. Clearly, 8 free moves mean a victory for B during this prelude and 4 free moves do the same during the continuation.

At last if B plays by a strategy then the ω-th move belongs to B. If the first move of B in F is an element of $A_M \cup U_M$ $(A_N \cup U_N)$ then W starts picking elements of U_N (U_M). W either can pick 5 elements of this (which means a victory because B has no 5-element winning set here), or B picks elements of U_N only, i.e., W can start playing in A_N.

Summarizing, the $\mathfrak{B}$-game $[B \cup A_E^0 \cup A_E^1 \cup F,\ B^0 \cup E^0 \cup F^0,\ B^1 \cup E^1 \cup F^1]^{\omega \cdot 2}$ is a non-determined $\mathfrak{B}$-game the existence of which was stated. ∎

In the last section of our paper we deal with the possible lengths of $\mathfrak{A}$-games in which I has WS.

Proposition 4.7. *Let $\alpha = \omega \cdot \beta + n$ $(n \in \omega)$ be an ordinal. Suppose there is an $\mathfrak{A}$-game $[A, F]^\alpha$ in which* I *has WS but has no α-bounded WS. Then the following cases cannot occur:*
(i) *$n > 0$ and n is even;*
(ii) *$\beta \neq 0$ and $n = 1$;*
(iii) *β is a limit ordinal and $n = 3$.*

Proof. I cannot win after II's move so (i) follows. If I wins by occupying an element at limit step then I can occupy that element before, which proves (ii). Now, suppose I wins by his second move after a limit step. He can do it only if he has a "valid" V as shown in Fig. 8. This V became valid previously. Since β is a limit ordinal, I has a limit step before the $(\omega \cdot \beta)$-th when this V is valid, here I can occupy the bottom of the V and win. ∎

Fig. 8

Theorem 4.8. *For every ordinal α not excluded in* 4.7 *there is an $\mathfrak{A}$-game $[A, F]^\alpha$ such that* I *has WS but has no α-bounded WS.*

PROOF. Let ξ be any ordinal and let $\xi=\omega\cdot\beta+n$ $(n\in\omega)$. ξ is even (odd) if n is even (odd). We define the ξ-train $\langle T, T^0, T^1\rangle$ for every odd ordinal ξ. The elements of T are the points P_δ for $\delta<\xi$. T^0 contains the subsets $\{P_\delta, P_{\delta+1}\}$ where $\delta+1<\xi$, δ is even, and if $n\geqq 5$ then the subset $\{P_{\omega\cdot\beta+k}\colon k<n, k \text{ is even}\}$ and if $n=3$ and $\beta=\gamma+1$ then the subset $\{P_{\omega\cdot\gamma}, P_{\omega\cdot\beta}, P_{\omega\cdot\beta+2}\}$. Finally, T^1 contains the subsets $\{P_\gamma, P_\delta\}$ where $\delta\leqq\gamma+1<\xi$, δ is even and γ is odd. The train is led by the 6-point engine E. The points of E are U_i, V_i for $i\leqq 2$, see Fig. 9. Let $U=\{U_0, U_1, U_2\}$ and $V=\{V_0, V_1, V_2\}$.

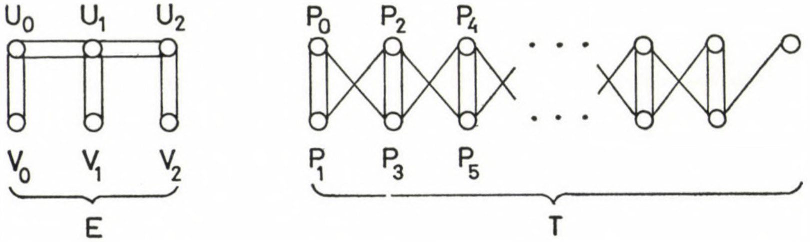


Fig. 9

Now let α be given and β, n be defined by $\alpha=\omega\cdot\beta+n$. We give the $\mathfrak{A}$-games which satisfy the requirements. The easy work of verification is left to the reader. We distinguish three cases.

Case A: α is finite. This case is trivial.

Case B: $\alpha=\omega\cdot\beta+n$ is infinite, n is odd and either $n>3$ or $n=3$ and $\beta=\gamma+1$.

Let E be the engine, $\langle T, T^0, T^1\rangle$ be the α-train. The board of the $\mathfrak{A}$-game is $E\cup T$ and the winning sets are $\{U_0, V_0\}$, $\{U_0, U_1, V_1\}$, $\{U_0, U_1, U_2, V_2\}$, $U\cup X$ for all $X\in T^0$ and $V\cup Y$ for all $Y\in T^1$. I may win picking $U_0, U_1, U_2, P_0, P_2, \ldots$ in succession. If I does not do so, II can either win or make a draw.

Case C: α is a limit. Let $\alpha=\sup\{\alpha_i\colon i\in I\}$, $\alpha_i=\omega\cdot\beta_i+n_i$ such that n_i is odd and $n_i>3$. Let $\langle T_i, T_i^0, T_i^1\rangle$ be the α_i-train, E be the engine and $\langle B, B^0, B^1, C^0, C^1, D^0, D^1\rangle$ be the broom associated with the index set I, and R be an entirely new point. The board of the $\mathfrak{A}$-game is $E\cup B\cup\{R\}\cup\bigcup\{T_i\colon i\in I\}$. The winning sets are $\{U_0, V_0\}$, $\{U_0, U_1, V_1\}$, $\{U_0, U_1, U_2, V_2\}$, $V\cup\{R\}$, the sets $U\cup X$ and $V\cup Y$ for all $X\in B^1$ and $Y\in B^0$, and the sets

$$U\cup C^1(i)\cup X\cup Y \quad \text{for} \quad X\in D^1,\ Y\in T_i^0$$

$$V\cup C^0(i)\cup X\cup Y \quad \text{for} \quad X\in D^0,\ Y\in T_i^1$$

for all $i\in I$.

After playing off the engine, I has to pick the point R, and II may choose any train to play in. Therefore there are plays of length at least α_i, i.e., I has no α-bounded WS. On the other side I can obviously win before the α-th step. ∎

The cardinality of the boards in these constructions is equal to the cardinality of α if α is infinite. This is the smallest possible value if α is not a successor cardinal. If $\alpha=\varkappa^+$ where $\varkappa$ is an infinite cardinal then the lower bound is $\varkappa$. We can construct $\mathfrak{A}$-games $[\varkappa, F]^{\varkappa^+}$ where I has WS but has no $\varkappa^+$-bounded WS, but the construction is too difficult to give here.

REFERENCES

[1] Csirmaz, L.: What do we know about amoeba, *Köz. Mat. Lapok* **57** (1978), 196—201 (in Hungarian).

[2] Davis, M.: Infinite games of perfect information, *Advances in game theory*, Annals of Math. Studies, no. **52**, Princeton University Press, Princeton, N. Y., 1964, 85—101.

[3] Devlin, K. J.: *Aspects of constructibility*, Lecture Notes in Mathematics, Vol. 354, Springer-Verlag, Berlin—New York, 1973, xii+240 pp.

[4] Gale, D. and Stewart, F. M.: Infinite games with perfect information, *Contributions to the theory of games*, Vol. 2, Annals of Math. Studies, no. **28**, Princeton University Press, Princeton, N. Y., 1953, 245—266.

[5] Hales, A. W. and Jewett, R. I.: On regularity and positional games, *Trans. Amer. Math. Soc.* **106** (1963), 222—229.

[6] Keisler, H. J.: Infinite quantifiers and continuous games, *Application of Model Theory to Algebra, Analysis and Probability*, Holt, Rinehart & Winston, New York, 1969, 228—264.

[7] Martin, D. A.: Measurable cardinals and analytic games, *Fund. Math.* **66** (1970), 287—291.

[8] Martin, D. A.: Borel determinacy, *Ann. of Math.* **102** (1975), 363—371.

[9] Mycielski, J.: On the axiom of determinateness, *Fund. Math.* **53** (1964), 205—224.

[10] Schoenfield, J. R.: *Mathematical Logic*, Addison-Wesley Publishing Co., Reading, Mass.—London—Don Mills, Ont., 1967, vii+344 pp.

(Received April 11, 1980)

MTA MATEMATIKAI KUTATÓ INTÉZETE
REÁLTANODA U. 13—15
H—1053 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 227—240

A GENERAL THEOREM ON STRONG MEANS

by

V. TOTIK

§ 1

In the course of the development of strong approximation theory one of the most general theorems appeared very early. Namely, in 1965 L. LEINDLER [2] proved an estimate (see below) which covered many interesting and important cases. The key point was the inequality*

$$\left\{\frac{1}{n}\sum_{k=n+1}^{2n}|s_k-f|^p\right\}^{1/p} \leqq K_p E_n(f), \tag{1.1}$$

where $s_k(x)=s_k(f;x)$ is the k-th partial sum of the Fourier series of the 2π-periodic continuous function f and $E_n(f)$ is the best uniform approximation of f by trigonometric polynomials of order at most n.

In the so called very strong approximation ([6, 7]) we used the following generalization of (1.1):

THEOREM A. *If $p>0$ and $1\leqq k_1<k_2<\ldots<k_r\leqq n$, $f\in C_{2\pi}$, then*

$$\left\{\frac{1}{r}\sum_{i=1}^{r}|s_{k_i}-f|^p\right\}^{1/p} \leqq K_p E_{k_1}(f)\log\frac{2n}{r}, \tag{1.2}$$

where K_p depends only on p.

(1.2) is not an essential generalization of (1.1), thus, it is surprising and remarkable that we cannot go further. The following theorems, which follow very easily from (1.2), cover every non-negative summation method known for us.

In order to formulate our first result we introduce some notations. Let $T=\{t_k\}_{k=0}^{\infty}$ be a non-negative sequence and $\mathcal{N}=\{N_m\}_{m=0}^{\infty}$ $(N_0=0,\ N_1=1)$ a subsequence of the natural numbers. Let

$$T(N_m, N_{m+1}) = \{t_k | N_m < k \leqq N_{m+1}\}$$

and

$$T^*(N_m, N_{m+1}) = \{t_k^* | N_m < k \leqq N_{m+1}\}$$

be the monotone decreasing rearrangement of the finite sequence $T(N_m, N_{m+1})$ finally let

$$\tau(p, N_m, N_{m+1}) = \sum_{k=1}^{N_{m+1}-N_m} t_{N_m+k}^*\left(\log\frac{N_{m+1}}{k}\right)^p \quad (m=1,2,\ldots)$$

$$\tau(p, N_0, N_1) = t_0+t_1.$$

---

* K, c and these with subscripts denote constants not necessarily the same at each occurrence.

THEOREM 1. *For any $f\in C_{2\pi}$ and $p>0$ we have*

$$T(f,p;x)=\left\{\sum_{k=0}^{\infty} t_k|s_k(x)-f(x)|^p\right\}^{1/p} \leqq K_p\left\{\sum_{m=0}^{\infty}\tau(p,N_m,N_{m+1})\big(E_{N_m}(f)\big)^p\right\}^{1/p},$$

where K_p depends only on p.

If $\{\chi(n)\}_{n=1}^{\infty}$ is a monotone decreasing positive sequence with $\chi(2n)\geqq c\chi(n)$, we set

$$W^{\gamma}\chi=\{f|E_n(f)=O(n^{-\gamma}\chi(n))\}\ (\gamma>0),$$

furthermore, in the case $\chi(n)\equiv 1$ we write W^{γ} instead of $W^{\gamma}\chi$.
We obtain from Theorem 1

COROLLARY 1. *If $f\in W^{\gamma}\chi$ then*

$$T(f,p;x)\leqq K\left\{\tau(p,N_0,N_1)+\sum_{m=1}^{\infty}\tau(p,N_m,N_{m+1})\big(N_m^{-\gamma}\chi(N_m)\big)^p\right\}^{1/p}\equiv$$

$$\equiv KT(W^{\gamma}\chi,\mathcal{N},p).$$

Now we prove that this is the best estimate in the most cases.

THEOREM 2. *Let us suppose that $\mathcal{N}=\{N_m\}$ satisfies*

$$N_m\left(1+\frac{1}{c}\right)\leqq N_{m+1}\leqq cN_m\quad (c>0;\ m=1,2,\dots).$$

There exists in every class $W^{\gamma}\chi$ a function f depending only on $\mathcal{N}$ such that for every non-negative sequence $T=\{t_k\}$ all the subsequences $T(N_m,N_{m+1})$ of which are monotonic, we have

$$T(f,p;0)\geqq K_{p,c}T(W^{\gamma}\chi,\mathcal{N},p), \tag{1.3}$$

and here $K_{p,c}$ depends only on p and c.

Let $T=\{T^i\}_{i\in I}$ $(i\to i_0)$ be a summation method such that each $T^i=\{t_k^i\}_{k=0}^{\infty}$ is a non-negative sequence of real numbers (I is an index set). We shall write

$$(T,W^{\gamma}\chi,p,i)\mathrel{\underset{\frown}{\overset{\smile}{\equiv}}}\delta_i$$

if for any $f\in W^{\gamma}\chi$ we have

$$\left\{\sum_{k=0}^{\infty} t_k^i|s_k-f|^p\right\}^{1/p}\leqq K\delta_i\ (K>0,\ i\to i_0)$$

furthermore, for an arbitrary subset $\{i_\nu\}\subseteq I$, $i_\nu\to i_0$ there is an $f\in W^{\gamma}\chi$ with the property

$$\left\{\sum_{k=0}^{\infty} t_k^{i'_\nu}|s_k(0)-f(0)|^p\right\}^{1/p}\geqq c\delta_{i'_\nu}\quad (c>0,\ \nu\to\infty),$$

where $\{i'_\nu\}\subseteq\{i_\nu\}$ is infinite.

This says exactly that $\{\delta_i\}_{i\in I}$ gives the exact strong approximation order of functions belonging to the class $W^\gamma\chi$. We obtain at once from Corollary 1 and Theorem 2

COROLLARY 2. *Let us suppose that there is a sequence* $\mathcal{N}=\{N_m\}$ *with the property*

$$N_m\left(1+\frac{1}{c}\right)\leqq N_{m+1}\leqq cN_m \quad (c>1,\ m=1,2,\dots)$$

such that for every $\{i_\nu\}_{\nu=1}^\infty\subseteqq I$, $i_\nu\to i_0$ *there is a subset* $\{i_{\nu'}\}_{\nu'=1}^\infty\subseteqq\{i_\nu\}$ *for which all of the sequences* $T^{i_{\nu'}}(N_m, N_{m+1})$ $(\nu', m=1,2,\dots)$ *are monotonic. In this case*

$$(T, W^\gamma\chi, p, i)\underset{\frown}{\equiv} T^i(W^\gamma\chi, \mathcal{N}, p).$$

A similar consideration gives that the same is true (e.g. with $N_m=2^{m-1}$) if each sequence T^i increases to a certain index and decreases after it.

Thus, Corollary 2 reduces the problem of the order of a strong mean to the mere computation of $T^i(W^\gamma\chi,\mathcal{N},p)$.

We shall illustrate our results by examining the most frequently used methods. For the sake of simplicity we consider only the classes W^γ, and mention that if $\gamma=r+\alpha$, r integer and $0<\alpha\leqq 1$, then

$$W^\gamma=\{f\,|\,f^{(r)}\in W^\alpha\},$$

where $W^\alpha=\mathrm{Lip}\,\alpha$ $(0<\alpha<1)$ and W^1 is the Zygmund-class.

Applications

I. Let us begin with a wide class of non-negative matrices introduced by L. LEINDLER [2]. Let

$$T_n(f, p, \|\alpha\|; x)=\left\{\frac{1}{A_n}\sum_{k=1}^n \alpha_{nk}|s_k(x)-f(x)|^p\right\}^{1/p},$$

where $A_n=\sum_{k=1}^n\alpha_{nk}$. We assume that there is an index-sequence $\{N_m\}$ $(N_0=0)$ and a $q'>1$ such that for $N_{m_0-1}<n\leqq N_{m_0}$ the inequalities

$$\left\{\sum_{k=N_{m-2}+1}^{N_m}(\alpha_{nk})^{q'}\right\}^{1/q'}\leqq KN_m^{-1/p'}A_n(N_{m-2}, N_m)$$

$$\left(A(N_{m-2}, N_m)=\sum_{k=N_{m-2}+1}^{N_m}\alpha_{nk},\quad n=1,2,\dots,\ m=2,\dots,m_0\right)$$

are satisfied with $p'=\dfrac{q'}{q'-1}$. In this case Theorem 1 gives

$$\frac{1}{A_n(N_{m-2}, N_m)}\sum_{k=N_{m-2}+1}^{N_m}\alpha_{nk}|s_k-f|^p\leqq K\big(E_{N_{m-2}}(f)\big)^p \tag{1.4}$$

and

$$T_n(f, p, \|\alpha\|; x) \leqq K\left\{\frac{1}{A_n}\sum_{m=2}^{m_0} A_n(N_{m-2}, N_m)\big(E_{N_{m-2}}(f)\big)^p\right\}^{1/p}. \tag{1.5}$$

The inequalities (1.4) and (1.5) were the basic ones of L. Leindler [2].

II. *Riesz means.* Let $\lambda=\{\lambda(n)\}_{n=0}^{\infty}$ be a monotonically increasing sequence of positive numbers tending to the infinity, and let

$$R_n^\lambda(f, p; x)=\left\{\frac{1}{\lambda(n)}\sum_{k=0}^{n-1}\big(\lambda(k+1)-\lambda(k)\big)|s_k(x)-f(x)|^p\right\}^{1/p}$$

(see [3]). We may suppose that the function $\lambda(x)$ is defined for all $x\geqq 0$ and that it is continuously differentiable.

1. If λ is convex and $\lambda'(x+1)\leqq K\lambda'(x)$, then

$$\begin{aligned}(R^\lambda, W^\gamma, p, n) \mathrel{\substack{\smile\\ \equiv \\ \frown}} &\left\{\frac{n^{1-\gamma p}}{\lambda(n)}\int_0^1 \lambda'(nx)|\log(1-x)|^p\,dx\right\}^{1/p}+\\ &+\left\{\frac{1}{\lambda(n)}\int_1^{n/2}\tau^{-\gamma p}\Big(\int_0^1 \lambda'(\tau x)|\log(1-x)|^p\,dx\Big)\,d\tau\right\}^{1/p}.\end{aligned} \tag{1.6}$$

Special cases:

(i) If $\lambda'(2x)=O\big(\lambda'(x)\big)$, then

$$(R^\lambda, W^\gamma, p, n) \mathrel{\substack{\smile\\ \equiv \\ \frown}} \left\{\frac{1}{\lambda(n)}\int_1^{n}\tau^{-\gamma p}\Big(\int_0^1 \lambda'(\tau x)|\log(1-x)|^p\,dx\Big)\,d\tau\right\}^{1/p}. \tag{1.7}$$

E.g., considering the means (see [2])

$$K_n(f, \beta, p; x)=\left\{\frac{1}{(n+1)^\beta}\sum_{k=0}^{n}(k+1)^{\beta-1}|s_k(x)-f(x)|^p\right\}^{1/p} \quad (\beta\geqq 1)$$

we may write $\lambda(x)=x^\beta$ and we obtain

$$(K, W^\gamma, p, n) \mathrel{\substack{\smile\\ \equiv \\ \frown}} H_n^{p,\beta,\gamma},$$

where

$$H_n^{p,\beta,\gamma}=\left\{n^{-\beta}\int_1^n x^{\beta-1-\gamma p}\,dx\right\}^{1/p}\sim\begin{cases} n^{-\gamma} & \text{if } \beta>\gamma p\\ n^{-\gamma}(\log n)^{1/p} & \text{if } \beta=\gamma p\\ n^{-\beta/p} & \text{if } \beta<\gamma p.\end{cases}$$

(ii) If

$$T_n(f, p; x)=\left\{e^{-Ax}\sum_{k=1}^{n}e^{Ak}|s_k(x)-f(x)|^p\right\}^{1/p} \quad (A>0),$$

then

$$(T, W^\gamma, p, n) \mathrel{\substack{\smile\\ \equiv \\ \frown}} n^{-\gamma}\log n. \tag{1.8}$$

2. If λ is concave and $\lambda'(x+1)\geqq c\lambda'(x)$ $(c>0)$ then

$$(R^{\lambda}, W^{\gamma}, p, n) \asymp \left\{\frac{1}{\lambda(n)} \int_1^n \lambda'(\tau)\tau^{-\gamma p}\,d\tau\right\}^{1/p}.$$

Specially, for

$$T_n^{\beta,\delta}(f, p; x) = \left\{\frac{1}{n^{\beta}(\log n)^{\delta}} \sum_{k=2}^{n} k^{\beta-1}(\log k)^{\delta}|s_k(x)-f(x)|^p\right\}^{1/p} \quad (0<\beta<1, \ -\infty<\delta<\infty)$$

we have

$$(T^{\beta,\delta}, W^{\gamma}, p, n) \asymp \begin{cases} n^{-\gamma} & \text{if } \beta>\gamma p \\ n^{-\gamma}(\log n)^{1/p} & \text{if } \beta=\gamma p \text{ and } \delta>-1 \\ n^{-\gamma}(\log n)^{1/p}(\log\log n)^{1/p} & \text{if } \beta=\gamma p \text{ and } \delta=-1 \\ n^{-\beta/p}(\log n)^{-\delta/p} & \text{if } \beta<\gamma p \text{ or } \beta=\gamma p \text{ and } \delta<-1. \end{cases}$$

III. *Nörlund means.* Let λ be as in II. The strong Nörlund means are the following

$$N_n^{\lambda}(f, p; x) = \left\{\frac{1}{\lambda(n)} \sum_{k=0}^{n-1} (\lambda(n-k)-\lambda(n-k-1))|s_k(x)-f(x)|^p\right\}^{1/p}.$$

1. If λ is convex and $\lambda'(x+1)\leqq K\lambda'(x)$ then

$$(N^{\lambda}, W^{\gamma}, p, n) \asymp \left\{\frac{1}{\lambda(n)} \int_1^n \lambda'(n-\tau)\tau^{-\gamma p}\,d\tau\right\}^{1/p}.$$

Specially, for

$$\sigma_n^{\beta}|f, p; x| = \left\{\frac{1}{A_n^{\beta}} \sum_{k=0}^{n} A_{n-k}^{\beta-1}|s_k(x)-f(x)|^p\right\}^{1/p} \quad \left(\beta\geqq 1, \ A_n^{\beta} = \binom{n+\beta}{n}\right)$$

we have (see also [2, 3])

$$(\sigma_n^{\beta}, W^{\gamma}, p, n) \asymp H_n^{p,1,\gamma}.$$

2. If λ is concave and $\lambda'(x+1)\geqq c\lambda'(x)$ $(c>0)$ then

$$(N^{\lambda}, W^{\gamma}, p, n) \asymp \left\{\frac{n^{1-\gamma p}}{\lambda(n)} \int_0^1 \lambda'(nx)|\log x|^p\,dx\right\}^{1/p} + \left\{\frac{1}{\lambda(n)} \int_1^{n/2} \tau^{-\gamma p}\lambda'(n-\tau)\,d\tau\right\}^{1/p}.$$

Special cases:

(i) The previous formula for σ^{β} is valid for $0<\beta\leqq 1$, too.

(ii) *The harmonic mean.* This is the mean

$$H(f, p; x) = \left\{\frac{1}{\log n} \sum_{k=0}^{n-1} \frac{1}{n-k}|s_k(x)-f(x)|^p\right\}^{1/p},$$

and for this we have

$$(H, W^{\gamma}, p, n) \asymp \begin{cases} n^{-\gamma}\log n & \text{if } \gamma p\leqq 1 \\ (n\log n)^{-1/p} & \text{if } \gamma p>1. \end{cases}$$

IV. *Generalized de la Vallée Poussin means.* These were defined by L. LEINDLER [4] as follows: Let $\lambda=\{\lambda_n\}$ be an increasing sequence of natural numbers with $\lambda_1=1$, $\lambda_{n+1}-\lambda_n\leqq 1$ $(n=1, 2, \dots)$, and let

$$V_n^\lambda(f, p; x)=\left\{\frac{1}{\lambda_n}\sum_{k=n-\lambda_n+1}^{n}|s_k(x)-f(x)|^p\right\}^{1/p}.$$

For these we have (see also [5, 7])

$$(V^\lambda, W^\gamma, p, n)\underset{\frown}{\overset{\smile}{\equiv}} n^{-\gamma}\log\frac{n}{\lambda_n}+\begin{cases}0 & \text{if } \gamma p<1\\ n^{-\gamma}\log\dfrac{n}{n-\lambda_n+1} & \text{if } \gamma p=1\\ n^{-1/p}(n-\lambda_n+1)^{\gamma-1/p} & \text{if } \gamma p>1.\end{cases} \tag{1.9}$$

V. *De la Vallée Poussin's method.* This is the following matrix-method

$$V_n(f, p, x)=\left\{\frac{1}{n+1}|s_0(x)-f(x)|^p+\frac{3n}{(n+1)(n+2)}|s_1(x)-f(x)|^p+\dots\right.$$

$$\left.+\dots+\frac{(2i+1)n(n-1)\dots(n-i+1)}{(n+1)\dots(n+i+1)}|s_i(x)-f(x)|^p+\dots\right\}^{1/p}.$$

For this the formula

$$(V, W^\gamma, p, n)\underset{\frown}{\overset{\smile}{\equiv}} H^{p,2,\gamma}_{\sqrt{n}}$$

holds true.

VI. *Euler's method.* Let $0<q<1$ and

$$E_n^q(f, p; x)=\left\{\sum_{k=0}^{n}\binom{n}{k}q^k(1-q)^{n-k}|s_k(x)-f(x)|^p\right\}^{1/p}.$$

We have

$$(E^q, W^\gamma, p, n)\underset{\frown}{\overset{\smile}{\equiv}} n^{-\gamma}\log n.$$

VII. *The circle method.* This is analogous to the Euler method and is defined by

$$C_n^q(f, p; x)=\left\{q^{n+1}\sum_{k=n}^{\infty}\binom{k}{n}(1-q)^{k-n}|s_k(x)-f(x)|^p\right\}^{1/p}.$$

For this we have also

$$(C^q, W^\gamma, p, n)\underset{\frown}{\overset{\smile}{\equiv}} n^{-\gamma}\log n.$$

VIII. *Abel's method.* For

$$A^t(f, p; x)=\left\{(1-t)\sum_{k=0}^{\infty}t^k|s_k(x)-f(x)|^p\right\}^{1/p}\quad (0<t<1,\ t\to 1)$$

the formula

$$(A, W^\gamma, p, t)\underset{\frown}{\overset{\smile}{\equiv}} H^{p,1,\gamma}_{1/(1-t)}$$

holds true.

IX. *Lindelöf's method.* If

$$L_t(f, p; x) = \left\{ \sum_{n=1}^{\infty} (n^{-nt} - (n+1)^{-(n+1)t}) |s_k(x) - f(x)|^p \right\}^{1/p} \quad (t > 0,\ t \to 0)$$

then

$$(L, W^\gamma, p, t) \asymp \begin{cases} \left(t \log \frac{1}{t}\right)^\gamma & \text{if } \gamma p < 1 \\ \left(t \log \frac{1}{t}\right)^\gamma \left(\log \frac{1}{t}\right)^{1/p} & \text{if } \gamma p = 1 \\ t^{1/p} & \text{if } \gamma p > 1. \end{cases}$$

X. *Borel's method.* For

$$B_t(f, p, x) = \left\{ e^{-t} \sum_{k=0}^{\infty} \frac{t^k}{k!} |s_k(x) - f(x)|^p \right\}^{1/p} \quad (t > 0,\ t \to \infty)$$

we have

$$(B, W^\gamma, p, t) \asymp t^{-\gamma} \log t.$$

XI. *Mittag-Leffler's method.* Actually, the following is a slight modification of Mittag-Leffler's method (the latter is not non-negative). Let $\Gamma(x)$ be the gamma function and

$$M_t(f, p; x) = \left\{ \sum_{k=1}^{\infty} \left(\frac{1}{\Gamma(2+kt)} - \frac{1}{\Gamma(2+(k+1)t)} \right) |s_k(x) - f(x)|^p \right\}^{1/p} \quad (t > 0,\ t \to 0)$$

(in the original mean stood $\Gamma(1+kt)$).

For this the order of the approximation is given by

(1.10) $$(M, W^\gamma, p, t) \asymp H_{1/t}^{p;1,\gamma}.$$

XII. *The R_2 method.* This is defined by

$$T_t(f, p; x) = \left\{ \frac{2}{\pi} \sum_{k=1}^{\infty} \frac{\sin^2 kt}{k^2 t} |s_k(x) - f(x)|^p \right\}^{1/p} \quad (t > 0,\ t \to 0).$$

For this

$$(T, W^\gamma, p, t) \asymp H_{1/t}^{p;1,\gamma}$$

holds true.

We remark that many of I—XII could be proved by other methods. We listed them only to illustrate the strength of Theorems 1 and 2. Also, anyone interested in the classes $W^r \operatorname{Lip} \alpha := \{f \mid f^{(r)} \in \operatorname{Lip} \alpha\}$ may feel a certain lack, namely if $\gamma = r+1$ with integer r then the class W^γ does not coincide with $W^r \operatorname{Lip} 1$, it is wider than the latter. Unfortunately, we cannot prove Theorem 2 for $W^r \operatorname{Lip} 1$ (perhaps it is not valid at all), but most of the corollaries I—XII hold if we replace $W^r \operatorname{Lip} 1$ for W^γ.

§ 2

PROOF OF THEOREM 1. We write E_k instead of $E_k(f)$. It is clear that

$$\sum_{k=0}^{1} t_k |s_k(x)-f(x)|^p \leqq K_p \tau(p, N_0, N_1) E_0^p.$$

Let $m \geqq 1$. For some permutation π of the set $\{N_m+1, \dots, N_{m+1}\}$ we have

$$\sum_{n=N_m+1}^{N_{m+1}} t_k |s_k(x)-f(x)|^p = \sum_{k=N_m+1}^{N_{m+1}} t_k^* |s_{\pi(k)}(x)-f(x)|^p =$$

$$= \sum_{k=N_m+1}^{N_{m+1}-1} (t_k^*-t_{k+1}^*) \sum_{i=N_m+1}^{k} |s_{\pi(i)}(x)-f(x)|^p + t_{N_{m+1}}^* \sum_{i=N_m+1}^{N_{m+1}} |s_{\pi(i)}(x)-f(x)|^p,$$

and thus, using Theorem A we obtain

$$\sum_{k=N_m+1}^{N_{m+1}} t_k |s_k(x)-f(x)|^p \leqq K_p (E_{N_m})^p \left(\sum_{k=1}^{N_{m+1}-N_m-1} (t_{N_m+k}^* - t_{N_m+k+1}^*) k \left(\log \frac{N_{m+1}}{k} \right)^p + \right.$$

$$\left. + t_{N_{m+1}}^* (N_{m+1}-N_m) \left(\log \frac{N_{m+1}}{N_{m+1}-N_m} \right)^p \right) =$$

$$= K_p (E_{N_m})^p \left(\sum_{k=2}^{N_{m+1}-N_m} t_{N_m+k}^* \left(k \left(\log \frac{N_{m+1}}{k} \right)^p - (k-1) \left(\log \frac{N_{m+1}}{k-1} \right)^p \right) + \right.$$

$$\left. + t_{N_m+1}^* (\log N_{m+1})^p \right) \leqq K_p (E_{N_m})^p \sum_{k=1}^{N_{m+1}-N_m} t_{N_m+k}^* \left(\log \frac{N_{m+1}}{k} \right)^p =$$

$$= K_p (E_{N_m})^p \tau(p, N_m, N_{m+1}),$$

by which we have proved our theorem.

PROOF OF THEOREM 2. For the sake of simplicity we prove Theorem 2 for the sequence $N_m = 2^{m-1}$ $(m=1, 2, \dots)$. The necessary changes in the general case are obvious.

Let

$$Q_{m,n}(x) = \sum_{l=1}^{n} \left(\frac{\cos (m-l)x}{l} - \frac{\cos (m+l)x}{l} \right) = 2 \sin mx \sum_{l=1}^{n} \frac{\sin lx}{l}$$

be the Fejér-polynomials and

$$f(x) = \cos 4x + \sum_{k=2}^{\infty} \chi(2^k) 2^{-k\gamma} Q_{2^k, 2^{k-2}}(x).$$

As $|Q_{m,n}| \leqq 8$, we have for $2^{m+1} < n \leqq 2^{m+2}$

$$E_n(f) \leqq |s_{2^{m+1}-2^{m-1}-1}(f) - f| \leqq \sum_{k=m+1}^{\infty} 8\chi(2^k) 2^{-k\gamma} \leqq 8\chi(2^{m+1}) \sum_{k=m+1}^{\infty} 2^{-k\gamma} =$$

$$= K\chi(2^{m+1}) 2^{-(m+1)\gamma} \leqq K\chi(n) n^{-\gamma},$$

i.e., $f \in W^\gamma \chi$.

For $0 \leqq \lambda \leqq n-1$

$$|s_{m\pm\lambda}(Q_{m,n}; 0)| = \sum_{l=\lambda+1}^{n} \frac{1}{l} \geqq c \log \frac{n}{\lambda},$$

by which

$$\sum_{k=2^{m-1}+1}^{2^m} t_k |s_k(f; 0) - f(0)|^p \geqq c(\chi(2^m))^p 2^{-m\gamma p} \left(\sum_{\lambda=1}^{2^{m-3}} t_{2^{m-1}+\lambda} \left(\log \frac{2^{m-3}}{\lambda} \right)^p + \right.$$

$$\left. + \sum_{\lambda=1}^{2^{m-2}} t_{2^m-\lambda} \left(\log \frac{2^{m-2}}{\lambda} \right)^p \right) \geqq c(\chi(2^m))^p 2^{-m\gamma p} \tau(p, 2^{m-1}, 2^m)$$

on the ground of which the proof of (1.3) is very easy.

Finally, we sketch very briefly the proofs of I—XII.

Ad I. The Hölder inequality gives (see also the proof of Theorem 1) with $p' = \dfrac{q'}{q'-1}$

$$\tau_n(p, N_{m-2}, N_m) = \sum_{k=1}^{N_m - N_{m-2}} \alpha^*_{n, N_{m-2}+k} \left(\log \frac{N_m}{k} \right)^p (A_n(N_{m-2}, N_m))^{-1} \leqq$$

$$\leqq \left\{ \sum_{k=N_{m-2}+1}^{N_{m+1}} (\alpha^*_{nk})^{q'} \right\}^{1/q'} \left\{ \sum_{k=1}^{N_m - N_{m-1}} \left(\log \frac{N_m}{k} \right)^{p'p} \right\}^{1/p'} (A_n(N_{m-2}, N_m))^{-1} \leqq K N_m^{-1/p'} N_m^{1/p'} \leqq K$$

where we used that

$$\sum_{k=1}^{N_m} \left(\log \frac{N_m}{k} \right)^p \leqq K \int_1^{N_m} \left(\log \frac{N_m}{x} \right)^p dx \leqq K_p N_m.$$

Theorem 1 and the previous estimate already give (1.4), which in turn implies (1.5).

Ad II. In the following we shall write $\gamma_i \sim \delta_i$ if $\dfrac{1}{c}\delta_i \leqq \gamma_i \leqq c\delta_i$ for some $c > 0$, and once for all let $\mathcal{N} = \{N_m\}$ be the sequence $N_m = 2^{m-1}$ $(m = 1, 2, \ldots)$.

1. When λ is convex then $\lambda' \nearrow$, $\lambda(k+1) - \lambda(k) \leqq \lambda(k+2) - \lambda(k+1)$, and thus, by $\lambda'(x+1) \sim \lambda'(x)$ we obtain

$$\tau_n(p, \mu/2, \mu) = (\lambda(n))^{-1} \sum_{k=\mu/2+1}^{\mu} (\lambda(k+1) - \lambda(k)) \left(\log \frac{\mu}{\mu - k + 1} \right)^p \sim$$

$$\sim (\lambda(n))^{-1} \sum_{k=\mu/2}^{\mu-1} (\lambda(k+1)) - \lambda(k)) \left(\log \frac{\mu}{\mu - k} \right)^p \sim (\lambda(n))^{-1} \sum_{k=\mu/2}^{\mu-1} \int_k^{k+1} \lambda'(x) \left(\log \frac{\mu}{\mu - x} \right)^p dx \sim$$

$$\sim (\lambda(n))^{-1} \int_{\mu/2}^{\mu} \lambda'(x) \left(\log \frac{\mu}{\mu - x} \right)^p dx \sim (\lambda(n))^{-1} \int_0^{\mu} \lambda'(x) \left(\log \frac{\mu}{\mu - x} \right)^p dx \sim$$

$$\sim (\lambda(n))^{-1} \mu \int_0^1 \lambda'(\mu x) |\log(1-x)|^p dx.$$

Now the expression on the right increases if μ does so, by which we obtain for $N_{m_0}<n\leqq N_{m_0+1}$

$$R_n^\lambda(W^\gamma, \mathcal{N}, p) \sim \left\{n^{-\gamma p}\tau_n(p, n/2, n)+\sum_{\mu=0}^{m_0-1} 2^{-\gamma\mu p}\tau_n(p, 2^\mu, 2^{\mu+1})\right\}^{1/p} \sim$$

$$\sim\left\{\frac{n^{1-\gamma p}}{\lambda(n)}\int_0^1 \lambda'(nx)|\log(1-x)|^p\,dx\right\}^{1/p}+\left\{\frac{1}{\lambda(n)}\int_1^{n/2}\tau^{-\gamma p}\left(\int_0^1 \lambda'(\tau x)|\log(1-x)|^p\,dx\right)d\tau\right\}^{1/p}$$

and this is the same as (1.6).

If $\lambda'(2x)\sim\lambda'(x)$, then the second term on the right of (1.6) majorizes the first one, thus (1.7) follows immediately.

If $\lambda(x)=e^{Ax}$, then $\lambda'(x)\sim e^{Ax}$ and

$$\mu\int_0^1 \lambda'(\mu x)|\log(1-x)|^p\,dx \sim e^{A\mu}(\log\mu)^p$$

on the ground of which the proof of (1.8) is easy.

2. If λ is concave, then $\lambda'\searrow$ and thus

$$\tau_n(p, \mu/2, \mu) \sim (\lambda(n))^{-1}\sum_{k=\mu/2}^{\mu-1}(\lambda(k+1)-\lambda(k))\left|\log\frac{\mu}{k-\frac{\mu}{2}+1}\right|^p \sim$$

$$\sim (\lambda(n))^{-1}\mu\int_1^2 \lambda'\left(\frac{\mu}{2}x\right)|\log(x-1)|^p\,dx.$$

Let the expression on the right be $h(\mu/2)$. The monotonicity of λ' gives

$$\sum_{\mu=0}^{m} 2^{-\mu\gamma p}h(2^\mu) \sim (\lambda(n))^{-1}\sum_{\mu=0}^{m} 2^{-\mu\gamma p}2^\mu\lambda'(2^\mu)\int_1^2 |\log(x-1)|^p\,dx \sim$$

$$\sim (\lambda(n))^{-1}\int_1^{2^m} \lambda'(\tau)\tau^{-\gamma p}\,d\tau$$

thus we obtain for $N_{m_0}<n\leqq N_{m_0+1}$

$$R_n^\mu(W^\gamma, \mathcal{N}, p) \sim \left\{n^{-\gamma p}\tau_n(p, n/2, n)+\sum_{\mu=0}^{m_0} 2^{-\mu\gamma p}\tau_n(p, 2^\mu, 2^{\mu+1})\right\}^{1/p} \sim$$

$$\sim\left\{\frac{1}{\lambda(n)}\int_1^n \lambda'(\tau)\tau^{-\gamma p}\,d\tau\right\}^{1/p}.$$

The proof of III is similar.

Ad IV. First observe that

$$\sum_{k=n-\lambda+1}^{n}\left(\log\frac{n+1}{n+1-k}\right)^p \sim \left(\log\frac{2n}{\lambda}\right)^p\lambda \tag{2.1}$$

thus we have either

$$V_n^\lambda(W^\gamma, \mathcal{N}, p) \sim n^{-\gamma} \log \frac{2n}{\lambda_n} \qquad \left(\lambda_n \leqq \frac{n}{2}\right)$$

or, using (2.1) with $\lambda = n/2$,

$$V_n^\lambda(W^\gamma, \mathcal{N}, p) \sim \left\{ \frac{1}{\lambda_n} \sum_{\mu=\log(n-\lambda_n)}^{\log n+1} 2^\mu 2^{-\mu\gamma p} \right\}^{1/p} \sim \left\{ \frac{1}{\lambda_n} \int_{n-\lambda_n}^{n} \tau^{-\gamma p}\, d\tau \right\}^{1/p},$$

which proves (1.9).

Ad V. As for large n and $i \leqq 2\sqrt{n}$

$$\frac{n(n-1)\cdot \ldots \cdot(n-i+1)}{(n+2)\cdot \ldots \cdot (n+i+1)} \geqq \prod_{i=1}^{2\sqrt{n}} \left(1 - \frac{2i}{n+i+1}\right) \geqq \frac{1}{2} \exp\left(-\sum_{i=1}^{2\sqrt{n}} \frac{2i}{n+i+1}\right) \geqq c > 0,$$

we obtain by (2.1)*

$$\text{(2.2)} \quad \left\{ \sum_{\mu=0}^{\log\sqrt{n}} 2^{-\mu\gamma p} \tau_n(p, 2^\mu, 2^{\mu+1}) \right\}^{1/p} \sim \left\{ \frac{1}{n} \sum_{\mu=0}^{\log\sqrt{n}} 2^{-\mu\gamma p} 2^\mu \sum_{k=2^\mu+1}^{2^{\mu+1}} (\log(2^{\mu+1}/k))^p \right\}^{1/p} \sim$$

$$\sim \left\{ \frac{1}{n} \sum_{\mu=0}^{\log\sqrt{n}} 2^{2\mu-\mu\gamma p} \right\}^{1/p} \sim \left\{ \frac{1}{n} \int_1^{\sqrt{n}} \tau^{1-\gamma p}\, d\tau \right\}^{1/p} \sim H_{\sqrt{n}}^{p,2,\gamma},$$

and so

$$V_n(W^\gamma, \mathcal{N}, p) \geqq c H_{\sqrt{n}}^{p,2,\gamma}.$$

On the other hand if $i > 2\sqrt{n}$, then $t_{n,i+1} < t_{n,i}$ and

$$t_{n,2i}/t_{n,i} = \frac{4i+1}{2i+1} \frac{(n-i)\cdot \ldots \cdot (n-2i+1)}{(n+i+2)\cdot \ldots \cdot (n+2i+1)} \leqq 2 \exp\left(-\sum_{j=i+1}^{2i} \frac{2j}{n+j+1}\right) \leqq$$

$$\leqq 2 \exp\left(-2\frac{i^2}{2n}\right) < \frac{1}{4},$$

by which

$$\sum_{\mu=\log\sqrt{n}}^{\log n} 2^{-\mu\gamma p} \tau_n(p, 2^\mu, 2^{\mu+1}) \leqq K(\sqrt{n})^{-\gamma p} \sum_{\mu=\log\sqrt{n}}^{\log n} t_{n,2^\mu} \sum_{k=2^\mu+1}^{2^{\mu+1}} (\log(2^{\mu+2}/k))^p \leqq$$

$$\leqq K(\sqrt{n})^{-\gamma p} \sum_{\mu=\log\sqrt{n}}^{\log n} t_{n,2^\mu} 2^\mu \leqq K(\sqrt{n})^{-\gamma p} t_{n,2\sqrt{n}} \sum_{\mu=1}^{\infty} \left(\frac{1}{4}\right)^\mu 2^{\mu+\log\sqrt{n}} \leqq K(\sqrt{n})^{-\gamma p},$$

and we have to remark only that $(\sqrt{n})^{-\gamma} \leqq K H_{\sqrt{n}}^{p,2,\gamma}$ (see the estimate of (2.2)).

Ad VI. Let $M_n = [(n+1)q]$. As

$$\sum_{|h|>\sqrt{n}} t_{n,M+h} < \frac{1}{2}$$

* From now on $\log(.)$ stands for $\log_2(.)$.

holds if n is large (see [1, p. 201]), we have

$$(2.3)\qquad E_n^q(W^\gamma, \mathcal{N}, p) \geqq cn^{-\gamma}\log n \sum_{|h|<\sqrt{n}} t_{n,M+h} > cn^{-\gamma}\log n.$$

On the other hand

$$\sum_{k=0}^{M/2}\binom{n}{k} q^k(1-q)^{n-k}(\log n)^p = O\left(e^{-nq^2/12}(\log n)^p\right) = o(n^{-\gamma}\log n)$$

(see [1, p. 201]) and

$$n^{-\gamma}\left\{\sum_{k=M/2}^{n}\binom{n}{k} q^k(1-q)^{n-k}(\log n)^p\right\}^{1/p} = O(n^{-\gamma}\log n),$$

by which

$$E_n^q(W^\gamma, \mathcal{N}, p) \leqq Kn^{-\gamma}\log n,$$

and this, together with (2.3), proves VI.

The proofs of VII and X are similar (see the estimates of [1, Theorems 137—139]).

Ad VIII. As

$$\tau_t(p, \mu, 2\mu) \sim (1-t)\int_\mu^{2\mu} t^x\left(\log\frac{\mu}{x-\mu}\right)^p dx \sim (1-t)\mu\int_1^2 t^{\mu x}|\log(x-1)|^p\,dx,$$

we have

$$A_t(W^\gamma, \mathcal{N}, p) \sim \left\{(1-t)\sum_{\mu=0}^\infty 2^{-\mu\gamma p}2^\mu\int_1^2 t^{2^\mu x}|\log(x-1)|^p\,dx\right\}^{1/p} \sim$$

$$\sim \left\{(1-t)\sum_{\mu=0}^\infty 2^{-\mu\gamma p}2^\mu t^{2^\mu}\int_1^2 |\log(x-1)|^p\,dx\right\}^{1/p} \sim$$

$$\sim \left\{(1-t)\int_1^\infty \tau^{-\gamma p}t^\tau\,d\tau\right\}^{1/p} \sim \left\{(1-t)\int_1^{1/(1-t)} \tau^{-\gamma p}\,d\tau\right\}^{1/p} \sim H_{1/(1-t)}^{p,1,\gamma}.$$

Ad IX. A simple computation gives that

$$n^{-nt}-(n+1)^{-(n+1)t} \sim n^{-nt}t\log n \qquad (t\log n \leqq 1)$$

and

$$n^{-nt}-(n+1)^{-(n+1)t} \sim t\log n \qquad (tn\log n \leqq 2)$$

by which

$$\tau_t(p, \mu, 2\mu) \sim t\log\mu\sum_{k=1}^{\mu}\left(\log\frac{\mu}{k}\right)^p \sim t\mu\log\mu \qquad (\mu \leqq 2(t\log 1/t)^{-1}),$$

$$\tau_t(p, \mu, 2\mu) \leqq K\mu^{-\mu t}\mu t\log\frac{1}{t} \qquad \left(2\left(t\log\frac{1}{t}\right)^{-1} \leqq \mu \leqq 4/t\right)$$

and

$$\tau_t(p, \mu, 2\mu) \leqq K\mu^{-\mu t}\mu \qquad (4/t < \mu).$$

Thus, with $\mu_t=[\log(2/(t\log 1/t))]$ we obtain

$$\left\{\sum_{\mu=0}^{\mu_t} 2^{-\mu\gamma t}\tau_t(p, 2^\mu, 2^{\mu+1})\right\}^{1/p} \sim \left\{t\int_1^{1/(t\log 1/t)} \tau^{-\gamma p}\log\tau\, d\tau\right\}^{1/p} \sim$$

$$\sim \begin{cases} (t\log 1/t)^\gamma & \text{if } \gamma p<1 \\ t^\gamma(\log 1/t)^{\gamma+1/p} & \text{if } \gamma p=1 \\ t^{1/p} & \text{if } \gamma p>1, \end{cases}$$

$$\left\{\sum_{\mu=\mu_t}^{\log(4/t)} 2^{-\mu\gamma p}\tau_t(p, 2^\mu, 2^{\mu+1})\right\}^{1/p} \leqq K\left(t\log\frac{1}{t}\right)^{1/p} 2^{-\gamma\mu_t}\left\{\sum_{\mu=\mu_t}^{\infty} 2^{-t\cdot 2^\mu\mu+\mu}\right\}^{1/p} \leqq$$

$$\leqq K\left(t\log\frac{1}{t}\right)^{1/p} 2^{-\gamma\mu_t}\{2^{-t2^{\mu_t}\mu_t+\mu_t}\}^{1/p} \leqq K\left(t\log\frac{1}{t}\right)^\gamma$$

and

$$\left\{\sum_{\mu=\log(4/t)}^{\infty} 2^{-\mu\gamma p}\tau(p, 2^\mu, 2^{\mu+1})\right\}^{1/p} \leqq Kt^\gamma \sum_{\mu=\log(4/t)}^{\infty} \exp[(-2^\mu t+1)\log 2^\mu] \leqq Kt^\gamma.$$

Collecting the above estimates we obtain (1.10).

Ad XI. If $n\leqq 2/t$ then

$$\frac{1}{\Gamma(2+nt)}-\frac{1}{\Gamma(2+(n+1)t)} = t\left(\frac{1}{\Gamma(x)}\right)'\Bigg|_{x=\vartheta} \sim t \qquad (2<\vartheta<4),$$

by which

$$\left\{\sum_{\mu=0}^{\log 1/t} 2^{-\mu\gamma p}\tau_t(p, 2^\mu, 2^{\mu+1})\right\}^{1/p} \sim \left\{t\int_1^{1/t} \tau^{-\gamma p}\, d\tau\right\}^{1/p} \sim H_{1/t}^{p,1,\gamma}. \tag{2.4}$$

On the other hand, using the formula $(1/\Gamma(x))'\sim\log x/\Gamma(x)$, we obtain with $\mu_t=[\log 1/t]$

$$\left\{\sum_{\mu=\mu_t}^{\infty} 2^{-\lambda\gamma p}\tau_t(p, 2^\mu, 2^{\mu+1})\right\}^{1/p} \leqq$$

$$\leqq K\left\{2^{-\mu_t\gamma p}\sum_{\mu=\mu_t}^{\infty}\left((\Gamma(2+2^\mu t))^{-1}-(\Gamma(2+(2^\mu+1)t))^{-1}\right)\sum_{k=2^\mu+1}^{2^{\mu+1}}\left(\log(2^{\mu+1}/(k-2^\mu))\right)^p\right\}^{1/p} \leqq$$

$$\leqq Kt^\gamma\left\{\sum_{\mu=\mu_t}^{\infty} t\frac{\log 2^\mu t}{\Gamma(2^\mu t)}2^\mu\right\}^{1/p} \leqq Kt^\gamma$$

and this, together with (2.4), proves (1.10).

Ad XII. On the ground of the estimates

$$\frac{(\sin nt)^2}{n^2t}\sim t \quad (nt\leqq 1); \qquad \frac{(\sin nt)^2}{n^2t}\leqq\frac{1}{n^2t} \quad (nt>1)$$

the proof is very easy (see the above computations).

REFERENCES

[1] HARDY, G. H.: *Divergent series*, Oxford, 1949.
[2] LEINDLER, L.: Über die Approximation im starken Sinne, *Acta Math. Acad. Sci. Hungar.* **16** (1965), 255—262.
[3] LEINDLER, L.: Bemerkung zur Approximation im starken Sinne, *Acta Math. Acad. Sci. Hungar.* **18** (1967), 273—277.
[4] LEINDLER, L.: On summability of Fourier series, *Acta Sci. Math.* **29** (1968), 147—162.
[5] LEINDLER, L.: On the strong approximation of Fourier series, *Acta Sci. Math.* **38** (1976), 317—324.
[6] TOTIK, V.: On the very strong and mixed approximation, *Acta Sci. Math.* **41** (1979), 419—428.
[7] TOTIK, V.: On the strong approximation of Fourier series, *Acta Math. Acad. Sci. Hungar.* **35** (1980), 151—172.

(Received April 18, 1980)

JÓZSEF ATTILA TUDOMÁNYEGYETEM
BOLYAI INTÉZET
ARADI VÉRTANÚK TERE 1
H—6720 SZEGED
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 241—246

A NEW PROOF OF S. A. TELJAKOVSKII'S APPROXIMATION THEOREM

by

T. M. MILLS and A. K. VARMA

1. Introduction

In this paper we present a new simple proof of S. A. TELJAKOVSKIĬ's [7] fundamental approximation theorem.

Let $U_{n-1}(x)$ be the Chebyshev polynomial of the second kind of degree $n-1$, and denote the zeros of $(1-x^2)U_{n-1}(x)$ by

$$x_{k,n} = \cos t_{k,n} = \cos\frac{k\pi}{n}, \quad k = 0, 1, \dots, n. \tag{1.1}$$

The fundamental polynomials, $l_{k,n}(x)$, of Lagrange interpolation based on the nodes $x_{k,n}$ are given as follows (where $x=\cos t$):

$$\begin{cases} l_{0,n}(x) = \dfrac{\sin nt \sin t}{2n(1-\cos t)}, \\ l_{n,n}(x) = \dfrac{(-1)^{n+1}\sin nt \sin t}{2n(1+\cos t)}, \end{cases} \tag{1.2}$$

$$l_{k,n}(x) = \frac{(-1)^{k+1}\sin nt \sin t}{n(\cos t-\cos t_k)} \qquad k = 1, 2, \dots, n-1. \tag{1.3}$$

Now consider the following interpolation process (first suggested by S. N. BERNSTEIN [2]) based on the nodes (1.1).

$$H_n[f, x] = \sum_{k=0}^{n} f(x_{k,n})\lambda_{k,n}(x) \tag{1.4}$$

where

$$\begin{cases} \lambda_{0,n} = l_{0,n}(x), \quad \lambda_{n,n}(x) = l_{n,n}(x), \\ \lambda_{k,n}(x) = \dfrac{l_{k-1,n}(x)+2l_{k,n}(x)+l_{k+1,n}(x)}{4}, \quad k = 2, 3, \dots, n-2 \\ \text{and} \\ \lambda_{1,n}(x) = \dfrac{3l_{1,n}(x)+l_{2,n}(x)}{4}, \quad \lambda_{n-1,n}(x) = \dfrac{3l_{n-1,n}(x)+l_{n-2,n}(x)}{4}. \end{cases} \tag{1.5}$$

We can now state the main result of this paper.

THEOREM 1. *Let $f \in C([-1, 1])$ with modulus of continuity $w(f; \delta)$. Then there exists a positive constant c which is independent of n, x and f such that*

$$|H_n[f, x]-f(x)| \leqq cw\left(f; \frac{\sqrt{1-x^2}}{n}\right)$$

for all $n \geqq 3$ and all $x \in [-1, 1]$.

2. Preliminaries

In this section we shall list various inequalities and formulae that will be useful in proving Theorem 1. Let us fix n and then, for simplicity, denote $x_{k,n}$ by x_k, $l_{k,n}(x)$ by $l_k(x)$, and $\lambda_{k,n}(x)$ by $\lambda_k(x)$.

D. L. BERMAN ([1], p. 251) has shown that

$$|l_k(x)| \leqq 2 \quad \text{for} \quad x \in [-1, 1] \quad \text{and} \quad k = 0, 1, \ldots, n. \tag{2.1}$$

From (2.1) and (1.5) it follows that

$$|\lambda_k(x)| \leqq 2 \quad \text{for} \quad x \in [-1, 1] \quad \text{and} \quad k = 0, 1, \ldots, n. \tag{2.2}$$

We deduce, from the definitions of $\lambda_k(x)$ and $l_k(x)$, that

$$\sum_{k=0}^{n} \lambda_k(x) \equiv \sum_{k=0}^{n} l_k(x) \equiv 1. \tag{2.3}$$

One can easily check that, for $k=2, 3, \ldots, n-2$,

$$\sin t_k \leqq 2 \sin t_{k+1} \leqq 2(\sin t + \sin t_{k+1}) \leqq 4 \sin (1/2)(t+t_{k+1}) \tag{2.4}$$

and

$$\sin \frac{(2k-1)\pi}{2n} \leqq 2 \sin t_{k-1} \leqq 4 \sin (1/2)(t+t_{k-1}). \tag{2.5}$$

For $r=0, 1, \ldots, n$ we have

$$|\sin nt| = |\sin nt - \sin nt_r| \leqq 2n |\sin (1/2)(t_r-t)|, \tag{2.6}$$

$$0 \leqq \sin t \leqq \sin t + \sin t_r \leqq 2 \sin (1/2)(t+t_r), \tag{2.7}$$

and

$$\sin (1/2)(t+t_r) \geqq |\sin (1/2)(t-t_r)|. \tag{2.8}$$

Now, let $x=\cos t$ where

$$\frac{(j-1)\pi}{n} \leqq t \leqq \frac{j\pi}{n}, \quad j = 1, 2, \ldots, n. \tag{2.9}$$

If $2 \leqq k \leqq j-3$ then $k=j-i$ where $3 \leqq i \leqq j-2$ and hence

$$|\sin (1/2)(t-t_{k+m})| \geqq (i-m-1)/n \quad \text{for} \quad m = 1, 0, -1. \tag{2.10}$$

If $j+2 \leqq k \leqq n-2$ then $k=j+i$ where $2 \leqq i \leqq n-j-2$ and hence

$$|\sin (1/2)(t-t_{k+m})| \geqq (i+m)/n \qquad \text{for} \quad m = 1, 0, -1. \tag{2.11}$$

3. Another representation of the fundamental functions

From (1.2), (1.3) and (1.5) it follows that, for $k=2, 3, \ldots, n-2$,

$$(3.1) \qquad l_{k-1}(x)+l_k(x) = \frac{2(-1)^k \sin nt \sin t \sin(\pi/(2n)) \sin((2k-1)\pi/(2n))}{n(x-x_k)(x-x_{k-1})}$$

and

$$(3.2) \qquad l_k(x)+l_{k+1}(x) = \frac{2(-1)^{k+1} \sin nt \sin t \sin(\pi/(2n)) \sin((2k+1)\pi/(2n))}{n(x-x_k)(x-x_{k+1})}.$$

If we write

$$\sin((2k+1)\pi/(2n)) = \sin((2k-1)\pi/(2n))+2\cos(k\pi/n)\sin(\pi/(2n))$$

in (3.2) and then add (3.1) to (3.2) then we obtain

$$(3.3) \qquad \lambda_k(x) = p_k(x)+q_k(x)$$

where

$$(3.4) \qquad p_k(x) = \frac{(-1)^{k+1} \sin nt \sin t \sin^2(\pi/(2n)) \cos(k\pi/n)}{n(x-x_k)(x-x_{k+1})}$$

and

$$(3.5) \qquad q_k(x) = \frac{(-1)^k \sin nt \sin t \sin(\pi/(2n)) \sin(\pi/n) \sin(k\pi/n) \sin((2k-1)\pi/(2n))}{n(x-x_{k-1})(x-x_k)(x-x_{k+1})}.$$

4. Estimates

In this section we derive certain estimates which will be necessary to prove Theorem 1.

LEMMA 4.1. *For* $-1 \leqq x \leqq 1$,

$$(4.1) \qquad \sum_{k=2}^{n-2} |\lambda_k(x)| \leqq c_1,$$

$$(4.2) \qquad \sum_{k=2}^{n-2} |x-x_k|\,|\lambda_k(x)| \leqq c_2 \sqrt{(1-x^2)}/n,$$

where c_1, c_2 *(and later* c_3, c_4, ...*) are absolute constants.*

PROOF. Let $x=\cos t$ where $(j-1)\pi/n \leqq t \leqq j\pi/n$ for some $j=1, 2, \ldots$, or n. Now

$$(4.3) \qquad \sum_{k=2}^{n-2} |\lambda_k(x)| = \sum_{k=2}^{j-3} |\lambda_k(x)| + \sum_{k=j-2}^{j+1} |\lambda_k(x)| + \sum_{k=j+2}^{n-2} |\lambda_k(x)|.$$

To prove (4.1) we shall estimate each of these sums.

From (3.4), (2.7), (2.8), and (2.10) it follows that

$$(4.4) \qquad \sum_{k=2}^{j-3} |p_k(x)| \leqq c_3 \sum_{i=3}^{j-2} \frac{1}{(i-1)(i-2)^2} \leqq c_4.$$

From (3.5), (2.4), (2.5), (2.7) and (2.10) we obtain

$$\sum_{k=2}^{j-3} |q_k(x)| \leqq c_5 \sum_{i=3}^{j-2} \frac{1}{i(i-1)(i-2)} \leqq c_6. \tag{4.5}$$

Therefore, from (3.3), (4.4) and (4.5) we obtain

$$\sum_{k=2}^{j-3} |\lambda_k(x)| \leqq c_4 + c_6. \tag{4.6}$$

From (2.2) it follows that

$$\sum_{k=j-2}^{j+1} |\lambda_k(x)| \leqq 8. \tag{4.7}$$

In the same way that (4.6) was established, one can show that

$$\sum_{k=j+2}^{n-2} |\lambda_k(x)| \leqq c_7. \tag{4.8}$$

Now (4.1) follows from (4.3), (4.6), (4.7) and (4.8).
To prove (4.2) we use (3.3)—(3.5), (2.4)—(2.5) to obtain

$$n \sum_{k=2}^{n-2} \frac{|x-x_k|\,|\lambda_k(x)|}{\sqrt{(1-x^2)}} \leqq \tag{4.9}$$

$$\leqq \sum_{k=2}^{n-2} \frac{|\sin nt| \sin^2\big(\pi/(2n)\big)}{|x-x_{k+1}|} + 4 \sum_{k=2}^{n-2} \frac{|\sin nt| \sin(\pi/n) \sin\big(\pi/(2n)\big)}{|\sin(1/2)(t-t_{k+1})|\,|\sin(1/2)(t-t_{k-1})|}.$$

From (2.6), (2.8)—(2.11) it follows (by proceeding as in the proof of (4.1)) that

$$n \sum_{k=2}^{n-2} \frac{|x-x_k|\,|\lambda_k(x)|}{\sqrt{(1-x^2)}} \leqq c_7 + c_8. \tag{4.10}$$

From this follows (4.2). Thus, Lemma 4.1 is established.

LEMMA 4.2. *For* $-1 \leqq x \leqq +1$,

$$|f(x)-f(x_k)|\,|\lambda_k(x)| \leqq c_9 w\left(f; \frac{\sqrt{(1-x^2)}}{n}\right), \tag{4.11}$$

$k=0, 1$ *and* $k=n-1, n$.

PROOF. If $x=\pm 1$ then equality holds in (4.11). Using the properties of the modulus of continuity, we obtain

$$|f(x)-f(x_k)|\,|l_k(x)| \leqq w(f; |x-x_k|)|l_k(x)| \leqq$$

$$\leqq w\left(f; \frac{\sqrt{1-x^2}}{n}\right)\left(\frac{n|x-x_k|}{\sqrt{1-x^2}}+1\right)|l_k(x)|.$$

From (2.1) and

$$n\frac{|x-x_k|\,|l_k(x)|}{\sqrt{1-x^2}} \leqq |\sin nt| \leqq 1$$

it follows that

$$|f(x)-f(x_k)|\,|l_k(x)| \leqq 3w\left(f;\ \sqrt{1-x^2}/n\right). \tag{4.12}$$

Therefore we have

$$|f(x)-f(1)|\,|\lambda_0(x)| \leqq 3w\left(f;\ \sqrt{(1-x^2)}/n\right), \tag{4.13}$$

$$|f(x)-f(-1)|\,|\lambda_n(x)| \leqq 3w\left(f;\ \sqrt{1-x^2}/n\right). \tag{4.14}$$

For $k=1$,

$$\begin{aligned}
|f(x)-f(x_1)|\,|\lambda_1(x)| &\leqq \frac{3}{4}|f(x)-f(x_1)|\,|l_1(x)|+(1/4)|f(x)-(x_1)|\,|l_2(x)| \leqq \\
&\leqq \frac{9}{4}w\left(f;\ \sqrt{1-x^2}/n\right)+(1/4)|f(x)-f(x_2)|\,|l_2(x)|+(1/4)|f(x_2)-f(x_1)|\,|l_2(x)| \leqq \\
&\leqq 3w\left(f;\ \sqrt{1-x^2}/n\right)+c_{10}w(f;\ 1/n^2)|l_2(x)| \leqq \\
&\leqq 3w\left(f;\ \sqrt{1-x^2}/n\right)+c_{10}w\left(f;\ \sqrt{1-x^2}/n\right)\left(1+\frac{1}{n\sqrt{1-x^2}}\right)|l_2(x)| \leqq \\
&\leqq 3w\left(f;\ \sqrt{1-x^2}/n\right)+c_{11}w\left(f;\ \sqrt{1-x^2}/n\right) \leqq \\
&\leqq c_{12}w\left(f;\ \sqrt{1-x^2}/n\right).
\end{aligned}$$

Similarly,

$$|f(x)-f(x_{n-1})|\,|\lambda_{n-1}(x)| \leqq c_{13}w\left(f;\ \sqrt{1-x^2}/n\right).$$

This proves Lemma 4.2.

5. Proof of Theorem 1

From (1.4) and (2.3) one immediately obtains

$$\begin{aligned}
H_n(f,x)-f(x) &= \sum_{k=0}^{n}\left(f(x_k)-f(x)\right)\lambda_k(x) = \\
&= \left(\sum_{k=0}^{1}+\sum_{k=2}^{n-2}+\sum_{k=n-1}^{n}\right)\left(f(x_k)-f(x)\right)\lambda_k(x).
\end{aligned} \tag{5.1}$$

Lemma 4.2 implies

$$\left(\sum_{k=0}^{1}+\sum_{k=n-1}^{n}\right)\left|\left(f(x_k)-f(x)\right)\lambda_k(x)\right| \leqq 4c_9w\left(f;\ \sqrt{1-x^2}/n\right). \tag{5.2}$$

Using Lemma 4.1 and the fact that $H_n(f, \pm 1) = f(\pm 1)$ one obtains

$$\sum_{k=2}^{n-2} |f(x_k)-f(x)|\,|\lambda_k(x)| \leqq \tag{5.3}$$

$$\leqq w\left(f;\, \sqrt{1-x^2}/n\right) \sum_{k=2}^{n-2} \left(1+\frac{n|x-x_k|}{\sqrt{1-x^2}}\right) |\lambda_k(x)| \leqq (c_1+c_2)\, w\left(f;\, \sqrt{1-x^2}/n\right).$$

By combining (5.1)—(5.3) we obtain (1.6) and Theorem 1 is established.

Related results on this problem are due to T. M. MILLS and A. K. VARMA [6], A. K. Varma [8], G. FREUD and A. SHARMA [4], P. VÉRTESI and O. KIS [9].

REFERENCES

[1] BERMAN, D. L.: On certain linear operators, *Dokl. Akad. Nauk SSSR* **73** (1950), 249—253 (in Russian).

[2] BERNSTEIN, S. N.: *Collected Works*, Vol. 2, Akad. Nauk SSSR, Moscow, 1954, pp. 130—140 (in Russian).

[3] DEVORE, R. A.: Degree of Approximation, *Approximation Theory* II, edited by G. G. Lorentz, C. K. Chui and L. L. Schumaker, Academic Press, New York, 117—161.

[4] FREUD, G. and SHARMA, A.: Some good sequences of interpolating polynomials, Addendum, *Canad. J. Math.* **29** (1977), 1163—1166.

[5] KIS, O.: Investigation of an interpolation process I, *Acta Math. Acad. Sci. Hungar.* **25** (1974), 363—386 (in Russian).

[6] MILLS, T. and VARMA, A. K.: A new proof of A. F. Timan's approximation theorem, *Israel J. Math.* **18** (1974), 39—44.

[7] TELJAKOVSKIĬ, S. A.: Two theorems on the approximation of functions by algebraic polynomials, *Mat. Sb.* **70** (112) (1966), 252—265.

[8] VARMA, A. K.: On an interpolation process of S. N. Bernstein, *Acta Math. Acad. Sci. Hungar.* **31** (1978), 81—87.

[9] VÉRTESI, P. and KIS, O.: On a new interpolation process, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **10** (1967), 117—128.

(Received April 22, 1980)

DEPARTMENT OF MATHEMATICS
BENDIGO COLLEGE OF ADVANCED EDUCATION
BENDIGO, VICTORIA 3550
AUSTRALIA

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF FLORIDA
GAINESVILLE, FL 32601
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 247—253

THE LAW OF LARGE NUMBERS FOR ADDITIVE FUNCTIONS

by

IMRE Z. RUZSA

1. Introduction

Let f_x, $x=1, 2, \dots$ be a sequence of additive arithmetical functions, where f_x is defined at least on the interval $[1, x]$. Our aim is to find the condition under which the relative frequencies

$$\frac{1}{x} |\{n: n \leqq x, f_x(n)-a_x \leqq u\}|$$

converge to the improper law as $x\to\infty$, with suitably chosen centering constants a_x.

Starting with TURÁN's proof (1934, 1936) of the Hardy—Ramanujan theorem on the normal order of the number of prime factors, much work has been devoted to this problem. (Even more attempt was made on the problem of convergence to a proper law; this is the main object of the books of KUBILIUS [4] and ELLIOTT [2].) A fairly general result has been obtained by Elliott [1] and MANSTAVIČIUS [5]; they solved the case

$$f_x(n) = f(n)/b_x,$$

i.e., one additive function with normalization, where the normalizing constants b_x satisfy

$$b_{x^2} = O(b_x).$$

We are going to solve the problem in complete generality.

2. Concepts and results

We shall use a probabilistic setting. An arithmetical function can be regarded as a random variable as soon as a probability measure is defined on $\mathbf{N}$. Throughout this paper this measure will always be a relative frequency measure ν_x for some x, which assigns the weight $1/x$ to each number $1, 2, \dots, x$. The symbol (f, ν_x) will be used to denote that f is regarded as a variable with respect to ν_x. For example, we shall use E to denote expectation and for arithmetical functions we shall write

$$\mathsf{E}(f, \nu_x) = x^{-1} \sum_{n\leqq x} f(n).$$

AMS (MOS) subject classifications (1980). Primary 10K20, 10H25.
Key words and phrases. Additive functions, law of large numbers, limiting distribution.

For a random variable ξ let

$$L(\xi)=\inf_{a,\varepsilon}\big(\varepsilon+\mathsf{P}(|\xi-a|>\varepsilon)\big),$$

i.e., $L(\xi)$ is the Lévy distance of ξ from the set of constants. For arithmetical functions we write $L(f, \nu_x)$.

For an additive function f put

$$U(f, x, \lambda)=\lambda^2+\sum_{p^k\le x} p^{-k}\min\big(1, (f(p^k)-\lambda\log p^k)^2\big),$$

$$U(f, x)=\min_{\lambda} U(f, x, \lambda),$$

$$V(f, x)=\min\big(1, U(f, x)\big).$$

THEOREM. *For every real-valued additive function f we have*

$$V(f, x)\ll L(f, \nu_x)\ll V(f, x)^{1/3}; \tag{2.1}$$

the implied constants are absolute.

Now we can answer our starting problem.

COROLLARY. *Let (f_x) be a sequence of real-valued additive functions. A necessary and sufficient condition for the distribution of (f_x-a_x, ν_x) to converge to the improper law with a suitable choice of the centering constants a_x is*

$$U(f_x, x)\to 0. \tag{2.2}$$

If (2.2) *holds and λ_x is a sequence of real numbers such that*

$$U(f_x, x, \lambda_x)\to 0,$$

then one may choose

$$a_x=\lambda_x\log x+\sum_{|f_x(p)-\lambda_x\log p|\le 1}\frac{1}{p}\big(f_x(p)-\lambda_x\log p\big).$$

The Corollary follows immediately from the Theorem except the choice of a_x, which is shown by the Lemma in the beginning of the next section.

3. Upper estimate

This will be very simple. It is a consequence of the lemma below, which is easily deduced from the Turán—Kubilius inequality.

LEMMA 3.1. *If f is a real-valued additive function, λ a real number and*

$$U(f, x, \lambda)=\delta,$$

then with

$$A=\lambda\log x+\sum_{|f(p)-\lambda\log p|\le 1}\frac{1}{p}\big(f(p)-\lambda\log p\big)$$

and with an absolute constant c_1 we have

$$\nu_x\big(|f(n)-A|\geqq 2\delta^{1/3}\big)\leqq c_1\delta^{1/3}. \tag{3.1}$$

PROOF. Put

$$f(n)=\lambda\log n+f_1(n)+f_2(n),$$

where f_1 and f_2 are additive functions, defined by $f_1(p^k)=f(p^k)-\lambda\log p^k$, $f_2(p^k)=0$ if $|f(p^k)-\lambda\log p^k|\leqq 1$, $f_1(p^k)=0$, $f_2(p^k)=f(p^k)-\lambda\log p^k$ otherwise.

Evidently, we have

$$\delta=\lambda^2+\sum_{p^k\leqq x}p^{-k}f_1^2(p^k)+\sum_{p^k\leqq x,\, f_2(p^k)\neq 0}p^{-k}=\lambda^2+\delta_1+\delta_2.$$

Obviously,

$$\nu_x\big(f_2(n)\neq 0\big)\leqq\delta_2\leqq\delta\leqq\delta^{1/3} \tag{3.2}$$

assuming $\delta\leqq 1$, but otherwise the lemma is trivial.

Denoting

$$A_1=\sum_{p\leqq x}f_1(p)/p=A-\lambda\log x$$

the Turán—Kubilius inequality yields

$$\mathsf{E}\big((f_1(n)-A_1)^2,\nu_x\big)\leqq c_2\delta_1$$

with an absolute constant c_2, and hence

$$\nu_x\big(|f_1(n)-A_1|\geqq\delta^{1/3}\big)\leqq c_2\delta_1\delta^{-2/3}\leqq c_2\delta^{1/3}. \tag{3.3}$$

Finally, utilizing $\lambda^2\leqq\delta$

$$\begin{aligned}\nu_x(|\lambda\log n-\lambda\log x|\geqq\delta^{1/3})&\leqq\exp(-\delta^{1/3}/|\lambda|)\\&\leqq\exp(-\delta^{-1/6})\leqq c_3\delta^{1/3}.\end{aligned} \tag{3.4}$$

Combining (3.3), (3.4) and (3.5) we get (3.2). □

4. Lower estimate

This is based on the following two lemmas.

LEMMA 4.1. *Let f be an additive function, and suppose*

$$\nu_x\big(f(n)\in[a,a+h]\big)\geqq q;$$

then there is a λ, $|\lambda|\leqq ch/q$ such that

$$\sum_{p\leqq x}\frac{1}{p}\min\big(h^2,(f(p)-\lambda\log p)^2\big)\leqq ch^2q^{-2}, \tag{4.1}$$

where c is an absolute constant.

(See RUZSA [6].)

LEMMA 4.2. *Let P be a set of primes and let $F(P, x)$ denote the number of squarefree numbers up to x which are divisible by no element of P. Suppose*

$$\sum_{p\in P} 1/p \leqq K.$$

We have

$$F(P, x) \geqq \gamma(K)x$$

with a positive constant $\gamma(K)$, depending only on K.

(See [3], Section 1, Corollary.)

PROOF OF THE LOWER ESTIMATE. Suppose that

$$\nu_x(|f(n)-a| < \delta) > 1-\delta; \tag{4.2}$$

we want to find an upper bound for $U(f, x)$.

We shall assume

$$\delta < c_1, \tag{4.3}$$

where c_1 is a small but positive absolute constant, to be specified later. As δ approaches 1, the estimate for U gets worse. (4.1) yields

$$U(f, x) \ll (1-\delta)^{-2}$$

and no essentially better estimate can be stated, that is why we use $V(f, x)$ in (2.1) instead of U.

For $\delta<1/2$ (4.1) yields

$$\sum_{p\leqq x} \frac{1}{p} \min\left(\delta^2, (f(p)-\lambda \log p)^2\right) \leqq c_2\delta^2, \tag{4.4}$$

where λ satisfies

$$|\lambda| < c_3\delta. \tag{4.5}$$

Now we regard a decomposition similar to that of the previous section. Let

$$f(n) = \lambda \log n+f_1(n)+f_2(n),$$

where

$$f_1(p^k) = f(p^k)-\lambda \log p^k, \quad f_2(p^k) = 0$$

if

$$|f(p^k)-\lambda \log p^k| \leqq \delta,$$

and

$$f_1(p^k) = 0, \quad f_2(p^k) = f(p^k)-\lambda \log p^k$$

otherwise. (4.4) means that

$$\sum f_1^2(p^k)p^{-k} \leqq c_2\delta^2, \tag{4.6}$$

$$\sum_{f_2(p)\neq 0} p^{-k} \leqq c_2, \tag{4.7}$$

the summands arising from higher powers can be estimated directly, in (4.6) utilizing $|f_1(p^k)| \leqq \delta$. (From the sums we omitted the trivial condition $p^k \leqq x$.)

Let

$$P = \{p: f_2(p) \neq 0, \; p \leqq x\}$$

and

$$P_1 = \{p^k: \; |f_2(p^k)| > \eta, \; p^k \leqq x\},$$

where $\delta < \eta < 1$, η will be chosen later. Let Q denote the set of squarefree numbers $n \leqq x$ which are divisible by no element of P, and let S be the set of numbers $n \leqq x$ of the form

$$n = p^k m, \; p^k \in P_1, \; p \nmid m, \; m \in Q. \tag{4.8}$$

Evidently, $f_2(n) = 0$ if $n \in Q$ and $|f_2(n)| \geqq \eta$ if $n \in S$.

We have

$$|Q| \geqq c_3 x$$

by (4.7) and Lemma 4.2. Put

$$\alpha = \sum_{p^k \in P_1} p^{-k}.$$

We are going to show

$$|S| \geqq c_4 \alpha x. \tag{4.9}$$

Namely, for a given $p^k \in P_1$ a "good" m must satisfy $m \leqq xp^{-k}$, it must be squarefree and must be free from prime factors from $P \cup \{p\}$. The number of such m's is

$$F(P \cup \{p\}, xp^{-k}) \geqq c_4 x p^{-k}$$

by Lemma 4.2. The numbers mp^k belonging to different $p^k \in P_1$ are different, thus we obtain

$$|S| \geqq \sum_{p^k \in P_1} c_4 x p^{-k} = c_4 \alpha x.$$

Now regard the function

$$f_3 = f - f_2 = \lambda \log + f_1.$$

If $n \in Q$ and $|f(n) - a| \leqq \delta$, then necessarily

$$|f_3(n) - a| \leqq \delta;$$

this happens for at least $(c_3 - \delta)x$ values of n ($\leqq x$). If $n \in S$ and $|f(n) - a| \leqq \delta$, then

$$|f_3(n) - a| \geqq \eta - \delta;$$

this happens for at least $(c_4 \alpha - \delta)x$ values of n.

For the variance (which we denote by D) this implies

$$\mathsf{D}(f_3, \nu_x) \geqq \frac{1}{2}(\eta - 2\delta)^2 (c_4 \alpha - \delta)(c_3 - \delta) > c_5 \eta^2 (c_4 \alpha - \delta) \tag{4.10}$$

supposing $\delta < c_3/2$, which is automatically satisfied if we choose $c_1 < c_3/2$, and also

$\eta > 3\delta$. On the other hand,

$$\mathsf{D}(f_3, \nu_x) = \mathsf{D}(\lambda \log + f_1, \nu_x) \leqq$$
$$\leqq 2(\lambda^2 \mathsf{D}(\log \nu_x) + \mathsf{D}(f_1, \nu_x)).$$

We have

$$\mathsf{D}(f_1, \nu_x) \ll \delta^2$$

by (4.6) and the Turán—Kubilius inequality and

$$\mathsf{D}(\log, \nu_x) \ll 1$$

by an easy computation. Combining these inequalities and (4.5) we obtain

(4.11) $$\mathsf{D}(f_3, \nu_x) \leqq c_6 \delta^2.$$

Now choose $\eta = \sqrt{\delta}$ (it satisfies our requirements if $\delta < 1/9$). (4.10) and (4.11) yield

$$c_5 \eta^2 (c_5 \alpha - \delta) \leqq c_6 \delta^2,$$

hence

$$\alpha \leqq c_7 \delta.$$

Therefore

$$\sum p^{-k} \min(1, f_2(p^k)^2) \leqq \sum_{|f_2(p^k)| > \eta} p^{-k} + \eta^2 \sum_{f_2(p^k) \neq 0} p^{-k} \leqq \alpha + c_2 \eta^2 \leqq c_8 \delta;$$

combined with (4.5) and (4.6) this gives

$$U(f, \lambda, x) \ll \delta. \qquad \square$$

5. Concluding remarks

(2.1) is best possible if we look for an estimate in terms of $U(f, x)$. A better result can be expected in the following more complicated form.

Let λ be a real number and let ξ_p (p prime) be independent random variables with the distribution

$$\mathsf{P}(\xi_p = f(p^k) - \lambda \log p^k) = p^{-k}(1 - 1/p).$$

Moreover, let ξ_0 be independent of the ξ_p's and with the distribution

$$\mathsf{P}(\xi_0 < u) = \begin{cases} e^u & \text{for} \quad u < 0, \\ 1 & \text{for} \quad u \geqq 0. \end{cases}$$

Finally, put

$$\eta = \eta(x, \lambda) = \lambda \xi_0 + \sum_{p \leqq x} \xi_p.$$

I think that $L(\eta)$ will be close to $L(f, \nu_x)$ for a suitably chosen value of λ; more exactly, I conjecture that

$$L(f, \nu_x) \asymp \min_{\lambda} L(\eta(x,\lambda))$$

holds. I can prove the upper estimate, the lower one seems to be considerably more difficult.

REFERENCES

[1] ELLIOTT, P. D. T. A.: The law of large numbers for additive arithmetic functions, *Math. Proc. Cambridge Phil. Soc.* **79** (1975), 33—71.
[2] ELLIOTT, P. D. T. A.: *Probabilistic number theory* I—II, Springer-Verlag, Berlin—Heidelberg—New York, 1979, 1980.
[3] ERDŐS, P. and RUZSA, I. Z.: On the small sieve I: Sifting by primes, *J. Number Theory* **12** (1980), 385—394.
[4] KUBILIUS, J.: *Probabilistic methods in the theory of numbers*, Vilnius, 1st ed. 1959, 2nd ed. 1962 (in Russian). English translation: *Amer. Math. Soc. Translations of Math. Monographs* **11** (1964).
[5] MANSTAVIČIUS, E.: The degenerate distribution for additive arithmetical functions, *Liet. Mat. Rinkinys=Litovsk. Mat. Sbornik* **16** (1976), 194—195 (in Russian).
[6] RUZSA, I. Z.: On the concentration of additive functions, *Acta Math. Acad. Sci. Hungar.* **36** (1980), 215—232.
[7] TURÁN, P.: On a theorem of Hardy and Ramanujan, *J. London Math. Soc.* **9** (1934), 274—276.
[8] TURÁN, P.: Über einige Verallgemeinerungen eines Satzes von Hardy und Ramanujan, *J. London Math. Soc.* **11** (1936), 125—133.

(Received May 5, 1980)

MTA MATEMATIKAI KUTATÓ INTÉZETE
REÁLTANODA U. 13—15
H—1053 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 255—271

THE MEAN-SQUARE DISCREPANCIES OF SOME TWO-DIMENSIONAL LATTICES

by

VERA T. SÓS and S. K. ZAREMBA

§ 1. Introduction

Let Q^2 denote the square defined by

$$0 \leqq x < 1; \quad 0 \leqq y < 1,$$

and let $\mathbf{Z}$ be any finite set of points $\mathbf{z}_0, \dots, \mathbf{z}_{m-1}$ contained in Q^2, $\mathbf{z}_i = (x_i, y_i)$ $(i = 0, \dots, m-1)$. The degree of equidistribution of $\mathbf{Z}$ can be described by the function

$$g(\mathbf{z}) = m^{-1}\nu(\mathbf{z}) - xy,$$

where $\mathbf{z} = (x, y)$ is in the closure $\bar{Q}^2$ of Q^2, and $\nu(\mathbf{z})$ is the number of points of $\mathbf{Z}$ for which $x_i < x$ and $y_i < y$.

Clearly, if the equidistribution of $\mathbf{Z}$ is good, $|g(\mathbf{z})|$ should be small throughout $\bar{Q}^2$. If we want a single number to measure the equidistribution in question, the obvious choice is a norm of $g(\mathbf{z})$. The two most natural norms are

(i) $$D^*(\mathbf{Z}) = \sup_{\mathbf{z} \in \bar{Q}^2} |g(\mathbf{z})|$$

and

(ii) $$D^{(2)}(\mathbf{Z}) = \Big(\int_{Q^2} g(\mathbf{z})^2 \, d\mathbf{z}\Big)^{1/2}.$$

The first norm is known as the *extreme discrepancy* of X, or, more simply, its discrepancy. For the second, the name of L^2 discrepancy, or *mean-square discrepancy* was introduced in 1968 Zaremba [10], although its concept appeared as early as 1954 Roth [7]. The definitions of $D^*(\mathbf{Z})$ and $D^2(\mathbf{Z})$ can be extended in an obvious manner to any number of dimensions; however, the present paper deals only with the case of two dimensions.

The two concepts of discrepancy, apart from their intrinsic number-theoretical interest, play an important part in numerical analysis: If we regard the expressions

$$m^{-1}\big(f(\mathbf{z}_0) + \dots + f(\mathbf{z}_{m-1})\big)$$

as approximate values of the integral

$$\int_{Q^2} f(\mathbf{z}) \, d\mathbf{z},$$

the absolute values of the errors have, under suitable conditions of smoothness imposed on f, upper bounds of the form

$$C^* D^*(\mathbf{Z}) \quad \text{or} \quad C^{(2)} D^{(2)}(\mathbf{Z}),$$

where the coefficients C^* and $C^{(2)}$ depend only on f (see, for instance HLAWKA [3], ZAREMBA [10] or [11] or KOROBOV [5]). If $\mathbf{Z}$ is a suitable lattice, then, depending on the smoothness of f, the error of integration can be of a much smaller order of magnitude than the bounds indicated above (see, for instance HLAWKA [4], KOROBOV [5], ZAREMBA [11] or VILENKIN [9]).

K. F. ROTH [7] proved that

$$D^{(2)}(\mathbf{Z}) \geqq c_2 m^{-1}(\log m)^{1/2}$$

for every finite set $\mathbf{Z}=\{\mathbf{z}_0, \dots, \mathbf{z}_{m-1}\}\subset Q^2$, where c_2 is an absolute constant. W. SCHMIDT [8] proved that for any such set $\mathbf{Z}$

$$D^*(\mathbf{Z}) \geqq cm^{-1}\log m,$$

where c is again an absolute constant.

Sequences of sets $\mathbf{Z}\subset Q^2$ for which

$$D^*(\mathbf{Z}) = O(m^{-1}\log m), \tag{1.1}$$

in particular sequences of such *lattices* $\mathbf{Z}$ are well-known (see, for instance HLAWKA [4], KOROBOV [5], ZAREMBA [11] or VILENKIN [9]). Sequences of sets $\mathbf{Z}\subset Q^2$ for which

$$D^2(\mathbf{Z}) = O\big(m^{-1}(\log m)^{1/2}\big) \tag{1.2}$$

have also been known (DAVENPORT [1], HALTON—ZAREMBA [2], VILENKIN [9]). But none of these sets formed a lattice, although the one considered by DAVENPORT [1] was a symmetric union of two lattices. In view of the theoretical and practical importance of lattices, it was felt that it was worth investigating which lattices $\mathbf{Z}$, if any, had an L^2 discrepancy of the order of $m^{-1}(\log m)^{1/2}$.

At this stage it should be recalled that if A is an upper bound of the partial quotients of the finite or infinite continued-fraction expansion of a number α, if m does not exceed the denominator of α in the case when α is rational, and if $\mathbf{Z}$ consists of the points

$$\langle 0,0\rangle,\ \langle m^{-1}, \{\alpha\}\rangle,\ \langle 2m^{-1}, \{2\alpha\}\rangle,\ \dots,\ \langle (m-1)m^{-1}, \{(m-1)\alpha\}\rangle,$$

$\{x\}$ denoting the fractional part of x, then

$$D^*(\mathbf{Z}) \leqq K^* m^{-1}\log m, \tag{1.3}$$

where K^* is a constant depending only on A; this is an immediate consequence of Proposition 4.3 in Zaremba [12].

The main purpose of the present paper is to show that if all the partial quotients of the finite or infinite continued fraction expansion of α are equal, $m\geqq 1$ not exceeding the denominator of α when α is rational, then

$$D^{(2)}(\mathbf{Z}) = O\big(m^{-1}(\log m)^{1/2}\big).$$

We obtain this result by examining the expressions

$$\frac{1}{m}\sum_{q=0}^{m-1} S_q^2, \tag{1.4}$$

where

$$S_q = \sum_{j=1}^{q}\left(\{j\alpha\}-\frac{1}{2}\right). \tag{1.5}$$

Propositions about the behaviour of (1.4) and its connection with $D^2(\mathbf{Z})$ may also be of some intrinsic interest.

In a forthcoming paper we are going to prove that if the partial quotients of the continued fraction are not all equal, even if they are bounded, $D^{(2)}(\mathbf{Z})$ can be of the order of $m^{-1}\log m$.

§ 2. A crucial lemma

LEMMA 2.1. *With the previously introduced notations, assuming that the partial quotients of the continued-fraction expansion of α are bounded, and that m does not exceed the denominator of α in the case when α is rational,*

$$D^2(\mathbf{Z}) = O\big(m^{-1}(\log m)^{1/2}\big)$$

if, and only if

$$\frac{1}{m}\sum_{q=0}^{m-1} S_q^2 = O(\log m),$$

where S_q is given by (1.5).

PROOF. We use a technique due to H. DAVENPORT [1]. To simplify some notations, we put

$$G(x, y) = mg(x, y) = v(x, y) - mxy$$

and

$$\psi(\eta) = \{\eta\} - \frac{1}{2}.$$

It is easily verified that for any β and any η in [0, 1]

$$\eta + \psi(\beta-\eta) - \psi(\beta) = \begin{cases} 0 & \text{if } \{\beta\} \geqq \eta, \\ 1 & \text{if } \{\beta\} < \eta. \end{cases}$$

Hence

$$v(x,y) = \sum_{0 \leqq j < mx} \big(y + \psi(j\alpha - y) - \psi(j\alpha)\big).$$

Clearly

$$|G(x, y) - \tilde{G}(x, y)| \leqq 1,$$

where

$$\tilde{G}(x, y) = \sum_{0 \leqq j < mx} \big(\psi(j\alpha - y) - \psi(j\alpha)\big).$$

Since it is well-known (ROTH [7]), that

$$\int_0^1\int_0^1 G(x, y)^2\, dx\, dy$$

is at least of the order of log m, the order of magnitude of the last integral is the same as that of

$$\int_0^1\int_0^1 \tilde{G}(x, y)^2\, dx\, dy.$$

Now we take advantage of the Fourier expansion

$$\psi(\alpha) = -\frac{1}{\pi}\sum_{n=0}^{\infty}\frac{\sin 2\pi n\alpha}{n}$$

valid for $\alpha\neq 0$. With this representation,

$$\tilde{G}(x, y) = \sum_{0\leq j<mx}\left(-\frac{1}{\pi}\sum_{n=1}^{\infty}\frac{\sin 2\pi n(\alpha j-y)}{n}-\psi(j\alpha)\right) =$$

$$(2.1)\qquad = -\frac{1}{\pi}\sum_{n=1}^{\infty}k^{-1}\cos 2\pi ny\sum_{0\leq j<mx}\sin 2\pi n\alpha j +$$

$$+\frac{1}{\pi}\sum_{n=1}^{\infty}k^{-1}\sin 2\pi ny\sum_{0\leq j<mx}\cos 2\pi n\alpha j-\sum_{0\leq j<mx}\psi(j\alpha).$$

Now we want to square this expression and integrate it with respect to y from 0 to 1. The three terms of the integrand being orthogonal to each other, the integral of $\tilde{G}(x, y)^2$ is equal to the sums of the integrals of the three squared terms. We shall denote these integrals by I_1, I_2 and I_3, respectively.

We begin with I_2. By the Parseval formula

$$(2.2)\qquad I_2 = \frac{1}{2\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2}\Big(\sum_{0\leq j<mx}\cos 2\pi n\alpha j\Big)^2.$$

Now we have to distinguish the cases of α being irrational and of α being rational. We begin with the former case, following DAVENPORT [1].

It is well-known (see, e.g., Lemma 6.5 in Zaremba [11]) that if $n\alpha$ is not an integer, then for any m

$$(2.3)\qquad \Big|\sum_{0\leq j<mx}\cos(2\pi n\alpha j)\Big| \leq \frac{1}{2\,\|n\alpha\|}$$

where $\|\xi\|$ denotes the distance of ξ from the nearest integer.

But also

$$\Big|\sum_{0\leq j<mx}\cos(2\pi in\alpha j)\Big| \leq [mx].$$

Thus

$$(2.4)\qquad I_2 \leq \frac{1}{2\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2}\min([mx]^2, 2^{-2}\|n\alpha\|^{-2}).$$

Let p_k/q_k be the successive convergents of the continued-fraction expansion of α, defined by $q_1=1$, $q_2=a_1$, ..., $q_{k+1}=a_kq_k+q_{k-1}$ ($k=2, 3, \ldots$), $p_1=0$, $p_2=1$, ..., $p_{k+1}=$

$=a_kp_k+p_{k-1}$, where $a_1, a_2, \ldots$ are the partial quotients in this expansion. If $q_{k-1}\leqq$ $\leqq n\leqq q_k$, by Lagrange's theorem,

$$\|n\alpha\| \geqq |q_{k-1}\alpha - p_{k-1}| \geqq (q_{k-1}+q_k)^{-1}.$$

Hence

$$n\|n\alpha\| \geqq q_{k-1}/(q_{k-1}+q_k) > (A+2)^{-1} = C, \tag{2.5}$$

where $A=\max\limits_i a_i$.

If $2^{r-1}\leqq n<2^r$, then by (2.5),

$$\|n\alpha\| > C/n > C/2^r.$$

But, for any given integer s, there can be at most two values of n, say n_1 and n_2 in $[2^{r-1}, 2^r)$ satisfying

$$sC\cdot 2^{-r} \leqq \|n_i\alpha\| < (s+1)C\cdot 2^{-r} \qquad (i=1,2). \tag{2.6}$$

Indeed, if there were a third one, we would have an n^* with $|n^*|<2^r$ and

$$\|n^*\alpha\| < C\cdot 2^{-r} < C/|n^*|,$$

which contradicts (2.5).

The two values of n in $[2^{r-1}, 2^r)$ satisfy

$$\|n\alpha\|^{-2} < C^{-2}s^{-2}2^r,$$

and according to (2.4), we find

$$\begin{aligned} I_2 &\leqq \frac{1}{\pi^2}\sum_{r=1}^{\infty} 2^{2-2r}\sum_{s=1}^{\infty}\min\left([mx]^2, C^{-2}s^{-2}2^{2r}\right) \leqq \\ &\leqq \frac{4}{\pi^2C^2}\sum_{r=1}^{[\log_2 m]}\sum_{1}^{\infty}\frac{1}{s^2}+\frac{1}{\pi^2}\sum_{r>[\log_2 m]}\sum_{s=1}^{\infty}[mx]^{\frac{1}{2}}C^{-\frac{3}{2}}2^{-\frac{r}{2}}s^{-\frac{3}{2}}. \end{aligned} \tag{2.7}$$

Since the first sum is $O(\log m)$ and the second is $O(1)$, we obtain

$$I_2 = O(\log m). \tag{2.8}$$

If α is rational, we denote by d its denominator, and we put

$$n = kd+l \quad \text{with} \quad 0\leqq l<d.$$

We have to single out the terms of (2.2) which correspond to values of n being multiples of d. The sum of these terms does not exceed

$$\frac{1}{2\pi^2}\sum_{k=1}^{\infty}\frac{[mx]^2}{k^2d^2} \leqq \frac{1}{2\pi^2}\sum_{k=1}^{\infty}\frac{1}{k^2} = \frac{1}{12}.$$

Concerning the other terms of (2.2), we note that $\|n\alpha\|=\|l\alpha\|$, while, as in the case of α irrational, $l\|l\alpha\|>C$, and all the more $n\|n\alpha\|>C$, or

$$\|n\alpha\| > \frac{C}{n}.$$

Argueing as in the case of an irrational α, we find that the sum of the terms of (2.2) which correspond to values of n other than multiples of d is smaller than the right-hand side of (2.7), and therefore is $O(\log m)$. Thus (2.8) holds both for rational and irrational values of α.

Concerning I_1, instead of (2.2) we have

$$I_1 = \frac{1}{2\pi} \sum_{k=1}^{\infty} \frac{1}{k^2} \sum_{0 \leqq j < mx} (\sin 2\pi k\alpha j)^2.$$

The treatment is exactly the same as that of I_2, the only difference being that in the case of α rational, all the terms corresponding to values of k divisible by d vanish. Thus in both cases

$$I_1 = O(\log m). \tag{2.9}$$

Both I_1 and I_2 have to be integrated with respect to x in $[0, 1]$. Since the upper bounds obtained for them do not depend on x, the double integrals are also $O(\log m)$.

I_3 is quite different. Since the square of the last term in the right-hand side of (2.1) is independent of y, it is equal to I_3. Now it has to be integrated with respect to x in $[0, 1]$; since it is a step function, in view of the definition of ψ, its integral reduces to the sum (1.4).

Thus, apart from a term which in any event is of a lower order of magnitude,

$$D^{(2)}(\mathbf{Z})^2 = \int_0^1 \int_0^1 G(x, y)^2 \, dx \, dy$$

is the sum of two terms which were found to be $O(\log m)$ and of the sum (1.4). This proves the lemma.

§ 3. Further lemmas about continued fractions

Let α be fixed and a_k, q_k, p_k $(k=1, \dots)$ have the same meaning as before.

Definition 3.1. A finite sequence $\langle b_r, \dots, b_s \rangle$, with $s<n'$ if $\alpha = p_{n'}/q_{n'}$ will be described as *admissible* (with respect to α) when

$$0 \leqq b_r < a_r \quad \text{and} \quad 0 \leqq b_i \leqq a_i,$$
$$\text{but} \quad b_{i-1} = 0 \quad \text{whenever} \quad b_i = a_i \quad (i = r+1, \dots, s).$$

The two lemmas and two corollaries which follow are well-known (they were exactly implied in Ostrowski [6]) and in any event are easy to prove.

Lemma 3.2. *If $\langle b_1, \dots, b_{n-1} \rangle$ is an admissible sequence,*

$$b_1 q_1 + \dots + b_{n-1} q_{n-1} < q_n.$$

Lemma 3.3. *Assuming that $n \leqq n'$ if $\alpha = p_{n'}/q_{n'}$, any nonnegative integer $p < q_n$ can be uniquely represented in the form*

$$q = b_1 q_1 + \dots + b_{n-1} q_{n-1} \tag{3.1}$$

where $\langle b_1, \dots, b_n \rangle$ is an admissible sequence.

COROLLARY 3.4. Assuming that $n \leqq n'$ if $\alpha = p_{n'}/q_{n'}$ there is a one-to-one correspondence between integers $0, 1, \ldots, q_n - 1$ and admissible sequences $\langle b_1, \ldots, b_{n-1} \rangle$ determined by (3.1).

COROLLARY 3.5. Under the same assumptions, the number of admissible sequences $\langle b_1, \ldots, b_{n-1} \rangle$ is equal to q_n.

It is well-known from the theory of continued fractions that for any $i \geqq 1$

$$\alpha = \frac{p_i}{q_i} + \frac{\Theta_i}{q_i q_{i+1}} \tag{3.2}$$

where $|\Theta_i| \leqq 1$. We consider now the various sums

$$S_q = \sum_{j=0}^{q} \left(\{j\alpha\} - \frac{1}{2} \right)$$

with $q < p_{n'}$ if $\alpha = p_{n'}/q_{n'}$. According to Lemma 3.3, q admits a unique representation (3.1) where $\langle b_i \rangle$ is an admissible sequence. Hence S_q can be represented uniquely in the form

$$S_q = \sum_{i=1}^{n-1} \sigma_i \tag{3.3}$$

where

$$\sigma_i = \sum_{\nu=0}^{b_i q_i - 1} \left(\{(\nu + t_i)\alpha\} - \frac{1}{2} \right)$$

when $b_i > 0$, and $\sigma_i = 0$ when $b_i = 0$, while

$$t_i = \sum_{k=1}^{i-1} b_k q_k \qquad (i = 2, \ldots, n-1); \; t_1 = 0.$$

According to OSTROWSKI [6] we have

$$\sigma_i = b_i \left(\frac{(-1)^i}{2} + \Theta_i \frac{b_i q_i + 2 \sum_{k=1}^{i-1} t_k q_k - 1}{2 q_{i+1}} \right), \quad i = 1, \ldots, k-1. \tag{3.4}$$

We consider now the special case when the number α has a finite or infinite continued-fraction expansion whose all partial quotients are equal to a positive integer a. The convergents of the expansion of α are easily found to be

$$p_i/q_i = v_{i-1}/v_i$$

where

$$v_j = (\beta^j + (-1)^{j+1} \beta^{-j})(a^2 + 4)^{-1/2} \qquad (j = 1, 2, \ldots) \tag{3.5}$$

and

$$\beta = \frac{1}{2}(a + (a^2 + 4)^{1/2}).$$

If $a = 1$, the sequence is that of the Fibonacci numbers, and (3.5) is nothing else but the Binet formula. Thus either

$$\alpha = v_{n'-1}/v_{n'} \tag{3.6}$$

for some integer n', or

$$\alpha = \lim_{n\to\infty} v_{n-1}/v_n = \beta^{-1} = \frac{1}{2}((a^2+4)^{1/2}-a). \tag{3.7}$$

In our case, (3.2) becomes

$$\alpha = \frac{v_{i-1}}{v_i} + \frac{\Theta_i}{v_i v_{i+1}} \tag{3.8}$$

with $i<n$ if α is given by (3.6).

LEMMA 3.6. *If α is given by* (3.6) *or* (3.7), *the sums $\Theta_r+\ldots+\Theta_s$ with $1\leqq r<s$ and $s<n'$ in the former case are bounded by a number depending only on a.*

PROOF. Since $|\Theta_i|\leqq 1$ for all i, it suffices to show that the sums

$$\sum_{k=1}^{\infty} |\Theta_{2k}+\Theta_{2k+1}|,$$

when α is irrational, and

$$\sum_{k=1}^{\left[\frac{n'-2}{2}\right]} |\Theta_{2k}+\Theta_{2k+1}|$$

when α is rational have an upper bound depending only on a.

By (3.2) with $q_i=v_i$, $q_{i+1}=v_{i+1}$ and $p_i=v_{i-1}$

$$\Theta_i = v_{i+1}(\alpha v_i - v_{i-1}), \tag{3.9}$$

and if α is irrational, we find

$$\Theta_i = \frac{(-1)^{i+1}}{a^2+4}(1+\beta^2+(-1)^i\beta^{-2i-2}+(-1)^i\beta^{-2i}).$$

Similarly

$$\Theta_{i+1} = \frac{(-1)^i}{a^2+4}(1+\beta^2+(-1)^{i+1}\beta^{-2i-4}+(-1)^{i+1}\beta^{-2i-2}).$$

Hence

$$|\Theta_i+\Theta_{i+1}| = \frac{(1+\beta^{-2})^2}{a^2+4}\beta^{-2i},$$

and further

$$\sum_{i=1}^{\infty} |\Theta_{2i}+\Theta_{2i+1}| = \frac{(1+\beta^{-2})^2}{a^2+4}\,\frac{1}{\beta^4-1}.$$

The case when α is rational, i.e., is given by (3.2), is slightly more complicated. Substituting (3.5) and (3.6) in (3.9), we find

$$\Theta_i =$$

$$= \frac{(\beta^{i+1}+(-1)^i\beta^{-i-1})((-1)^{n'}\beta^{i-n'+1}+(-1)^{n'}\beta^{i-n'-1}+(-1)^{i+1}\beta^{n'-i-1}+(-1)^{i+1}\beta^{n'-i+1})}{(a^2+4)^{3/2}v_{n'}}$$

and a similar expression for Θ_{i+1}, with $i+1$ substituted for i everywhere. After some simplifications, we obtain

$$|\Theta_i+\Theta_{i+1}| = \frac{(1+\beta^2)^2\beta^{2i-n'}+(1+\beta^{-2})\beta^{n'-2i}}{(a^2+4)^{3/2}v_{n'}}.$$

Since $(a^2+4)^{3/2}v_{n'}/\beta^{n'}$ is bounded, to prove that $\Theta_r+\ldots+\Theta_s$ is bounded, it suffices to show the boundedness of

$$\sum_{\nu=1}^{\left[\frac{n'-2}{2}\right]}((1+\beta^2)^2\beta^{4\nu-2n'}+(1+\beta^{-2})\beta^{-4\nu}),$$

which is trivial.

§ 4. A probabilistic interpretation of the problem

The sum (1.4) with $m=q_n$, i.e., the sum

$$\frac{1}{q_n}\sum_{q=1}^{q_n-1}S_q^2,$$

where $n<n'$ if $\alpha=p_{n'}/q_{n'}$, can be regarded as the expectation $\mathsf{E}(S_q^2)$, q being a random variable taking each of the values $0, 1, \ldots, q_n-1$ with the same probability $\frac{1}{q_n}$. To compute

$$\mathsf{E}(S_q^2) = \mathsf{E}(S_q)^2+\operatorname{var} S_q,$$

we need the first and second order moments of the joint probability distribution of $\sigma_1, \ldots, \sigma_{n-1}$. Owing to (3.4), this will be deduced from the relevant moments of $b_1, \ldots, b_{n-1}$.

We consider the case $a_i=a$, $i=1, 2, \ldots$. We begin with the probability $\mathsf{P}[b_i=k]$ with $0<k<a_j$; it is, of course, equal to 0 when $a=1$. If $a>1$, $b_1, \ldots, b_{i-1}$ can form any admissible sequence and according to Corollary (3.5) there are v_i such sequences. Independently of them, $b_{i+1}, \ldots, b_{n-1}$ can be any admissible sequence, which gives v_{n-i} possibilities. Thus the total number of admissible sequences featuring $b_i=k$ is v_iv_{n-i}; since $v_1=1$, this is true, in particular for $i=1$ and for $i=n-1$. Since each of the sequences in question has probability $1/v_n$, we have

$$\mathsf{P}(b_i = k) = v_iv_{n-i}/v_n \quad 0<k<a; \quad i = 1, \ldots, n-1. \tag{4.1}$$

If $b_i=a$ with $i>1$, we must have $b_{i-1}=0$. Hence the factor v_i in (4.1) has to be replaced by v_{i-1} and so

$$\mathsf{P}(b_i = a) = v_{i-1}v_{n-i}/v_n \quad (i = 2, \ldots, n-1) \tag{4.2}$$

while necessarily $\mathsf{P}(b_1=a)=0$.

From (4.1) and (4.2) we obtain

$$\mathsf{E}(b_i) = \frac{v_{n-i}}{v_n}\left(\frac{a(a-1)}{2}v_i+av_{i-1}\right). \tag{4.3}$$

This is still valid for $i=1$ since $v_0=0$. Substituting here, for v_i, v_{i-1}, v_{n-i} and v_n their expressions in terms of β, we find

$$\mathsf{E}(b_i) = A+O(\beta^{-2i})+O(\beta^{2i-2n}), \tag{4.4}$$

where

$$A = \frac{a(a-1)+2a\beta^{-1}}{2(a^2+4)^{1/2}}.$$

Similarly, we deduce from (4.1) and (4.2)

$$\mathsf{E}(b_i^2) = \frac{v_{n-1}}{v_n}\left(\frac{2a^3-3a^2+a}{6}v_i+a^2v_{i-1}\right),$$

and eventually, in terms of β,

$$\mathsf{E}(b_i^2) = B+O(\beta^{-2i})+O(\beta^{2i-2n}), \tag{4.5}$$

where

$$B = \frac{2a^3-3a^2+a+6a^2\beta^{-1}}{6(a^2+4)^{1/2}}.$$

In view of (3.4), we also need $\mathsf{E}(b_hb_i)$. Assuming $h<i$, and argueing as before, we find

$$\mathsf{P}(b_h = k,\ b_i = l) = v_hv_{i-h}v_{n-i}/v_n \qquad 0<k<a;\quad 0<l<a, \tag{4.6}$$

and similarly

$$\mathsf{P}(b_h = a,\ b_i = l) = v_{h-1}v_{i-h}v_{n-i}/v_n \qquad (0<l<a) \tag{4.7}$$

$$\mathsf{P}(b_h = k,\ b_i = a) = v_hv_{i-h-1}v_{n-i}/v_n \qquad (0<k<a) \tag{4.8}$$

$$\mathsf{P}(b_h = b_i = a) = v_{h-1}v_{i-h-1}v_{n-i}/v_n. \tag{4.9}$$

Consequently,

$$\begin{aligned}\mathsf{E}(b_hb_i) = {}& \frac{a^2(a-1)^2v_hv_{i-h}v_{n-i}}{4v_n}+\\ &+\frac{a^2(a-1)(v_hv_{i-h-1}+v_{h-1}v_{i-h})v_{n-i}}{2v_n}+\frac{a^2v_{h-1}v_{i-h-1}v_{n-i}}{v_n}\end{aligned} \tag{4.10}$$

and eventually, in terms of β,

$$\mathsf{E}(b_hb_i) = A^2+(-1)^{i-h+1}C\beta^{2h-2i}+O(\beta^{-2h})+O(\beta^{2i-2n}), \tag{4.11}$$

where

$$C = \frac{a^2((a-1)^2-2(a-1)(\beta-\beta^{-1})-4)}{4a^2+4}.$$

Now we can prove the following proposition:

LEMMA 4.1. *With the previous notations, $\mathsf{E}(S_q)$ has an upper bound which depends only on a.*

PROOF. According to (3.3) and (3.4)

$$S_q = \sigma^{(1)} + \sigma^{(2)} + \sigma^{(3)} + \sigma^{(4)},$$

where

$$\sigma^{(1)} = \sum_{i=1}^{n-1} \frac{(-1)^i}{2} b_i; \qquad \sigma^{(2)} = \sum_{i=1}^{n-1} \frac{\Theta_i v_i b_i^2}{2v_{i+1}};$$

$$\sigma^{(3)} = \sum_{i=1}^{n-1} \frac{\Theta_i}{v_{i+1}} \sum_{h=1}^{i-1} v_h b_h b_i; \quad \sigma^{(4)} = \sum_{i=1}^{n-1} \frac{\Theta_i b_i}{2v_{i+1}}.$$

Now, according to (4.4),

$$\mathsf{E}(\sigma^{(1)}) = A \sum_{i=1}^{n-1} \frac{(-1)^i}{2} + \sum_{i=1}^{n-1} O(\beta^{-2i}) + \sum_{i=1}^{n-1} O(\beta^{2i-2n}),$$

which is obviously bounded.

Concerning $\sigma^{(2)}$, we observe that

$$\frac{v_i}{v_{i+1}} = \frac{\beta^i + (-1)^{i+1}\beta^{-i}}{\beta^{i+1} + (-1)^i \beta^{-i-1}} = \beta^{-1} + O(\beta^{-2i}).$$

Consequently, since $b_i^2 \leqq a^2$ and $|\Theta_i| \leqq 1$, we find

$$\mathsf{E}(\sigma^{(2)}) = \frac{1}{2\beta} \sum_{i=1}^{n-1} \Theta_i \mathsf{E}(b_i^2) + \sum_{i=1}^{n-1} O(\beta^{-2i}),$$

and further, by (4.5),

$$\mathsf{E}(\sigma^{(2)}) = \frac{B}{2\beta} \sum_{i=1}^{n-1} \Theta_i + \sum_{i=1}^{n-1} O(\beta^{-2i}) + \sum_{i=1}^{n-1} O(\beta^{2i-2n}).$$

Here, the last two terms are immediately seen to be bounded, and so is the first by Lemma 3.6. Thus $\mathsf{E}(\sigma^{(2)})$ is bounded.

Passing to $\sigma^{(3)}$, we observe that, with $h < i$,

$$\frac{v_h}{v_{i+1}} = \frac{\beta^h + (-1)^{h+1}\beta^{-h}}{\beta^{i+1} + (-1)^i \beta^{-i-1}} = \beta^{h-i-1} + O(\beta^{-h-i}).$$

Consequently

$$\mathsf{E}(\sigma^{(3)}) = \sum_{i=1}^{n-1} \Theta_i \sum_{h=1}^{i-1} \beta^{h-i-1} \mathsf{E}(b_h b_i) + \sum_{i=1}^{n-1} \Theta_i \sum_{h=1}^{i-1} \mathsf{E}(b_h b_i) O(\beta^{-h-i}).$$

Here, Θ_i and $\mathsf{E}(b_h b_i)$ being bounded, the second term is easily seen to be bounded. Substituting (4.11) in the first term, we find

$$A^2 \sum_{i=1}^{n-1} \Theta_i \sum_{h=1}^{i-1} \beta^{h-i-1} + C \sum_{i=1}^{n-1} \Theta_i \sum_{h=1}^{i-1} (-1)^{i-h+1} \beta^{3h-3i-1} +$$

$$+ \sum_{i=1}^{n-1} \Theta_i \sum_{h=1}^{i-1} O(\beta^{-h-i-1}) + \sum_{i=1}^{n-1} \Theta_i \sum_{n=1}^{i-1} O(\beta^{h+i-2n}).$$

The first term of this expression is easily seen to be bounded in view of Lemma 3.6. A similar argument shows that the second term is also bounded. Since Θ_i is bounded, there is no difficulty over the boundedness of the last two terms.

The boundedness of $\mathsf{E}(\sigma^{(4)})$ is trivial, since $\sigma^{(4)}$ itself is bounded.

§ 5. The variance of S_q

Throughout this section we assume that all the partial quotients in the continued fraction expansion of α are equal to an integer a. Some variances and covariances have to be computed before attacking $\operatorname{var} S_q$. There would be no difficulty in writing down an exact expression for $\operatorname{var} b_i$ on the basis of (4.4) and (4.5). However, it suffices for our purpose to note that $\operatorname{var} b_i$ is obviously bounded, say

$$\operatorname{var} b_i \leqq V \qquad (i=1, \dots, n-1). \tag{5.1}$$

Similarly, it suffices to know that for some W

$$\operatorname{var} b_i^2 \leqq W \qquad (i=1, \dots, n-1). \tag{5.2}$$

We need to know more about $\operatorname{cov}(b_h, b_i) = \mathsf{E}(b_h b_i) - \mathsf{E}(b_h)\mathsf{E}(b_i)$. We can rewrite (4.10) in the following form:

$$\mathsf{E}(b_h b_i) = a^2\left(v_h \frac{a-1}{2} + v_{h-1}\right)\left(v_{i-h}\frac{a-1}{2} + v_{i-h-1}\right) v_{n-i}/v_n. \tag{5.3}$$

In view of (4.3), we have therefore

$$\operatorname{cov}(b_h, b_i) = a^2 \frac{\left(v_h \frac{a-1}{2} + v_{h-1}\right) v_{n-i}}{v_n} \times \tag{5.4}$$
$$\times \frac{\left(v_{i-h}\frac{a-1}{2} + v_{i-h-1}\right) v_n - v_{n-h}\left(v_i \frac{a-1}{2} + v_{i-1}\right)}{v_n}.$$

It is easily seen that, here, the first fraction is $O(\beta^{h-i})$. In the numerator of the second fraction, if we express it in term of β, we find a linear combination of β^{h-i+n}, $\beta^{h-i+n+1}$, β^{i-h-n}, β^{n-i}, $\beta^{i-h-n-1}$, $\beta^{h-i-n+1}$, β^{n-i-h}, $\beta^{n-i-h+1}$, β^{i+h-n}, β^{h-i-n}, $\beta^{h+i-n-1}$, and $\beta^{h-i-n+1}$. Taking into account the denominator v_n, which is exactly of the order of β^n, it can be seen that

$$\operatorname{cov}(b_h, b_i) = O(\beta^{2h-2i}) \quad \text{when} \quad h < i.$$

In fact

$$|\operatorname{cov}(b_h, b_i)| \leqq C\beta^{2h-2i} \quad \text{when} \quad h < i, \tag{5.5}$$

where C has the same value as in (4.11), but the precise value of this coefficient is irrelevant from our viewpoint.

In an exactly similar way, we evaluate $\operatorname{cov}(b_h^2, b_i^2)$, obtaining, for a D depending only on a,

$$|\operatorname{cov}(b_h^2, b_i^2)| \leqq D\beta^{2h-2i} \quad \text{when} \quad h < i. \tag{5.6}$$

Now we proceed to compute $\mathsf{E}(b_h b_i b_k b_j)$ for $0<h<i\leqq k<j\leqq n$. We have for K, L, R and S in $(0, a)$

$$\mathsf{P}(b_h = K,\ b_i = L,\ b_k = R,\ b_j = S) = v_h v_{i-h} v_{k-i} v_{j-k} v_{n-j}/v_n,$$

$$\mathsf{P}(b_h = K,\ b_i = L,\ b_k = R,\ b_j = a) = v_h v_{i-h} v_{k-i} v_{j-k-1} v_{n-j}/v_n,$$

and soon. Eventually we find

$$\mathsf{E}(b_h b_i b_k b_j) = a^4 \left(v_h \frac{a-1}{2} + v_{h-1}\right)\left(v_{i-h}\frac{a-1}{2} + v_{i-h-1}\right)\cdot$$

$$\cdot\left(v_{k-i}\frac{a-1}{2} + v_{k-i-1}\right)\left(v_{j-k}\frac{a-1}{2} + v_{j-k-1}\right)v_{n-j}/v_n,$$

and, in view of (5.3),

$$\operatorname{cov}(b_h b_i, b_k b_j) =$$

$$= a^3\left(v_h\frac{a-1}{2} + v_{h-1}\right)\left(v_{i-h}\frac{a-1}{2} + v_{i-h-1}\right)\left(v_{j-k}\frac{a-1}{2} + v_{j-k-1}\right)v_{n-j}v_n^{-1}\times$$

$$\times a\left(\left(v_{k-i}\frac{a-1}{2} + v_{k-i-1}\right)v_n - v_{n-i}\left(v_k\frac{a-1}{2} + v_{k-1}\right)\right)v_n^{-1}.$$

The first line above is easily seen to be $O(\beta^{i-k})$. In the second line, if we express the v's in terms of β, we find, after crucial simplifications, a linear combination of β^{i-k}, β^{k-i-2n}, β^{-i-k}, β^{k+i-2n} and β^{i-k-2n}. Under our assumption, $i-k$ is the biggest exponent of β; hence the second line is also $O(\beta^{i-k})$. Thus there exists a constant M depending only on a such that

$$\operatorname{cov}(b_h b_i, b_k b_j) \leqq M\beta^{2(i-k)} \quad \text{when} \quad 0 < h < i \leqq k < j < n. \tag{5.7}$$

Obviously, there exists also a number N such that

$$\operatorname{cov}(b_h b_i, b_k b_j) \leqq N \quad \text{for} \quad h, i, k, j \quad \text{between } 0 \text{ and } n. \tag{5.8}$$

LEMMA 5.1. *With our previous notations*

$$\operatorname{var} S_q = O(n) = O(\log m),$$

where $m=v_n$.

PROOF. Owing to Schwarz inequality, we only need to show that

$$\operatorname{var}\sigma^{(k)} = O(n) \qquad (k = 1, 2, 3, 4).$$

According to (5.1) and (5.5),

$$\operatorname{var}\sigma^{(1)} \leqq \frac{n-1}{4}V+\frac{C}{2}\sum_{i=2}^{n-1}\sum_{h=1}^{i-1}\beta^{2h-2i} < \frac{n-1}{4}V+\frac{C(n-2)}{2(\beta^2-1)} = O(n).$$

Since $\beta v_i/(2v_{i+1})$ is bounded in view of (5.6), $\sigma^{(2)}$ can be treated exactly like $\sigma^{(1)}$, yielding

$$\operatorname{var}\sigma^{(2)} = O(n).$$

Now

$$\sigma^{(3)} = \sum_{i=1}^{n-1} Q_i$$

where

$$Q_i = \Theta_i \sum_{h=1}^{i-1} v_h v_{i+1}^{-1} b_h b_i.$$

Assuming $i \leqq j$, we have, according to (5.7) and (5.8),

$$|\operatorname{cov}(Q_i, Q_j)| \leqq \frac{1}{v_{i+1}v_{j+1}} \sum_{h=1}^{i-1}\sum_{k=1}^{j-1} v_h v_k |\operatorname{cov}(b_h b_i, b_k b_j)| \leqq$$

$$\leqq \frac{M}{v_{i+1}v_{j+1}} \sum_{h=1}^{i-1} v_h \sum_{k=1}^{j-1} v_k \beta^{2i-2k} + \frac{N}{v_{i+1}v_{j+1}} \sum_{h=1}^{i-1} v_h \sum_{k=1}^{i-1} v_k,$$

it being understood that when $i=j$, 0 should be substituted for the first term of the last expression. When $i<j$, this term is of order of

$$\beta^{-i-j}\sum_{h=1}^{i-1}\beta^h \sum_{k=i}^{j-1}\beta^{2i-k} = \beta^{i-j}\sum_{h=1}^{i-1}\beta^h\sum_{k=1}^{j-1}\beta^{-k} = O(\beta^{i-j}).$$

The second term is of the order of

$$\beta^{-i-j}\sum_{h=1}^{i-1}\beta^h\sum_{k=1}^{i-1}\beta^k = O(\beta^{i-j}).$$

Thus there exists a constant R, depending only on a, such that

$$\operatorname{cov}(Q_i, Q_j) \leqq R\beta^{i-j} \quad \text{when} \quad i \leqq j.$$

Eventually,

$$\operatorname{var}\sigma^{(3)} = \sum_{i,j=1}^{n-1}\operatorname{cov}(Q_i, Q_j) \leqq (n-1)R + 2R\sum_{j=2}^{n-1}\sum_{i=1}^{j-1}\beta^{i-j} = O(n).$$

Finally, $\sigma^{(4)}$ being obviously bounded, so is $\operatorname{var}\sigma^{(4)}$.

THEOREM 5.2. *Let all the partial quotients in the continued-fraction expansion of α be equal to a positive integer a. Then there exists a constant Λ depending only on a, and such that*

$$\frac{1}{m}\sum_{q=0}^{m-1} S_q^2 \leqq \Lambda \log m$$

for any positive integer $m>1$ and not exceeding its denominator if α is rational.

PROOF. As an immediate consequence of Lemmas 4.1 and 5.1, there exists a constant λ depending only on a and satisfying

$$\frac{1}{v_n}\sum_{q=0}^{v_n-1} S_q^2 \leqq \lambda \log v_n$$

for all n if α is irrational and for all $n \leqq n'$ if $\alpha = v_{n'-1}/v_{n'}$. If $v_{n-1} < m < v_n$, then

$$\frac{1}{m}\sum_{q=0}^{m-1} S_q^2 < \frac{1}{v_{n-1}}\sum_{q=0}^{v_n-1} S_q^2 < \frac{a+1}{v_n}\sum_{q=0}^{v_n-1} S_q^2 \leqq (a+1)\lambda \log v_n,$$

and if we put, for instance

$$\Lambda = \lambda(a+1)\log\big(a(a+1)\big)(\log a)^{-1},$$

we have (5.9), at least when $m \geqq v_2 = a$.

If necessary, an adjustment of the value of Λ will take care of the case $1 < m < v_2$.

As a corollary to Lemma 2.1 and Theorem 5.2 we have the following proposition:

THEOREM 5.3. *If all the partial quotients in the continued-fraction expansion of α are equal to a positive integer a and if $\mathbf{Z}$ is the sequence of points*

$$\langle 0, 0\rangle, \left\langle \frac{1}{m}, \{\alpha\}\right\rangle; \left\langle \frac{2}{m}, \{2\alpha\}\right\rangle, \ldots, \left\langle \frac{m-1}{m}, \{(m-1)\alpha\}\right\rangle,$$

m being an arbitrary positive integer if α is irrational and not exceeding its denominator if α is rational, then the mean-square discrepancy $D^{(2)}(\mathbf{Z})$ of $\mathbf{Z}$ satisfies

$$D^{(2)}(\mathbf{Z}) = O\big(m^{-1}(\log m)^{1/2}\big),$$

where the constant implied in the right-hand side depends only on a.

It may be worth returning for a moment to the behaviour of S_q.

LEMMA 5.4. *Under the conditions of Theorem 5.2, to any $\varepsilon > 0$ there corresponds a number c depending only on a and ε, and such that*

$$S_q < c\sqrt{\log m}$$

holds for all but at most εm values of $q \in [2, m)$, $m \geqq 2$ being arbitrary.

PROOF. We return to the probabilistic interpretation of our problem. According to the Chebyshev inequality, for any positive K, $\mathsf{P}[S_q^2 \geqq K] \leqq K^{-2}\mathsf{E}(S_q^2)$. In view of (5.9), putting $K=\sqrt{\Lambda \log m/\varepsilon}$, we find

$$\mathsf{P}\left[|S_q| \geqq \sqrt{\varepsilon^{-1}\Lambda \log m}\right] \leqq \varepsilon,$$

and the Lemma holds with $c=\sqrt{\Lambda/\varepsilon}$.

THEOREM 5.5. *Under the conditions of Theorem* 5.2, *to any* $\varepsilon>0$ *there corresponds a number C depending only on a and* ε, *and such that*

$$S_q < C\sqrt{\log q}$$

holds for all but at most εm *values of q in the interval* $[2, m)$, $m \geqq 2$ *being still an arbitrary integer.*

PROOF. We divide $[2, m)$ into the intervals $[2, 2^2)$, $[2^2, 2^3)$, ..., $[2^{r-1}, 2^r)$, and $[2^r, m)$, where $2^r < m \leqq 2^{r+1}$. According to the preceding Lemma, to any $\varepsilon' > 0$ there corresponds a number c' depending only on a and ε', and such that $S_q \geqq c'\sqrt{\log 2^\nu}$ holds in the interval $[2^{\nu-1}, 2^\nu)$ for not more than $2^\nu\varepsilon'$ values of q. But for these values of q, $\log 2^\nu \leqq 2 \log q$, and so we have

$$S_q \geqq C'\sqrt{2\log q} \tag{5.10}$$

for at most $2^\nu\varepsilon'$ values of q in $[2^{\nu-1}, 2^\nu)$. Similarly, (5.10) holds for at most $m\varepsilon'$ values of q in $[2^r, m)$. Thus, in all, the number of values of q in $[2, m)$ for which (5.10) holds is at most $(2^2+2^3+\ldots+2^r+m)\varepsilon' < 3m\varepsilon'$, and if we put $\varepsilon'=\varepsilon/3$ and $C=c'\sqrt{2}$, we obtain the conclusion of the theorem.

REMARK. The above Theorem might be surprising knowing the following:

($*$) *For* S_N *we have the same best possible* Ω*-estimation, as for* D_N:

$$S_N = \Omega(\log N)$$

and

$$D_N = \Omega(\log N).$$

($**$) *For* D_N *a much stronger result is true: for an arbitrary* α

$$D > c \log N$$

holds for all but at most N^ε *values of q;* $1 \leqq q \leqq N$ *where* $\varepsilon \to 0$ *with* $c \to 0$. (See V. T. Sós [13].)

Acknowledgement. The second author's contribution to this paper was a result of his work done within the Research Project 3468 on "Gleichverteilte und nicht gleichverteilte Zufallszahlen" of the Austrian Fund for the Furtherance of Scientific Research at the Statistics Institute of the Technical University in Graz.

REFERENCES

[1] DAVENPORT, H.: Note on irregularities of distribution, *Mathematika* **3** (1956), 131—135.
[2] HALTON, J. H. and ZAREMBA, S. K.: The extreme and L^2 discrepancies of some plane sets, *Monatsh. Math.* **73** (1969), 316—328.
[3] HLAWKA, E.: Funktionen von beschränkter Variation in der Theorie der Gleichverteilung, *Ann. Mat. Pura Appl.* (IV) **54** (1961), 325—333.
[4] HLAWKA, E.: Zur angenäherten Berechnung mehrfacher Integrale, *Monatsh. Math.* **66** (1962), 140—151.
[5] KOROBOV, N. M.: Number-theoretical Methods in Approximate Analysis, Fizmatgiz, Moscow, 1963 (in Russian).
[6] OSTROWSKI, J. A.: Bemerkungen zur Theorie der Diophantischen Approximationen, *Abh. Math. Sem. Hamburg* **1** (1922), 77—98.
[7] ROTH, K. F.: On irregularities of distribution, *Mathematika* **1** (1954), 73—79.
[8] SCHMIDT, W. M.: Irregularities of distribution VII, *Acta Arith.* **21** (1972), 45—50.
[9] VILENKIN, I. V.: Plane nets of integration, *Z. Vyčisl. Mat. i Fiz.* **7** (1967), 189—196 (in Russian); Engl. transl.: *USSR Comp. Math. and Math. Phys.* **7** (1) (1967), 256—267.
[10] ZAREMBA, S. K.: Some applications of multidimensional integration by parts, *Ann. Polon. Math.* **21** (1968), 85—96.
[11] ZAREMBA, S. K.: La méthode des bons treillis pour le calcul intégrales multiple, in S. K. Zaremba (Ed.), *Applications of Number Theory to Numerical Analysis,* Academic Press, New York, 1972, 39—119.
[12] ZAREMBA, S. K.: Good lattice points, discrepancy and numerical integration, *Ann. Mat. Pura Appl.* (IV) **73** (1966), 293—318.
[13] Sós, V. T.: On strong irregularities of the distribution of $\{n\alpha\}$ sequences, *Studies in Pure Mathematics*, To the memory of Paul Turán, Akadémiai Kiadó, Budapest, 1982, 474—484.

(Received May 9, 1980)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
ANALÍZIS I. TANSZÉK
MÚZEUM KRT. 6—8
H—1088 BUDAPEST
HUNGARY

STATISTISCHES INSTITUT DER
TECHNISCHEN UNIVERSITÄT IN GRAZ
A—8010 GRAZ
AUSTRIA

Present address:
YR HEN YSGOL
ABERFFRWD
ABERYSTWYTH SY23 3ND
WALES
GREAT BRITAIN

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 273—292

EXTREMUM PROBLEMS FOR THE MOTIONS OF A BILLIARD BALL IV. A HIGHER-DIMENSIONAL ANALOGUE OF KEPLER'S STELLA OCTANGULA

by

I. J. SCHOENBERG

Contents

1. Introduction. The following pages describe what would be merely an exercise in Descriptive Geometry, if it were not for the fact that we are in the space $\mathbf{R}^n$, and are thereby forced to use analytic geometry. The 8-pointed star called *Stella Octangula* (abbreviated to SO) mentioned in the title, is shown in Figure 1.
It is composed of the union of the surfaces of the two regular tetrahedra

$$\Pi_3 = ABCD, \quad \text{and} \quad \tilde{\Pi}_3 = \tilde{A}\tilde{B}\tilde{C}\tilde{D}, \tag{1.1}$$

both inscribed in the cube γ_3. Notice that Π_3 and $\tilde{\Pi}_3$ are regular simplices of $\mathbf{R}^3$, and that they are symmetric to each other with respect to the center O of γ_3.

My colleagues Carl DE BOOR and Donald CROWE observed that Kepler's SO is an analogue in $\mathbf{R}^3$ of the star of David, the tetrahedra Π_3, $\tilde{\Pi}_3$, playing the role of the two regular triangles of David's star. An analogue of SO in $\mathbf{R}^n$ now seems obvious: In $\mathbf{R}^n$ we consider two congruent regular simplices α_n and $\tilde{\alpha}_n$ having a common center O and so placed, that they are symmetric to each other with respect to O. This, however, is not the analogue of SO in $\mathbf{R}^n$, that we have in mind.

We denote by SO_n the analogue in $\mathbf{R}^n$ of SO to be now defined, writing $\mathrm{SO}_3=\mathrm{SO}$ if $n=3$. Let

$$\gamma_n = \{-1 \leqq x_i \leqq 1,\ i = 1, \ldots, n\} \tag{1.2}$$

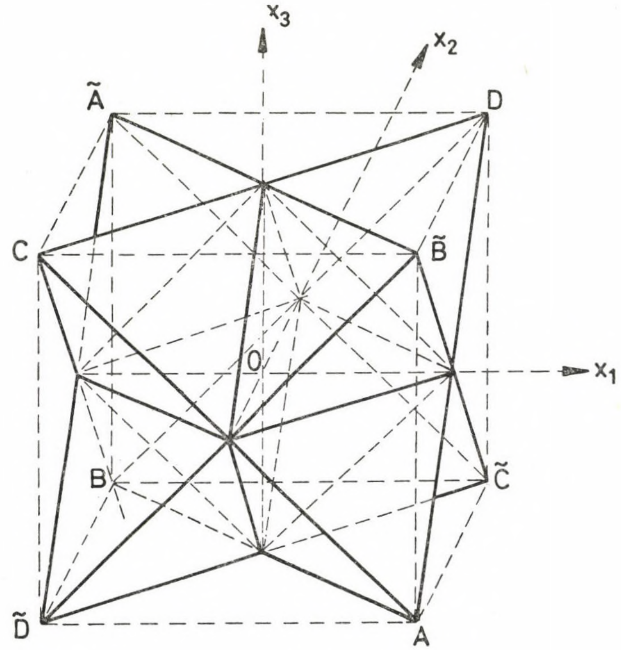


Fig. 1

be our fundamental measure polytope, or hypercube. For the *natural number j* we consider the cross-polytope

$$O_n^j = \left\{(x_i);\ \sum_1^n |x_i| = 2j-1\right\} \tag{1.3}$$

for values of j such that

$$2j-1 < n. \tag{1.4}$$

Its 2^n facets are in the hyperplanes (we abbreviate "hyperplane" to HP, plural HPs).

$$\sum_1^n \varepsilon_i x_i = 2j-1, \quad \text{where} \quad \varepsilon_i = \pm 1. \tag{1.5}$$

Its intersection with γ_n we denote by

$$F_j(\varepsilon_1, \ldots, \varepsilon_n) = \gamma_n \cap \left\{\sum_1^n \varepsilon_i x_i = 2j-1\right\}. \tag{1.6}$$

Notice that this intersection is a *non-degenerate convex* $(n-1)$*-dimensional polytope.* The reason for this is that *the vertex* $(\varepsilon_1, \ldots, \varepsilon_n)$ of γ_n, *and its center* $O=(0, \ldots, 0)$ *are on opposite sides of the HP* (1.5), *because we assume that* $n>2j-1$.

The $F_j(\varepsilon_1, \ldots, \varepsilon_n)$ are by definition the facets of SO_n; *to define* SO_n *we merely have to form their union*

$$SO_n = \bigcup_{2j-1<n} \bigcup_{\varepsilon_i=\pm 1} F_j(\varepsilon_1, \ldots, \varepsilon_n). \tag{1.7}$$

This, then, is our analogue of SO in $\mathbf{R}^n$.

Let us look at the simplest examples.

1. $n=2$. The inequality $2j-1<n$ is satisfied by the single value $j=1$. By (1.6) we obtain the edge

$$F_1(\varepsilon_1, \varepsilon_2) = \gamma_2 \cap \{\varepsilon_1 x_1 + \varepsilon_2 x_2 = 1\},$$

and so (1.7) reduces to

$$SO_2 = \bigcup_{\varepsilon_i=\pm 1} F_1(\varepsilon_1, \varepsilon_2) = \Pi_2, \tag{1.8}$$

which is the slanting square of Figure 2.

2. $n=3$. Again $2j-1<n$ has the only solution $j=1$. We see that (1.3), (1.4), reduces to the single octahedron

$$O_3^1 = \left\{(x_1, x_2, x_3);\ \sum_1^3 |x_i| = 1\right\}, \tag{1.9}$$

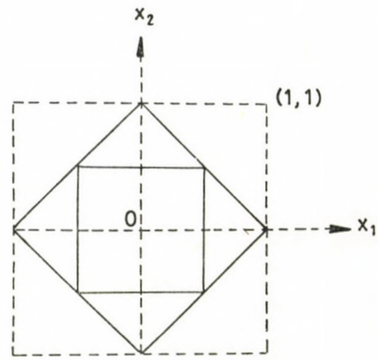


Fig. 2

and this is the regular octahedron whose 6 vertices are the centers of the 6 facets of the cube of Figure 1. According to (1.6) and (1.7) we are to take the 8 HPs of the facets of O_3^1 and form the union of their intersections with γ_3. In this way we get the 8 facets of

$$SO_3 = \bigcup_{\varepsilon_i=\pm 1} F_1(\varepsilon_1, \varepsilon_2, \varepsilon_3) = \Pi_3 \cup \tilde{\Pi}_3. \tag{1.10}$$

The definition (1.7) *is seen to lead to Kepler's SO if* $n=3$.

The two cases of $n=2$ and $n=3$ are typical of the general situation. The following results will be established.

1. *If* $n=2k$ *is even, then*

$$SO_n = \Pi_n \tag{1.11}$$

is a connected skew polytope in $\mathbf{R}^n$ *having* $n2^{n-1}$ *facets.*

2. *If* $n=2k+1$ *is odd, then*

$$SO_n = \Pi_n \cup \tilde{\Pi}_n. \tag{1.12}$$

Here Π_n and $\tilde{\Pi}_n$ are connected skew polytopes in $\mathbf{R}^n$, which are symmetric to each other with respect to the center O, hence

$$\tilde{\Pi}_n = -\Pi_n. \tag{1.13}$$

Π_n is composed of $(n-1)2^{n-2}$ facets.

The way the facets $F_j(\varepsilon_1, \ldots, \varepsilon_n)$ of SO_n are distributed among Π_n and $\tilde{\Pi}_n$ is described by the representations

$$\Pi_n = \bigcup_{j=1}^{k} \bigcup_{\prod_1^n \varepsilon_i = (-1)^j} F_j(\varepsilon_1, \ldots, \varepsilon_n), \tag{1.14}$$

and

$$\tilde{\Pi}_n = \bigcup_{j=1}^{k} \bigcup_{\prod_1^n \varepsilon_i = (-1)^{j+1}} F_j(\varepsilon_1, \ldots, \varepsilon_n). \tag{1.15}$$

The case when $n=3$, hence $k=1$, already shows clearly this structure: We divide the 8 facets of the octahedron (1.9) into two classes depending on the sign of the product $\varepsilon_1\varepsilon_2\varepsilon_3$, to obtain

$$\Pi_3 = \bigcup_{\varepsilon_1\varepsilon_2\varepsilon_3 = -1} F_1(\varepsilon_1, \varepsilon_2, \varepsilon_3) \tag{1.16}$$

and

$$\tilde{\Pi}_3 = \bigcup_{\varepsilon_1\varepsilon_2\varepsilon_3 = 1} F_1(\varepsilon_1, \varepsilon_2, \varepsilon_3). \tag{1.17}$$

The 4 facets of O_3^1 with $\varepsilon_1\varepsilon_2\varepsilon_3 = -1$, and the 4 facets with $\varepsilon_1\varepsilon_2\varepsilon_3 = 1$, form a kind of "checkerboard" design on the surface of O_3^1.

We may assume that $\mathbf{R}^n = \mathbf{R}^{n+1} \cap \{x_{n+1}=0\}$, and it then follows that $\gamma_n \subset \gamma_{n+1}$. From (1.6) we conclude that

$$F_j(\varepsilon_1, \ldots, \varepsilon_n) \subset F_j(\varepsilon_1, \ldots, \varepsilon_n, \varepsilon_{n+1}),$$

and finally *the definition* (1.7) *of SO_n proves the inclusion*

$$SO_n \subset SO_{n+1}. \tag{1.18}$$

For $n=2$ the inclusion $SO_2 \subset SO_3$, hence $\Pi_2 \subset \Pi_3 \cup \tilde{\Pi}_3$, is nicely exhibited in Figure 1, where the square Π_2, of Figure 2, appears as the intersection of $SO_3 = \Pi_3 \cup \tilde{\Pi}_3$ with the plane $x_3 = 0$.

We come to the characteristic properties of Π_n. The polytope Π_n was derived in our previous paper [4, equation (1.15)], and was there denoted by $\tilde{\Pi}_n^{n-1}$. It was there shown that Π_n is a so-called *König—Szűcs polytope*, and that among all such polytopes *in general position*, it is the one that *stays farthest away from the origin.* These concepts and results will be discussed in § 2. However, our discussion will not be independent of the paper [4], because we take over from [4] the parametric representation of Π_n stated in Theorem 1 below.

The contents of the chapters, sections, and appendix are deemed to be sufficiently explained by their headings.

I am grateful to Professor H. S. M. Coxeter for suggesting the present study of the geometric structure of SO_n based on the parametric representation of Π_n as given in [4]. An altogether different approach to SO_n was given by Coxeter in his paper [1].

I. The skew polytope Π_n

2. A characterization of the skew polytope Π_n of $\mathbf{R}^n$. Let

$$\gamma_n:\ -1 \leqq x_\nu \leqq 1, \quad (\nu = 1, \ldots, n) \tag{2.1}$$

denote the cube I^n, where $I=[-1, 1]$. In $\mathbf{R}^n$ we consider the hyperplane in parametric form

$$L_n':\ x_\nu = \sum_{i=1}^{n-1} \lambda_\nu^i u_i + a_\nu \quad (\nu = 1, \ldots, n), \tag{2.2}$$

where $\mathbf{a}=(a_\nu)$ is an interior point of γ_n. It helps to think of L_n' as carrying an $(n-1)$-dimensional pencil of light-rays emanating from the point $\mathbf{a}=(a_\nu)$ and spreading uniformly through L_n'. We think of the $2n$ facets $x_\nu=\pm 1$ of γ_n as mirrors which reflect back into γ_n any rays that may strike them, as well as any reflected rays. The complete path of these reflected rays is a skew $(n-1)$-dimensional polytope Π_n' such that

$$\Pi_n' \subset \gamma_n. \tag{2.3}$$

For the two special cases when $n=2$ and $n=3$, the skew polygons Π_2' and skew polyhedra Π_3' were first considered by D. König and A. Szűcs in their pioneering paper [3]. For this reason *we refer to Π_n' as a K—S polytope*. Observe that Π_2' is also the path of a billiard ball moving within the square "table" γ_2.

Our first task is to find a convenient representation for the polytope Π_n'. This is nicely obtained by using the *reflecting function* $R(u)$ defined as follows:

$$R(u) = \begin{cases} u & \text{if} \quad -1 \leqq u \leqq 1, \\ 2-u & \text{if} \quad 1 \leqq u \leqq 3, \end{cases} \quad \textit{and} \quad R(u+4) = R(u) \quad \textit{for} \quad u. \tag{2.4}$$

This function is an appropriate normalization of a so-called *linear Euler spline*; its graph is shown in Figure 3. Using $R(u)$ it is found that a parametric representation of the K—S polytope Π_n' obtained by reflecting the HP (2.2) is given by the equations

$$\Pi_n':\ x_\nu = R\left(\sum_{i=1}^{n-1} \lambda_\nu^i u_i + a_\nu\right), \quad (\nu = 1, \ldots, n). \tag{2.5}$$

The reasons are briefly as follows: By (2.4) we have $R(u)=u$ in the interval $-1\leqq u\leqq 1$, and this implies, by (2.2), that the intersection $L_n'\cap\gamma_n$ is left pointwise unchanged in passing from (2.2) to (2.5). The remaining portion of Π_n' is obtained

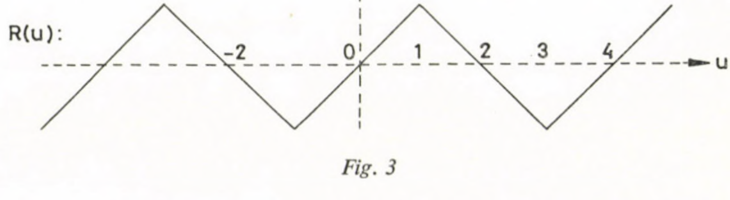


Fig. 3

by successive reflections of $L_n' \cap \gamma_n$ in the facets of γ_n, due to the zig-zag nature of the graph of Figure 3.

In order to avoid essentially lower-dimensional problems, we assume that the HP (2.2) is not parallel to any of the coordinate axes. The conditions for this are that

(2.6) *The* $n \times (n-1)$ *matrix* $\|\lambda_\nu^i\|$ *has no vanishing minor of order* $n-1$,

and we then say that L_n', as well as Π_n', are in *general position.*

Our problem is as follows.

PROBLEM 1. *Among all K—S polyhedra* Π_n', *defined by* (2.5), *which are in general position, to find those which stay away "as far as possible" from the center* O *of* γ_n.

What does "as far as possible" mean? We use here the Minkowskian norm

$$\|\mathbf{x}\|_\infty = \max(|x_1|, \ldots, |x_n|)$$

and determine the open neighbourhood $\|\mathbf{x}\|_\infty < \varrho$, with maximal ϱ, which contains no point of Π_n'. We find that

$$\max \varrho = \frac{1}{n}, \tag{2.7}$$

$\langle u \rangle$:

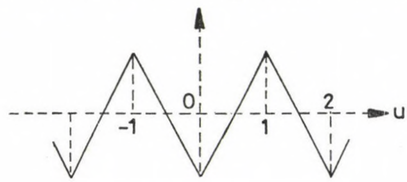


Fig. 4

and determine the corresponding extremizing Π_n', which we denote by Π_n. This extremum problem, resulting in Theorem 1 below, was solved in our previous paper [4]. In describing its solution, it is convenient to use a different linear Euler spline denoted by $\langle u \rangle$, and related to the function (2.4) by

$$\langle u \rangle = R(2u-1).$$

This function may also be defined by

$$\langle u \rangle = \begin{cases} 2u-1 & \text{if} \quad 0 \leqq u \leqq 1, \\ -2u-1 & \text{if} \quad -1 \leqq u \leqq 0, \end{cases} \quad \textit{and} \quad \langle u+2 \rangle = \langle u \rangle \quad \textit{for all} \quad u. \tag{2.8}$$

Its graph is shown in Figure 4

It should be clear that also the equations

$$x_\nu = \left\langle \sum_{i=1}^{n-1} \lambda_\nu^i u_i + a_\nu \right\rangle, \quad (\nu = 1, \ldots, n), \tag{2.9}$$

always define a K—S polytope, except that it no longer arises by reflexions of (2.2), but by reflexions of a HP simply related to (2.2). In terms of the function (2.8), the solution of Problem 1 as given in [4, § 6], is described by the following

THEOREM 1. 1. *The special K—S polytope*[1]

$$(2.10)\qquad \Pi_n: \begin{array}{ll} x_i = \langle u_i \rangle & (i = 1, \ldots, n-1) \\ x_n = \left\langle \sum_1^{n-1} u_i + \frac{n-1}{2} \right\rangle, & (0 \leqq u_i \leqq 2,\ i = 1, \ldots, n) \end{array}$$

is a finite skew polytope which has no point in common with the open cube

$$(2.11)\qquad C_n:\ \|x\|_\infty < \frac{1}{n},$$

so that

$$(2.12)\qquad \Pi_n \cap C_n = \emptyset;$$

while all 2^n *vertices* $\left(\pm\frac{1}{n}, \ldots, \pm\frac{1}{n}\right)$ *of* C_n *are points of* Π_n.

2. *If* (2.2) *is a HP* L_n' *in general position which is different from the HPs of the facets of* Π_n, *and also different from the HPs of the facets of*

$$(2.13)\qquad \tilde{\Pi}_n = -\Pi_n,$$

then the K—S polytope Π_n', *obtained by reflecting* L_n', *must satisfy*

$$(2.14)\qquad \Pi_n' \cap C_n \neq \emptyset,$$

which means that Π_n' *penetrates into the cube* (2.11).

REMARK. Let p satisfy $1 \leqq p \leqq \infty$, and let

$$(2.15)\qquad \|\mathbf{x}\|_p = \left(\sum_1^n x_\nu^p\right)^{1/p} = \left(\sum_1^n \left(\frac{1}{n}\right)^p\right)^{1/p} = \frac{1}{n^{1-\frac{1}{p}}}$$

denote the p-sphere circumscribed to the cube (2.11).

In § 11 *we show that Theorem* 1 *remains correct if we replace in its statement the cube* C_n *by the open p-neighborhood*

$$(2.16)\qquad \|\mathbf{x}\|_p < \frac{1}{n^{1-\frac{1}{p}}} \quad (1 \leqq p \leqq \infty).$$

[1] In [4, equations (1.3)] we used a different normalization of our present function (2.8); let us denote the old function for the moment by $\langle\langle u \rangle\rangle$. In terms of the present function $\langle u \rangle$ its expression is $\langle\langle u \rangle\rangle = \frac{1}{2}(\langle\langle u \rangle\rangle + 1)$. This reflecting function was adapted to the measure-polytope γ_n': $\{0 \leqq x_\nu \leqq 1,\ \nu = 1, \ldots, n\}$, which we now abandon in favor of (2.1). It should not be surprising that our equations (2.10) are identical with the old equations (1.15) of [4].

II. Symmetry properties of the polytope Π_n

3. Π_n is invariant on permutations of the axes. We state this property as

LEMMA 1. *If*

$$(x_1, \ldots, x_n) \in \Pi_n, \tag{3.1}$$

and $(x_{i_1}, \ldots, x_{i_n})$ *is a permutation of the coordinates* $(x_1, \ldots, x_n)$, *then*

$$(x_{i_1}, \ldots, x_{i_n}) \in \Pi_n. \tag{3.2}$$

PROOF. This will become evident as soon as we rewrite the equations (2.10) in a symmetric form. We introduce the new parameter u_n by the equation

$$\sum_1^n u_\nu = -\frac{n-1}{2}, \tag{3.3}$$

and (2.10) show that

$$x_n = \left\langle \sum_1^{n-1} u_i + \frac{n-1}{2} \right\rangle = \langle -u_n \rangle = \langle u_n \rangle.$$

Therefore the equations (2.10) may be replaced by the symmetric system

$$x_\nu = \langle u_\nu \rangle, \quad (\nu = 1, \ldots, n), \tag{3.4}$$

where the n parameters u_ν are connected by the equation (3.3). Since (3.4) is symmetric in the u, the lemma has become evident.

4. The symmetries of the polytope Π_n. These depend strongly on the parity of n.

LEMMA 2. *We assume that*

$$n = 2k \quad \textit{is even.} \tag{4.1}$$

The polytope Π_n remains invariant if we perform a reflexion in the hyperplane $x_\nu = 0$.

PROOF. By Lemma 1 we may assume that $\nu = 1$, and we are to show that

$$(x_1, \ldots, x_n) \in \Pi_n \quad \text{implies that} \quad (-x_1, x_2, \ldots, x_n) \in \Pi_n. \tag{4.2}$$

By (2.10) and (4.1) let

$$x_i = \langle u_i \rangle, \; x_n = \left\langle \sum_1^{n-1} u_i + k - \frac{1}{2} \right\rangle$$

and

$$x_i' = \langle u_i' \rangle, \; x_n' = \left\langle \sum_1^{n-1} u_i' + k - \frac{1}{2} \right\rangle$$

be two points of Π_n, where

$$u_1' = 1 - u_1, \quad u_i' = -u_i \qquad (i = 2, \ldots, n-1).$$

From Figure 4 $\langle u\rangle$ is seen to be odd about the point $u=\frac{1}{2}$, and so

$$x_1' = \langle u_1'\rangle = \langle 1-u_1\rangle = -\langle u_1\rangle = -x_1,$$

while

$$x_i' = \langle u_i'\rangle = \langle -u_i\rangle = \langle u_i\rangle = x_i \quad \text{for} \quad i = 2, \ldots, n-1.$$

Finally

$$x_n' = \left\langle \sum_1^{n-1} u_i' + k - \frac{1}{2}\right\rangle = \left\langle 1 - \sum_1^{n-1} u_i + k - \frac{1}{2}\right\rangle = \left\langle k + \frac{1}{2} - \sum_1^{n-1} u_i\right\rangle =$$

$$\left\langle \sum_1^{n-1} u_i - k - \frac{1}{2}\right\rangle = \left\langle \sum_1^{n-1} u_i + k - \frac{1}{2}\right\rangle = x_n,$$

proving (4.2).

A consequence of Lemma 2 is

COROLLARY 1. *If n is even, then Π_n has the origin O as center of symmetry, hence*

(4.3) $$\Pi_n = -\Pi_n.$$

For odd n we have

LEMMA 3. *Let*

(4.4) $$n = 2k+1 \quad \textit{be odd.}$$

The polytope Π_n remains invariant if we perform a reflexion in the hyperplane $x_i=0$ followed by a reflexion in $x_j=0$ ($j\neq i$).

PROOF. Again in view of Lemma 1 we may assume that $i=1$ and $j=2$, and we are to show that the mapping

(4.5) $$(x_1, x_2, \ldots, x_n) \to (-x_1, -x_2, x_3, \ldots, x_n) \quad \textit{leaves } \Pi_n \textit{ invariant.}$$

By (4.4) the equations (2.10) become

$$x_i = \langle u_i\rangle, \; x_n = \left\langle \sum_1^{n-1} u_i + k\right\rangle.$$

However, the identity $\langle u+1\rangle = -\langle u\rangle$ shows that $\langle u+k\rangle = (-1)^k\langle u\rangle$, and so we may replace (2.10) by

(4.6) $$\Pi_n: \begin{array}{ll} x_i = \langle u_i\rangle, & (i = 1, \ldots, n-1), \\ x_n = (-1)^k \left\langle \sum_1^{n-1} u_i\right\rangle. & \end{array}$$

Besides the point (x_ν) defined by (4.6), we consider a second point (x_ν') corresponding to parameters u_i' defined by

$$u_1' = u_1+1, \quad u_2' = u_2-1, \quad u_i' = u_i \qquad (i = 3, \ldots, n-1).$$

From the equations

$$x_1' = \langle u_1'\rangle = \langle u_1+1\rangle = -\langle u_1\rangle = -x_1,$$
$$x_2' = \langle u_2'\rangle = \langle u_2-1\rangle = -\langle u_2\rangle = -x_2,$$
$$x_i' = x_i \qquad (i = 3, \ldots, n-1),$$
$$x_n' = (-1)^k\left\langle \sum_1^{n-1} u_i'\right\rangle = (-1)^k\left\langle \sum_1^{n-1} u_i\right\rangle = x_n,$$

we see that (4.5) indeed holds.

We have the

COROLLARY 2. *Let* (4.4) *hold, and let us define the symmetric image of* Π_n *with respect to the origin by*

$$\tilde{\Pi}_n = -\Pi_n. \tag{4.7}$$

Then a reflexion of Π_n *in a coordinate hyperplane carries* Π_n *into* $\tilde{\Pi}_n$.

PROOF. By Lemma 3, and because n is odd, we have

$$(x_1, x_2, \ldots, x_n)\in \Pi_n \tag{4.8}$$

if and only if $(x_1, -x_2, \ldots, -x_n)\in\Pi_n$. Now (4.7) shows that the last inclusion holds iff

$$(-x_1, x_2, \ldots, x_n)\in \tilde{\Pi}_n. \tag{4.9}$$

The equivalence of (4.8) with (4.9) proves our corollary.

III. The geometric structure of Π_n revealed

We separate the discussion according to the parity of n.

5. The case when $n=2k$ is even. From $n=2k$ and properties of $\langle u\rangle$, the equations (2.10) defining Π_n may be written

$$\Pi_{2k}: \quad \begin{aligned} x_i &= \langle u_i\rangle, \qquad (i = 1, \ldots, n-1),\\ x_n &= (-1)^k\left\langle \frac{1}{2} - \sum_1^{n-1} u_i\right\rangle. \end{aligned} \tag{5.1}$$

On the other hand we have defined the analogue of Kepler's Stella Octangula by the equation (1.7). As j ranges over the natural numbers satisfying $2j-1<n=2k$, we find that $j=1, 2, \ldots, k$, and so we may rewrite (1.7) as

$$SO_{2k} = \bigcup_{j=1}^{k} \bigcup_{\varepsilon_i=\pm 1} F_j(\varepsilon_1, \ldots, \varepsilon_{2k}), \tag{5.2}$$

where

$$F_j(\varepsilon_1, \ldots, \varepsilon_{2k}) = \gamma_{2k}\cap\left\{\sum_1^{2k} \varepsilon_i x_i = 2j-1\right\}. \tag{5.3}$$

Our main result for even n is

THEOREM 2. *We have*

$$\Pi_{2k} = SO_{2k}. \tag{5.4}$$

PROOF. Observe that Π_{2k} as defined by (5.1) and expressed in terms of $\langle u\rangle$, is automatically a K—S polytope. On the other hand SO_{2k}, as defined by (5.2), appears as just a collection of $k2^{2k}$ facets about which, a priori, we have no idea how they hang together, if at all. We will show, however, that *the facets of* SO_{2k} *are precisely the facets of* Π_{2k}.

We start by identifying an obvious facet of Π_{2k}, and we do this by restricting the u_i to satisfy the inequalities

$$0 \leqq u_i \quad (i = 1, \ldots, n-1), \qquad \sum_1^{n-1} u_i \leqq \frac{1}{2}. \tag{5.5}$$

Figure 4 and (5.1) show that the x_ν may be explicitly expressed in terms of the u_i by using $\langle u\rangle = 2u-1$ if $0 \leqq u \leqq 1$. It follows that the image of the simplex (5.5) is in the HP

$$x_i = 2u_i - 1, \ (-1)^k x_n = 1 - 2\sum_1^{n-1} u_i - 1 = -2\sum_1^{n-1} u_i.$$

Eliminating the u_i we find that $(-1)^k x_n = \sum_1^{n-1}(-1-x_i) = -(n-1) - \sum_1^{n-1} x_i$, and it follows that the HP

$$-x_1 - x_2 - \ldots - x_{n-1} - (-1)^k x_n = n-1$$

contains a facet of Π_{2k}. Applying to it the arbitrary reflexions in the coordinate HPs allowed by Lemma 2, we conclude that the 2^n HPs

$$\sum_1^n \varepsilon_i x_i = n-1 \quad (= 2k-1) \tag{5.6}$$

contain facets of Π_n. Notice that *these are precisely the facets*

$$\bigcup_{\varepsilon_i = \pm 1} F_k(\varepsilon_1, \ldots, \varepsilon_{2k}) \tag{5.7}$$

of the last term (for $j=k$*) of the union* (5.2).

The question arises:

(5.8) *Where do we go from here by reflexions?*

An answer depends on which facets of γ_n the facets of (5.7) intersect. And by intersect we mean *strictly*, i.e., intersecting the *interior* of the facet $\gamma_n \cap \{x_\nu = \eta\}$, $(\eta = \pm 1)$. Incidentally, we will just write $x_\nu = \eta$, meaning thereby the facet $\gamma_n \cap \{x_\nu = \eta\}$.

An answer to the question (5.8) is provided by

LEMMA 4. *The hyperplane* (5.6) *intersects strictly the facets*

$$x_i = \varepsilon_i, \quad (i = 1, \ldots, n), \tag{5.9}$$

but (5.6) *has no point in common with the remaining* n *facets of* γ_n.

PROOF. 1. We may assume that $i=1$. The intersection of (5.6) with the facet $x_1=\varepsilon_1$ is in the HP

$$\sum_2^n x_i\varepsilon_i = n-2,$$

and this equation has solutions in the open cube

$$-1 < x_2 < 1,\ -1 < x_3 < 1, \ldots, -1 < x_n < 1,$$

for instance the point $x_i=(n-2)(n-1)^{-1}\varepsilon_i$ $(i=2, \ldots, n)$.

2. On the other hand, the intersection of (5.6) with the facet $x_1=-\varepsilon_1$ is in the HP

$$\sum_2^n \varepsilon_i x_i = n$$

and this equation has evidently no solutions in the closed cube

$$-1 \leqq x_2 \leqq 1, \ldots, -1 \leqq x_n \leqq 1,$$

because its left side has only $n-1$ terms, all of absolute value $\leqq 1$.

Let the HP

$$\sum_1^n \varepsilon_i x_i = C \tag{5.10}$$

intersect the interior of γ_n so as to produce the facet

$$F = \gamma_n \cap \left\{\sum \varepsilon_i x_i = C\right\}.$$

What are the reflexions of F in the facet of γ_n? The answer is given by

LEMMA 5. 1. *If* (5.10) *intersects strictly the facet* $x_1=\varepsilon_1$, *then its reflexion in* $x_1=\varepsilon_1$ *is in the HP*

$$-\varepsilon_1 x_1 + \sum_2^n \varepsilon_i x_i = C-2. \tag{5.11}$$

2. *If* (5.10) *intersects strictly the facet* $x_1=-\varepsilon_1$, *then its reflexion in* $x_1=-\varepsilon_1$ *is in the HP*

$$-\varepsilon_1 x_1 + \sum_2^n \varepsilon_i x_i = C+2. \tag{5.12}$$

PROOF. 1. To perform the reflexion it is convenient to shift the origin to the point $(\varepsilon_1, 0, \ldots, 0)$ by writing (5.10) in the form

$$\varepsilon_1(x_1-\varepsilon_1) + \sum_2^n \varepsilon_i x_i = C-1.$$

We now obtain the equation of the reflected HP by changing the sign of the factor $(x_1-\varepsilon_1)$ to obtain $-\varepsilon_1(x_1-\varepsilon_1)+\sum_2^n \varepsilon_i x_i = C-1$, and finally (5.11). 2. Likewise, to reflect (5.10) in $x_1=-\varepsilon_1$, we write (5.10) as $\varepsilon_1(x_1+\varepsilon_1)+\sum_2^n \varepsilon_i x_i=C+1$, to obtain the reflected HP $-\varepsilon_1(x_1+\varepsilon_1)+\sum_2^n \varepsilon_i x_i=C+1$, and finally (5.12).

By Lemma 4 we can reflect (5.6) in $x_1=\varepsilon_1$; setting $C=n-1=2k-1$ we obtain from (5.11), by applying the reflexions of Lemma 2, the entire collection of 2^n HPs

$$\sum_1^n \varepsilon_i x_i = 2k-3, \quad \text{for arbitrary} \quad \varepsilon_i = \pm 1. \tag{5.13}$$

If we reflect this HP in $x_1=-\varepsilon_1$, we return to the HPs (5.6). However, if $n \geqq 6$, and if we reflect (5.13) in $x_1=\varepsilon_1$, we obtain, again via Lemma 2, the collection of 2^n HPs

$$\sum_1^n \varepsilon_i x_i = 2k-5, \quad \text{for arbitrary} \quad \varepsilon_i = \pm 1. \tag{5.14}$$

We can continue this process until we reach the 2^n HPs

$$\sum_1^n \varepsilon_i x_i = 1. \tag{5.15}$$

We claim that *from this point on no further HPs will appear by reflexions*. Indeed, reflecting (5.15) in $x_1=\varepsilon_1$, we obtain by Lemma 5, for $C=1$, the HP

$$-\varepsilon_1 x_1+\sum_2^n \varepsilon_i x_i = -1,$$

which is already among the HPs (5.15). Reflexion in $x_1=-\varepsilon_1$ will lead to a HP (5.10) with $C=3$, which was already obtained before.

Our discussion shows that SO_{2k}, of (5.2), a K—S polytope which is identical with Π_{2k}, proving Theorem 2.

6. The case when $n=2k+1$ is odd. We found in (4.6) that we can write

$$\Pi_{2k+1}: \begin{aligned} x_i &= \langle u_i \rangle, \qquad (i = 1, \dots, n-1), \\ x_n &= (-1)^k \left\langle \sum_1^{n-1} u_i \right\rangle. \end{aligned} \tag{6.1}$$

In (1.7) we have defined SO_n, which in our case when $n=2k+1$, becomes

$$SO_{2k+1} = \bigcup_{j=1}^{k} \bigcup_{\varepsilon_i=\pm 1} F_j(\varepsilon_1, \dots, \varepsilon_{2k+1}). \tag{6.2}$$

In (1.14) and (1.15) we have decomposed this union into two parts

$$SO_{2k+1} = \Sigma_0 \cup \Sigma_1, \tag{6.3}$$

where

$$\Sigma_0 = \bigcup_{j=1}^{k} \bigcup_{\Pi \varepsilon_i=(-1)^j} F_j(\varepsilon_1, \dots, \varepsilon_{2k+1}) \tag{6.4}$$

and

$$\sum_1 = \bigcup_{j=1}^{k} \bigcup_{\Pi\varepsilon_i=(-1)^{j+1}} F_j(\varepsilon_1, \ldots, \varepsilon_{2k+1}). \tag{6.5}$$

Here we wish to prove

THEOREM 3. *We have*

$$\Pi_{2k+1} = \sum_0 \quad \textit{and} \quad \tilde{\Pi}_{2k+1} = \sum_1. \tag{6.6}$$

PROOF. This is a variation of our proof of Theorem 2. We begin by identifying a certain set of facets of Π_{2k+1}. Restricting the u_i to the simplex

$$0 \leqq u_i, \quad \sum_1^{n-1} u_i \leqq 1, \tag{6.7}$$

and expressing the x_ν of (6.1), using (2.8), in terms of the u_i, we find on eliminating the u_i between these n equations, that the simplex (6.7) is mapped by (6.1) into the HP

$$-x_1-x_2-\ldots-x_{n-1}+(-1)^k x_n = n-2(=2k-1). \tag{6.8}$$

Notice that the product of the coefficients of the left side $=(-1)^k$, because $n-1$ is even. Also, because $n=2k+1$, we can no longer use Lemma 2, but must appeal to Lemma 3, with the result that from (6.8) we get the collection of 2^{n-1} HPs

$$\sum_1^n \varepsilon_i x_i = 2k-1, \quad \textit{where} \quad \prod_1^n \varepsilon_i = (-1)^k. \tag{6.9}$$

Lemma 5 remains valid. From (5.11) and (5.12), we see that a reflexion in $x_1=\varepsilon_1$, or $x_1=-\varepsilon_1$, will change the fixed sign of the product $\prod_1^n \varepsilon_i$ for the successive families of HPs (5.13) and (5.14) thus reached.

In this way we find that the collection of facets (6.4) is closed with respect to reflexions. It follows that Σ_0 is a finite K—S polytope which must be identical with the K—S polytope Π_{2k+1}. This proves the first identity (6.6). Finally, it should be clear from (6.4) and (6.5) that

$$\Sigma_1 = -\Sigma_0.$$

In view of $\tilde{\Pi}_{2k+1} = -\Pi_{2k+1}$, the second identity (6.6) follows from the first.

7. The polytope Π_n and the cube C_n are disjoint. It seems worthwhile to point out that the results of Part III immediately imply the property (2.12) of Theorem 1, to the effect that Π_n does not penetrate into the open hypercube

$$C_n\colon \ \|\mathbf{x}\|_\infty < \frac{1}{n}. \tag{7.1}$$

Let $n=2k$ be even. Among the facets of SO_n as exhibited in (5.2), the facets $F_1(\varepsilon_1, \ldots, \varepsilon_n)$ are nearest the origin O. The HP of $F_1(\varepsilon_1, \ldots, \varepsilon_n)$ has the equation

$$\sum_1^n \varepsilon_i x_i = 1, \tag{7.2}$$

and it evidently contains the vertex

$$\left(\frac{\varepsilon_1}{n}, \frac{\varepsilon_2}{n}, \ldots, \frac{\varepsilon_n}{n}\right) \quad \text{of} \quad C_n. \tag{7.3}$$

Therefore (7.2) is seen to be the HP through (7.3) and perpendicular to the diagonal of C_n joining its center O to its vertex (7.3).

If $n=2k+1$ *is odd*, the situation is similar, in view of (6.4), the only difference being that we consider only such HPs (7.2), and vertices (7.3), that satisfy the condition

$$\varepsilon_1 \varepsilon_2 \ldots \varepsilon_n = +1.$$

IV. The true shape and size of the facets of Π_n

8. A choice of coordinates in the parameter space $\mathbf{R}^{n-1}$. As throughout this paper, our foundation is the representation

$$\Pi_n: \begin{aligned} x_i &= \langle u_i \rangle, \qquad\qquad (i = 1, \ldots, n-1), \\ x_n &= \left\langle \sum_1^{n-1} u_i + \frac{n-1}{2} \right\rangle, \end{aligned} \tag{8.1}$$

our objective being to describe geometrically the mapping

$$F\colon (u_i) \mapsto (x_\nu). \tag{8.2}$$

This will be a piecewise isometry, provided that we select in $\mathbf{R}^{n-1}$ a coordinate system as follows.

Let α_{n-1} be a regular simplex in $\mathbf{R}^{n-1}$ such that

$$\textit{all edges of} \quad \alpha_{n-1} \quad \textit{are} = \sqrt{8}. \tag{8.3}$$

Let O be one of its vertices and let $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_{n-1}$ denote the vectors representing its $r-1$ edges issuing from O. The point $\mathbf{u}=(u_i)$ is then represented by

$$\mathbf{u} = \sum_1^{n-1} \mathbf{f}_i u_i. \tag{8.4}$$

From our choice of the $\mathbf{f}_i$ we have, in terms of inner products, the equations

$$\mathbf{f}_i \cdot \mathbf{f}_i = \mathbf{f}_i^2 = 8, \quad \mathbf{f}_i \cdot \mathbf{f}_j = \sqrt{8}\,\sqrt{8} \cos 60^\circ = 8\frac{1}{2} = 4. \tag{8.5}$$

The mapping (8.2), explicitly given by (8.1), is piecewise linear due to the presence of the function $\langle u \rangle$. In particular (8.1) is continuous, and has everywhere continuous (in fact constant) partial derivatives $\partial x_\nu / \partial u_i$, with the exception of the hyperplanes on which the expressions inside the function $\langle \cdot \rangle$ assumes integer values. These HPs are

$$u_i = j \quad (j \in \mathbf{Z},\ i = 1, \ldots, n-1) \tag{8.6}$$

and

$$\sum_1^{n-1} u_i + \frac{n-1}{2} = j, \quad (j \in \mathbf{Z}). \tag{8.7}$$

If the point $\mathbf{u}=(u_i)$ is in none of these HPs we have by (8.1) and (2.8) that

$$dx_i = \pm 2\, du_i, \quad dx_n = \pm 2 \sum_1^{n-1} du_i$$

and therefore

$$(dx)^2 = \sum_1^n (dx_\nu)^2 = 4 \sum_1^{n-1} (du_i)^2 + 4 \left(\sum_1^{n-1} du_i\right)^2$$

and finally

$$(dx)^2 = 8 \sum_1^{n-1} (du_i)^2 + 8 \sum_{i<j} du_i\, du_j. \tag{8.8}$$

On the other hand, from (8.4)

$$(du)^2 = \left(\sum_1^{n-1} \mathbf{f}_i\, du_i\right)^2 = \sum_1^{n-1} \mathbf{f}_i^2 (du_i)^2 + 2 \sum_{i<j} (\mathbf{f}_i \cdot \mathbf{f}_j)\, du_i\, du_j,$$

and the equations (8.5) show that

$$(du)^2 = 8 \sum_1^{n-1} (du_i)^2 + 8 \sum_{i<j} du_i\, du_j. \tag{8.9}$$

The identity of the quadratic forms (8.8) and (8.9) shows that $|dx|=|du|$, and therefore the mapping (8.2) is an *isometry* in each of the cells into which the HPs (8.6) and (8.7) divide the space $\mathbf{R}^{n-1}$.

We state our result as

LEMMA 6. *The cells into which the HPs* (8.6) *and* (8.7) *dissect the space* $\mathbf{R}^{n-1}$ *represent in true shape and size the facets of the skew polytope* Π_n.

Let us look more closely at the dissection of $\mathbf{R}^{n-1}$ by the HPs (8.6), (8.7). The HPs of the system (8.6) being parallel to the oblique coordinate HPs, divide $\mathbf{R}^{n-1}$ into a lattice of congruent acute rhombohedra, the fundamental one being

$$Rh_0 = \{(u_i);\ 0 \leqq u_i \leqq 1,\ i = 1, \ldots, n-1\}. \tag{8.10}$$

Figure 5 represents in parallel projection Rh_0 for the case $n=4$. The location of the second system (8.7) depends on the parity of n. Accordingly our discussion branches out into two cases.

9. The dimension $n=2k$ is even. Now (8.7) becomes

$$\sum_1^{n-1} u_i = j + \frac{1}{2}, \quad (j \in \mathbf{Z}). \tag{9.1}$$

Since the sum $\sum u_i$ varies in the rhombohedron (8.10) from the value zero, at 0, to the value $n-1$ at the opposite vertex $\sum \mathbf{f}_i$, we see that the HPs (9.1) that intersect

Rh_0 are the $n-1$ HPs

$$(9.2)\qquad \sum_1^{n-1} u_i = j + \frac{1}{2} \quad \text{for} \quad j = 0, 1, \dots, n-2.$$

These $n-1$ HPs dissect Rh_0 into n cells. By Lemma 6 these n cells represent in true shape and size n facets of the polytope Π_n.

What are all the facets of Π_n, for $n=2k$, and how many are there? Rh_0 contains just n of these facets. To obtain them all, we recall that $\langle u \rangle$ has the period 2. We would therefore expect a fundamental region of the mapping (8.1) to be the domain

$$(9.3)\qquad D_{n-1} = \{(u_i);\ 0 \leqq u_i \leqq 2,\ i = 1, \dots, n-1\}.$$

This is again a rhombohedron of twice the linear size of Rh_0 and expressible as

$$D_{n-1} = \bigcup_{\eta_i=0,1} \left(Rh_0 \oplus \sum_1^{n-1} \mathbf{f}_i \eta_i \right).$$

Since each of these 2^{n-1} unit rhombohedra contains n cells, we conclude that

(9.4) *the total number of facets of* Π_n *is* $= n2^{n-1}$.

This agrees with the number given in Theorem 2 and therefore shows that all these facets are different.

Since the mapping (8.1) has the period 2 in each of the variables u_i, we conclude that opposite facets of D_{n-1} are to be identified, and we obtain the following:

THEOREM 4. *If $n=2k$ is even, then the skew polytope Π_n is topologically a torus T^{n-1}.*

For $n=4$ Figure 5 shows that Rh_0 is divided by the 3 planes (9.2) into 4 cells of which the first and fourth are regular tetrahedra having edges$=\sqrt{2}$, while the second and third are truncated tetrahedra, each bounded by 4 regular triangles and 4 regular hexagons. By (9.4) Π_4 has a total of $4\times 8=32$ facets of which 16 are tetrahedra and 16 are truncated tetrahedra.

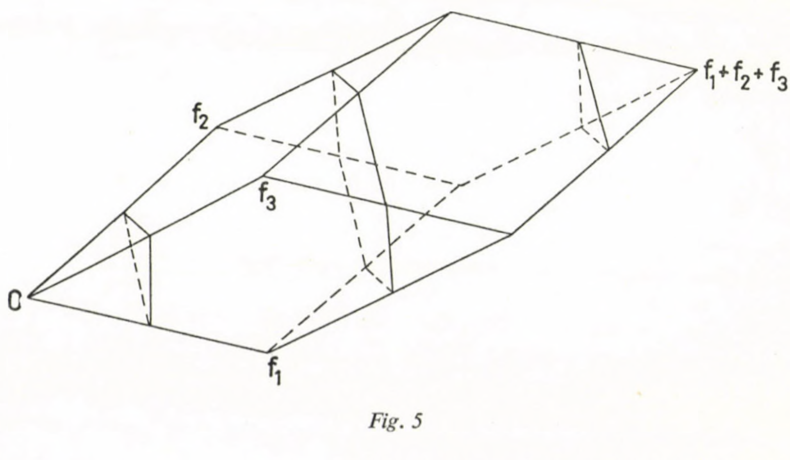


Fig. 5

10. What is the topological structure of Π_{2k+1}? In this case of n odd (8.7) becomes $\sum u_i = j$ $(j \in \mathbf{Z})$, and exactly $n-2$ among them, namely

$$\sum_{1}^{n-1} u_i = j \qquad (j = 1, 2, \ldots, n-2) \tag{10.1}$$

dissect the rhombohedron (8.10) into $n-1$ cells. We would expect a fundamental domain of the mapping (8.1) to be given by (9.3). However, this is not the case due to the following

LEMMA 7. *Let $n=2k+1$ be odd. The mapping* (8.1) *is even about every lattice point* $(u_i)=(j_1, j_2, \ldots, j_{n-1})$, $(j_i \in \mathbf{Z})$.

PROOF. We found in (6.1) that (8.1) may be written as

$$\Pi_n: \quad \begin{aligned} x_i &= \langle u_i \rangle \\ x_n &= (-1)^k \left\langle \sum_{1}^{n-1} u_i \right\rangle. \end{aligned} \tag{10.2}$$

The points (u_i) and (u_i') are symmetric in the point (j_i) provided that

$$u_i' = 2j_i - u_i, \quad (i = 1, \ldots, n-1),$$

and we are to show that this implies that

$$x_\nu' = x_\nu, \quad (\nu = 1, \ldots, n).$$

This follows from

$$x_i' = \langle u_i' \rangle = \langle 2j_i - u_i \rangle = \langle -u_i \rangle = \langle u_i \rangle = x_i$$

if $i<n$, and

$$x_n' = (-1)^k \left\langle \sum_{1}^{n-1} u_i' \right\rangle = (-1)^k \left\langle 2 \sum_{i=1}^{n-1} j_i - \sum_{1}^{n-1} u_i \right\rangle = (-1)^k \left\langle \sum_{1}^{n-1} u_i \right\rangle = x_n.$$

In particular, the mapping (10.2) is even about the point $(u_i)=(1, 1, \ldots, 1)$, which is the center of the rhombohedron (9.3). But then we can certainly reduce D_{n-1} to one of its two halves, namely

$$D_{n-1}^* = \{(u_i);\ 0 \leqq u_i \leqq 2,\ i = 1, \ldots, n-2,\ 0 \leqq u_{n-1} \leqq 1\}, \tag{10.3}$$

and still obtain the complete Π_n as the image of D_{n-1}^*.

In D_{n-1}^* we have the union of 2^{n-2} unit rhombohedra. Because (10.1) and their analogues, dissect each of these into $n-1$ cells, *we get for Π_n a total of $(n-1)2^{n-2}$ facets*. This agrees with the number given in Theorem 3 and shows that all these facets are different.

We turn now to the topological structure of Π_n. In the parallelepiped D_{n-1}^* of "height"$=1$ we consider $n-2$ pairs of opposite facets

$$u_i = 0 \quad \text{and} \quad u_i = 2, \quad (i = 1, \ldots, n-2), \tag{10.4}$$

and also the *top* $u_{n-1}=1$ and the *bottom* $u_{n-1}=0$. By the periodicity of (10.2), and by Lemma 7, we are to

1. Identify pairs of opposite facets (10.4);

2. Identify two points of the top $u_{n-1}=1$ that are symmetric in its center $(1, 1, \ldots, 1, 1)$. Likewise identify two points of $u_{n-1}=0$ that are symmetric in its center $(1, \ldots, 1, 0)$.

I am unable to identify the topological structure of the fundamental domain D^*_{n-1} with the above identifications of its boundary. Accordingly, we close Part IV with the following unsolved

PROBLEM 2. *To determine the topological structure of the polytope* Π_{2k+1}.

V. Appendix

11. Replacing the norm $\|\mathbf{x}\|_\infty$ in Theorem 1 by $\|\mathbf{x}\|_p$. Here we wish to justify the remark at the end of § 2. Let us first circumscribe a p-sphere $\|\mathbf{x}\|_p=\varrho_p$ to our cube

$$C_n\colon\ \|\mathbf{x}\|_\infty < \frac{1}{n}. \tag{11.1}$$

The p-norm $(1 \leqq p < \infty)$ of its vertices $\left(\pm\frac{1}{n}, \ldots, \pm\frac{1}{n}\right)$ is

$$\varrho_p = \left(n\frac{1}{n^p}\right)^{1/p} = 1/n^{1-(1/p)}, \tag{11.2}$$

and so the open p-sphere circumscribed to C_n is

$$S_p\colon\ \|\mathbf{x}\|_p < 1/n^{1-(1/p)}. \tag{11.3}$$

Since S_p is convex we certainly have the inclusion

$$C_n \subset S_p. \tag{11.4}$$

Let us show that[2]

$$S_p \subset S_1 = \left\{\sum_1^n |x_i| < 1\right\}. \tag{11.5}$$

PROOF. This follows from the monotonicity of the ordinary means

$$M_p(|x_i|) = \left(\frac{1}{n}\sum_1^n |x_i|^p\right)^{1/p} \quad \text{as functions of } p$$

as shown in [2, § 2.9, p. 26]. This monotonicity implies that $M_1(|x_i|) \leqq M_p(|x_i|)$, hence that

$$\frac{1}{n}\sum_1^n |x_i| \leqq \left(\frac{1}{n}\sum_1^n |x_i|^p\right)^{1/p}$$

or

$$\|\mathbf{x}\|_p = \left(\sum_1^n |x_i|^p\right)^{1/p} \geqq \frac{1}{n^{1-\frac{1}{p}}}\sum_1^n |x_i|.$$

[2] This is also evident geometrically from the convexity of S_p and because the 2^n facets of S_1 are the HPs of support of S_p at the 2^n vertices of C_n.

But then the inclusion $\mathbf{x}=(x_1, \ldots, x_n)\in S_p$, which means that $\|\mathbf{x}\|_p<1/n^{1-1/p}$, surely implies that $\sum_1^n |x_i|<1$ and (11.5) is established.

A third preliminary remark is that

$$\Pi_n \cap S_1 = \emptyset. \tag{11.6}$$

PROOF. This should be clear from (5.4) and (5.2) for *even n*, and from (6.6) and (6.4) for *odd n*. Indeed, observe that in either case the facets of Π_n, which are nearest the origin O, are in HPs of facets of the open cross-polytope

$$S_1 = \left\{\sum_1^n |x_i| < 1\right\}.$$

In order to show that $\|\mathbf{x}\|_\infty$ may be replaced by $\|\mathbf{x}\|_p$ in Theorem 1 we have to establish the following

LEMMA 8. 1. *That if* $1\leqq p<\infty$, *then*

$$\Pi_n \cap S_p = \emptyset. \tag{11.7}$$

2. *If* Π_n' *is a K—S polytope in general position with facets different from the facets of* SO_n, *then*

$$\Pi_n' \cap S_p \neq \emptyset. \tag{11.8}$$

PROOF. 1. From (11.6) and (11.5), the equation (11.7) follows immediately. 2. From Theorem 1 we know that

$$\Pi_n' \cap C_n \neq \emptyset, \tag{11.9}$$

hence Π_n' intersects C_n. But then (11.9) and (11.4) clearly imply (11.8), and our proof is complete.

REFERENCES

[1] COXETER, H. S. M.: The derivation of Schoenberg's star-polytopes from Schoute's simplex nets, *Proceedings of the Coxeter Symposium* held in May, 1979, in Toronto, to appear under the title „The geometric way", to be published by the Springer-Verlag.

[2] HARDY, G. H.—LITTLEWOOD, J. E. and PÓLYA, G.: *Inequalities*, Cambridge University Press, 1934.

[3] KÖNIG, D. and SZÜCS, A.: Mouvement d'un point abandonné à l'intérieur d'un cube, *Rendiconti del Circ. Mat. di Palermo* **38** (1913), 79—90.

[4] SCHOENBERG, I. J.: Extremum problems for the motions of a billiard ball III. The multi-dimensional case of König and Szücs, MRC TSR # 1880, September 1978, Madison, Wisconsin; *Studia Sci. Math. Hungar.* **13** (1978), 53—78.

(Received May 11, 1980)

UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER
610 WALNUT STREET
MADISON, WI 53706
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 293—297

TOPOLOGICAL COMPLETE SUBGRAPHS IN RANDOM GRAPHS

by

MIKLÓS AJTAI, JÁNOS KOMLÓS and ENDRE SZEMERÉDI

1. Introduction

We say that a graph contains a topological complete subgraph of order k (TCk), if there are k vertices such that any two of them are connected by a path, and these paths do not intersect. Let Σ denote the largest k for which the graph contains a TCk.

HAJÓS [7] conjectured that Σ is at least as large as the chromatic number χ of the graph.

This has been disproved recently by CATLIN [3].

ERDŐS and FAJTLOWICZ [5] remarked that almost all graphs provide a counter-example, for a typical graph has

$$\chi \approx n/\log n \quad \text{and} \quad \Sigma = O(\sqrt{n}).$$

BOLLOBÁS and CATLIN [2] made this latter bound precise by showing that one actually has

$$\Sigma \sim 2\sqrt{n}$$

for almost all graphs. They also mentioned that if the edges of a random graph are drawn independently with a (fixed) probability p, then one has

$$\Sigma \sim \sqrt{\frac{2n}{1-p}}$$

almost surely.

On the other hand, the case of scarce graphs has already been investigated by Erdős and RÉNYI. In their fundamental paper [6] they proved that the threshold for a topological complete 4-gon is $n/2$. More precisely, if we draw cn edges at random, then for $c>1/2$ (and n large) the graph almost surely contains a cycle with two "crossing" diagonals (a topological 4-gon), while in the case $c<1/2$ every component of a typical graph is either a tree or a component with exactly one cycle (i.e., $\Sigma \leqq 3$).

They also asked for the determination of the threshold for TCk in case $k>4$.

In this paper we determine the asymptotic behavior of Σ for random graphs with an arbitrarily prescribed number of edges, in particular, we show that for any fixed k the above threshold is $n/2$.

2. The theorem

Random graph will always mean a graph, in which the edges have been chosen at random with a prescribed probability p (possibly dependent on n, the number of vertices), and with mutually independent choices. We will mostly characterize the graph by the parameter $v=(n-1)p$, the expected valency of a vertex, rather than by p itself.

In these terms the Bollobás—Catlin result says that if v is of order n, then

$$\Sigma \sim c\sqrt{n}$$

almost surely, where

$$c = \sqrt{\frac{2}{1-v/n}}.$$

Let μ denote the maximal valency of the graph. Although μ is a random variable, it is very stable, almost surely near to the number $\mu(v)$ determined by the equality

$$\sum_{k \geqq \mu(v)} \frac{e^{-v} v^k}{k!} = \frac{1}{n}.$$

THEOREM. *For $v>1$, $v=o(\sqrt{n})$ we have*

$$\Sigma \sim \mu.$$

More precisely, we have for any fixed $\varepsilon>0$ and $0<\delta<1/100$:

$$(1-\delta)\mu < \Sigma \leqq \mu$$

with probability tending to 1 as $n\to\infty$ uniformly for

$$1+\varepsilon \leqq v \leqq \delta^4 \sqrt{n}.$$

The inequality $\Sigma \leqq \mu$ obviously holds for any graph, and thus the Theorem says that for typical scarce graphs this theoretical bound is achieved.

Since the valencies follow a Poisson distribution of parameter v (more precisely a binomial distribution), one has

$$\mu \sim v \quad \text{for} \quad v/\log n \to \infty$$

(the graph is almost regular)

$$\mu \sim d \log n \quad \text{for} \quad v \sim c \log n, \quad d = d(c)$$

$$\mu \sim \frac{\log n}{\log\left(\dfrac{\log n}{v}\right)} \quad \text{for} \quad v = o(\log n)$$

(i.e., $\mu\sim\mu(v)$, that is why we can treat μ as if it were a number, identify it with $\mu(v)$).

Thus the following more detailed picture arises:

COROLLARY. *One almost surely has*

$$\Sigma \sim d_0 \sqrt{n} \quad \textit{for} \quad \nu \sim cn, \; d_0 = d_0(c) = \sqrt{\frac{2}{1-c}}$$

(Bollobás—Catlin)

$$d_1 \sqrt{n} < \Sigma < d_2 \sqrt{n} \quad \textit{for} \quad \nu > c\sqrt{n}, \; \nu \ll n$$

$d_1 = d_1(c)$, $d_2 = d_2(c)$

$$\Sigma \sim \nu \quad \textit{for} \quad \nu \gg \log n, \; \nu \ll \sqrt{n}$$

$$\Sigma \sim d \log n \quad \textit{for} \quad \nu \sim c \log n, \; d = d(c)$$

$$\Sigma \sim \frac{\log n}{\log \left(\frac{\log n}{\nu}\right)} \quad \textit{for} \quad 1+\varepsilon \leqq \nu \ll \log n.$$

In particular,

$$\Sigma \sim \frac{\log n}{\log \log n}$$

for any $\nu > 1$, ν fixed (i.e., for fixed $\nu > 1$ and n large, Σ is practically independent of the value of ν), and $\Sigma = 2$ or 3 for $\nu < 1$, ν fixed.

This is an abrupt change (from 3 to $\log n / \log \log n$) in the structure of a random graph at the point $\nu = 1$, similar to the ones discovered by Erdős and Rényi.

3. The proof

The proof will heavily depend on the following recent result of the authors:

For $\nu > 1$ the random graph almost surely contains a path of length cn, $c = c(\nu)$. Further, c is close to 1 for large ν. (See [1]. Simultaneously proved by FERNANDEZ DE LA VEGA [4] to appear in *Studia Sci. Math. Hungarica.*)

PROOF of the Theorem. Let $\nu_1 = \nu_1(\delta)$ be such a large constant (independent of n) that almost all graphs on n vertices with average degree ν_1 have a path of length $(1-\delta/10)n$.

We restrict ourselves first to the case of large ν, the case of small ν will only need little modification.

Case I. $\nu > \nu_0 = 10\nu_1/\delta$. Set

$$\nu_2 = \nu - \nu_1 - \nu\delta/10 > \nu(1-\delta/5),$$

$$\nu_3 = \nu\delta/10,$$

$$p_i = \nu_i/(n-1), \quad i = 1, 2, 3.$$

By first throwing $nv_1/2$ edges (precisely, making randomizations with probability p_1) we almost surely get a path of length $m=(1-\delta/10)n$. Cut this path into $\delta^2 vm/\sqrt{n}$ small intervals $J_1, J_2, \dots$ each of length

$$l = \sqrt{n}/(\delta^2 v).$$

Now throw the next $nv_2/2$ edges (randomize with p_2). Choose a vertex v out of the remaining $\delta n/10$ ones. The probability that at this stage v and J_k gets connected, is

$$1-(1-p_2)^l \sim \frac{v_2}{\delta^2 v}\frac{1}{\sqrt{n}}.$$

Thus the number of intervals J_k connected to v follows a Poisson distribution of parameter $>(1-\delta/3)v$ (more precisely, a binomial distribution). Setting $\mu_2=(1-\delta/2)\mu$, we have, with a probability $1-o_n(1)$, μ_2 vertices out of the above $\delta n/10$ ones such that each one is connected to at least μ_2 intervals J_k.

Indeed, μ is determined by the condition that

$$\sum_{k\geq\mu}\frac{e^{-v}v^k}{k!}$$

be of order $1/n$, and this implies that for $\bar{v}=(1-\delta/3)v$ the probability

$$\sum_{k\geq\mu_2}\frac{e^{-\bar{v}}\bar{v}^k}{k!}$$

is larger than $\mu_2/(n\delta/10)$; now apply the laws of large numbers.

The condition $v\leq\delta^4\sqrt{n}$ guarantees that we can choose these intervals J_k disjoint.

Now consider $\mu_3=(1-\delta)\mu$ of these μ_2 vertices, and let $I_{i,1}, \dots, I_{i,\mu_2}$, $i=1, \dots, \mu_3$ denote the corresponding disjoint intervals.

Let us now throw down the last $nv_3/2$ edges $\left(p_3\sim\frac{\delta}{10}\frac{v}{n}\right)$. The probability that two of the above intervals are connected by at least one edge in this last stage, is

$$p_4 = 1-(1-p_3)^{n/(v^2\delta^4)} \sim \frac{1}{10\delta^3 v} \gg \frac{v}{n} = p,$$

thus these intervals (interpreted as vertices) form a very dense graph (cf. Shrinking lemma in AJTAI—KOMLÓS—SZEMERÉDI [1]).

We claim that one can almost surely find a "complete disjoint representative system" for these intervals, namely that for any pair (i, j), $i\neq j$, one can find indices k, t in such a way that $I_{i,k}$ and $I_{j,t}$ are connected by edges, and among these $\binom{\mu_3}{2}$ pairs each interval occurs at most once. Indeed, if this was not possible, then we would have two indices $i\neq j$ and $2(\mu_2-\mu_3)$ intervals

$$I_{i,k},\ k\in K\subset\{1, \dots, \mu_2\},\ |K| = \mu_2-\mu_3$$
$$I_{j,t},\ t\in T\subset\{1, \dots, \mu_2\},\ |T| = \mu_2-\mu_3$$

such that no $I_{i,k}$, $k\in K$, is connected to any of the intervals $I_{j,t}$, $t\in T$.

But the probability of this latter event is certainly less than

$$\binom{\mu_3}{2} 2^{2\mu_2}(1-p_3)^{(\mu_2-\mu_3)^2 n/(\nu^2\delta^4)} < 4^{\mu} e^{-\mu^2/(40\delta\nu)} < e^{-\mu} = o(1).$$

Case II. $\nu \leqq \nu_0$ (but $\nu \geqq 1+\varepsilon$).

Only change is that now we do not have a long path at the first stage, but after throwing $(1+\varepsilon/2)n/2$ edges we still get a path of length cn (say). We have $\varepsilon n/4$ additional edges to be thrown down (randomization with $p'=\varepsilon/(2n)$). Only change in the previous proof is that the valencies in the vertex-interval even "graph" do not follow a Poisson distribution of parameter almost ν, but only of small parameter $\nu'=\varepsilon c/2$. But the maximum of n Poisson variables is $\log n/\log\log n$ anyway, regardless how small ν' is. (The only point we needed $\nu>1$ was the first stage, when we found a path of length cn.)

REFERENCES

[1] Ajtai, M., Komlós, J. and Szemerédi, E.: The longest path in a random graph, *Combinatorica* **1** (1981), 1—12.

[2] Bollobás, B. and Catlin, P. A.: Topological cliques of random graphs, *J. Combinatorial Theory Ser. B* **30** (1981), 224—227.

[3] Catlin, P. A.: Hajós' graph-coloring conjecture: Variations and counterexamples, *J. Combinatorial Theory Ser. B* **26** (1979), 268—274.

[4] De la Vega, W. F.: Long paths in random graphs, *Studia Sci. Math. Hungar.* (to appear).

[5] Erdős, P. and Fajtlowicz, S.: On the conjecture of Hajós, *Combinatorica* **1** (1981), 141—143.

[6] Erdős, P. and Rényi, A.: On the evolution of random graphs, *Publ. Math. Inst. Hung. Acad. Sci.* **5A** (1960), 17—61.

[7] Hajós, G.: Über eine Konstruktion nicht n-färbbarer Graphen, *Wiss. Z. Martin-Luther-Univ. Halle—Wittenberg Math.-Natur. Reihe* **10** (1961), 116—117.

(Received May 19, 1980)

MTA MATEMATIKAI KUTATÓ INTÉZETE
REÁLTANODA U. 13—15
H—1053 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 299—301

MEASURE PRESERVING TRANSFORMATIONS

by

E. GALANIS

Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be measure spaces. A *bipartite measure graph* with vertex sets X and Y is a triple $(X, Y; M)$, where M is a subset of $X \times Y$ which is measurable with respect to $\mu \otimes \nu$. NASH—WILLIAMS [6] (see also [1], p. 101) asked for measure graph versions of well-known theorems in graph theory. Recently some of these questions were answered by BOLLOBÁS [2] and KATONA [4]. The aim of this note is to prove a measure graph analog of the theorem of PHILIP HALL.

As remarked in [2], the most natural bipartite measure graph analog of a complete matching in a bipartite graph is a measure preserving transformation $f: X \to Y$ whose graph is contained in M. It is also pointed out by Bollobás [2] that because of MAHARAM's theorem [5] on homogeneous measure algebras, when looking for measure preserving transformations $f: X \to Y$, we may assume that $\mathcal{A}$ and $\mathcal{B}$ are homogeneous measure algebras of the same density character. Even more, one is very likely to need topological restrictions on the (edge) set $M \subset X \times Y$. In this paper we shall consider the case when both X and Y are the unit interval [0, 1] with the Lebesgue measure.

For simplicity the Lebesgue measure of a set $A \subset [0, 1]$ is denoted by $|A|$; our graph theoretic terminology and notation is that of the book [1].

Before stating and proving our theorem, we have to make some preliminary remarks. Given a metric space X with distance d, a subset $A \subset X$ and a positive real number ε, put

$$A(\varepsilon) = \{y \in X: d(x, y) \leqq \varepsilon \text{ for some } x \in A\}.$$

The *Hausdorff distance* of two non-empty subsets $A, B \subset X$ is defined as

$$\varrho(A, B) = \inf \{\varepsilon : A \subset B(\varepsilon),\ B \subset A(\varepsilon)\}.$$

It is immediate that on the set $H(X)$ of all non-empty closed subsets of X the function ϱ is indeed a distance, so endowed with ϱ, $H(X)$ is a metric space. In the proof of our theorem we shall make use of the fact that $H(I)$ is *compact*, where I is the unit interval [0, 1] with the Euclidean distance. To prove the compactness of $H(I)$ it suffices to show that $H(I)$ is totally bounded and complete. The space $H(I)$ is totally bounded since the $2^n - 1$ sets of the form

$$A(S, n) = \left\{\frac{k}{n}: k \in S\right\}, \quad \emptyset \neq S \subset \{1, 2, \ldots, n\}$$

form a $1/n$-net in $H(I)$. Indeed, if $A \subset I$ is a non-empty closed set and

$$S = \left\{k: \left[\frac{(k-1)}{n}, \frac{k}{n}\right] \cap A \neq \emptyset\right\} \subset \{1, 2, \ldots, n\}$$

then

$$A \subset A(S, n) \subset A(1/n)$$

so

$$\varrho(A, A(S, n)) \leqq 1/n.$$

When proving the completeness of $H(I)$, it suffices to consider Cauchy sequences $(A_n)_{n=1}^{\infty}$ with $\varrho(A_n, A_m) < 2^{-n}$ for $m > n$. Put $B_n = \bigcup_{k \geqq n} A_k$. Then $\bar{B}_1 \supset \bar{B}_2 \supset \dots$ is a nested sequence of compact sets so $B = \bigcap_{n=1}^{\infty} \bar{B}_m$ is a closed subset of I. Since $B \subset A_n \subset \subset B_n \subset A_n(2^{-n})$, $\varrho(A_n, B) \leqq 2^{-n}$ and so $A_n \to B$ in the Hausdorff metric.

Let I_1 and I_2 be two copies of $I = [0, 1]$. For $M \subset I_1 \times I_2$ and $A \subset I_1$, put $M_A = \{y\colon (x, y) \in M,\ x \in A\}$. It is clear that if M contains the graph of a measure preserving transformation then $|M_A| \geqq |A|$ for every Lebesgue measurable set A. Our theorem shows that slightly more than this natural necessary condition is, in fact, sufficient.

THEOREM. *Let M be an open subset of $I_1 \times I_2$ such that $M_{I_1} = I_2$ and whenever A is a closed subset of I_1 and $\emptyset \neq A \neq I_1$ we have $|M_A| > |A|$. Then there is a piecewise linear measure preserving $1-1$ transformation $f\colon I_1 \to I_2$ whose graph is contained in M, that is*

$$(x, f(x)) \in M \quad \textit{for every } x, \quad 0 \leqq x \leqq 1.$$

PROOF. For every natural number n we shall construct an n by n bipartite graph G_n. The vertex classes of G_n are $X_n = \{x_1, x_2, \dots, x_n\}$ and $Y_n = \{y_1, y_2, \dots, y_n\}$, and $x_i y_j$ is an edge of G_n if and only if the square

$$Q_{ij}^{(n)} = \{(x, y)\colon i-1 \leqq nx \leqq i,\ j-1 \leqq ny \leqq j\}$$

is contained in M. It is clear that if the graph G_n contains a complete matching from X_n to Y_n then there is a piecewise linear measure preserving $1-1$ transformation $f\colon I_1 \to I_2$ whose graph is contained in M: we can use the diagonals of the squares $Q_{ij}^{(n)}$ corresponding to the edges of the complete matching to define such a transformation. Thus to prove the theorem it suffices to show that there is an n for which G_n does contain a complete matching from X_n to Y_n. Suppose no G_n contains a complete matching. Then by Philip Hall's well-known theorem ([3], see also [1], p. 9) for every natural number n there is a set $W_n \subset X_n$ such that

$$|\Gamma(W_n)| < |W_n|,$$

where $\Gamma_n(W_n) = \{y_k\colon x_i y_k \in E(G_n) \text{ for some } x_i \in W_n\}$.

Put

$$Z_n = \{x \in I_1\colon i-1 \leqq nx \leqq i,\ x_i \in W_n\},$$

that is let Z_n be the subset of the interval I_1 corresponding to W_n. By definition, each Z_n is a non-empty closed subset of I_1. Since the space $H(I_1)$ of non-empty closed subsets of I_1 with the Hausdorff metric is compact, there is a subsequence (Z_{n_k}) of (Z_n) converging to a closed subset $Z \subset I_1$ in the Hausdorff metric. To complete the proof we shall show that the existence of this set Z contradicts our assumptions.

Let us note first that $Z \neq I_1$. Indeed, as M is open and $M_I = I_1$, we can find a finite collection of open rectangles in M, whose projections into I_1 cover the whole of I_1. This implies immediately that there are $\varepsilon > 0$ and n_0 such that if $n \geqq n_0$, $W_n \subset X_n$ and $\varrho(Z_n, I_1) < \varepsilon$, where Z_n is the subset of I_1 corresponding to W_n, then $\Gamma_n(W_n) = Y_n$ so

$$|\Gamma_n(W_n)| = |Y_n| \geqq |W_n|.$$

Now, since $\emptyset \neq Z \neq I_1$

$$|M_Z| \geqq |Z| + 4\varepsilon$$

for some $\varepsilon > 0$. Since M is open there is a finite subset F of Z such that

$$|M_F| \geqq |Z| + 3\varepsilon.$$

For a natural number n, denote by $M_F^{(n)}$ the collection of points $y \in I_1$ for which there is an $x \in F$ such that the square with centre (x, y) and side length $4/n$ is contained in M. Since $|M_F^{(n)}| \to |M_F|$ as $n \to \infty$, there is an n_0 such that

$$|M_F^{(n_0)}| \geqq |Z| + 2\varepsilon.$$

We are about to obtain our desired contradiction. There is an $n \geqq n_0$ such that $\varrho(Z_n, Z) \leqq 1/n_0$ and $|Z_n| \leqq |Z| + \varepsilon$. Then the subset $W_n \subset \{1, 2, \ldots, n\}$ clearly satisfies

$$|\Gamma_n(W_n)| \geqq |W_n|$$

contrary to our assumption. The proof of our theorem is completed.

In conclusion let us note the following reformulation of our theorem.

Let $M \subset I_1 \times I_2$ be open and suppose that for every closed set $A \subset I_1$ there is a closed set $B \subset M_A \subset I_2$ such that

$$|B| \geqq |A|.$$

Then there is a piecewise linear measure preserving $1-1$ transformation $f\colon I_1 \to I_2$ whose graph is contained in M.

I am grateful to Dr. Bollobás for his advice concerning the problem discussed in this note.

REFERENCES

[1] Bollobás, B.: *Extremal graph theory*, Academic Press, 1978.
[2] Bollobás, B.: Measure graphs (to appear).
[3] Hall, P.: On representatives of subsets, *J. London Math. Soc.* **10** (1935), 26—30.
[4] Katona, G.: Continuous versions of some extremal hypergraph problems (to appear).
[5] Maharam, D.: On homogeneous measure algebras, *Proc. Nat. Acad. Sci. USA* **28** (1942), 108—111.
[6] Nash-Williams, C. St—J. A.: Unexplored and explored territories in graph theory, in "*New Direction in Graph Theory*", ed. F. Harary, Academic Press, 1973, 149—186.

(Received September 4, 1978)

DEPARTMENT OF MATHEMATICS
NATIONAL TECHNICAL UNIVERSITY OF ATHENS
24, PATISSION STREET
ATHENS 147
GREECE

PRINTED IN HUNGARY
Szegedi Nyomda, Szeged

Die *Studia Scientiarum Mathematicarum Hungarica* ist eine Halbjahrsschrift der Ungarischen Akademie der Wissenschaften. Sie veröffentlicht Originalbeiträge aus dem Bereich der Mathematik in deutscher, englischer, französischer oder russischer Sprache. Es erscheint jährlich ein Band.

Adresse der Redaktion: 1053 Budapest V., Reáltanoda u. 13—15, Ungarn.
Technischer Redaktor: J. Merza

Bestellbar bei Buch- und Zeitungs-Aussenhandelsunternehmen *Kultura* (Budapest 62, P. O. B. 149), oder bei den Vertretungen im Ausland.

Austauschabmachungen können mit der Bibliothek des Mathematischen Instituts (1053 Budapest V., Reáltanoda u. 13—15) getroffen werden.

Die zur Veröffentlichung bestimmten Manuskripte sind in zwei Exemplaren an die Redaktion zu schicken.

Studia Scientiarum Mathematicarum Hungarica est une revue biannuelle de l'Académie Hongroise des Sciences publiant des essais originaux, en francais, anglais, allemand ou russe, du domaine des mathématiques.

Rédaction: 1053 Budapest V., Reáltanoda u. 13—15, Hongrie.
Rédacteur technique: J. Merza

On s'abonne chez *Kultura*, Société pour le Commerce de Livres et Journaux (Budapest 62, P. O. B. 149) ou chez ses représentants à l'étranger.

Pour établir des relations d'échange on est prié de s'adresser à la Bibliothèque de l'Institut de Mathématique (1053 Budapest V., Reáltanoda u. 13—15).

On est prié d'envoyer les articles destinés à la publication en deux exemplaires à l'adresse de la Rédaction.

Studia Scientiarum Mathematicarum Hungarica — выходит два раза в год в Издании Академии Наук Венгрии. Журнал публикует оригинальные исследования в области математики на русском, немецком, английском, и французском языках. Отдельные выпуски составляют ежегодно один том.

Адрес редакции: 1053 Budapest V., Reáltanoda u. 13—15, Венгрия.
Технический редактор: J. Merza

Подписка на журнал принимается Внешнеторговым предприятием „Култура" (Budapest 62, P. O. B. 149) или его представительствами за границей.

По поводу отношения обмена просим обрашаться к Библиотеке Института Математики (1053 Budapest V., Reáltanoda u. 13—15).

Работы, предназначенные для опубликования в журнале следует направлять по адресу редакции в двух экземплярах.

INDEX

Index: 26 748

HU ISSN 0081—6906

Studia Scientiarum Mathematicarum Hungarica

AUXILIO
CONSILII INSTITUTI MATHEMATICI
ACADEMIAE SCIENTIARUM HUNGARICAE

REDIGIT
L. FEJES TÓTH

MUS XIV.
SC. 4.
9

AKADÉMIAI KIADÓ, BUDAPEST

Studia Scientiarum Mathematicarum Hungarica

A Magyar Tudományos Akadémia matematikai folyóirata

Szerkesztőség: 1053 Budapest V., Reáltanoda u. 13—15.

Technikai szerkesztő: Merza József

Kiadja az Akadémiai Kiadó, 1054 Budapest V., Alkotmány u. 21.

A *Studia Scientiarum Mathematicarum Hungarica* angol, német, francia vagy orosz nyelven közöl eredeti értekezéseket a matematika tárgyköréből. Félévenként jelenik meg, évi egy kötetben.

Előfizetési ára belföldre 120,— Ft, külföldre 165,— Ft. Megrendelhető a belföld számára az Akadémiai Kiadónál, a külföld számára pedig a Kultúra Könyv és Hírlap Külkereskedelmi Vállalatnál (1011 Budapest I., Fő u. 32.).

Cserekapcsolatok felvétele ügyében kérjük az MTA Matematikai Kutató Intézete Könyvtárához (1053 Budapest V., Reáltanoda u. 13—15.) fordulni.

Közlésre szánt dolgozatokat kérjük két példányban a szerkesztőség címére küldeni.

Studia Scientiarum Mathematicarum Hungarica is a journal of the Hungarian Academy of Sciences publishing original papers on mathematics, in English, German, French or Russian.

It is published semiannually, making up one volume per year.

Editorial Office: 1053 Budapest V., Reáltanoda u. 13—15, Hungary.
Technical Editor: J. Merza

Orders may be placed with *Kultura* Trading Co. for Books and Newspapers, Budapest 62, P.O.B. 149 or with its representatives abroad.

For establishing exchange relations please write to the Library of the Mathematical Institute (1053 Budapest V., Reáltanoda u. 13—15.)

Papers intended for publication should be sent to the Editor in 2 copies.

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 303—309

RADICALS OF ADDITIVELY INVERSIVE HEMIRINGS

by

TALAT SHAHEEN and S. M. YUSUF

In this paper we obtain a generalization of radical classes of a ring as discussed by WIEGANDT [4]. The first section contains some definitions and basic results. In the second section we define radical class of additively inversive hemirings. It is shown that every radical class satisfies the extension property. After defining an *e*-potent element and an *e*-set, a class **N** of additively inversive hemirings is defined which is shown to be a radical class and is in fact, a generalization of the nil radical class of rings.

1. Introduction

Two elements a and b of a semigroup S are said to be *inverses* of *each other* if $aba=a$ and $bab=b$. An *inverse semigroup* is a semigroup in which every element has a unique inverse. By [1] (Theorem 1.17, p. 28), the following two conditions on a semigroup S are equivalent:

(i) *S is regular, and any two idempotent elements of S commute with each other.*

(ii) *S is an inverse semigroup.*

The unique inverse of $a\in S$ is denoted by a^{-1}. A subsemigroup T of an inverse semigroup S is called an *inverse subsemigroup* of S if $a\in T$ implies $a^{-1}\in T$. If a, b are elements of an inverse semigroup, we have $(a^{-1})^{-1}=a$ and $(ab)^{-1}=b^{-1}a^{-1}$.

If we denote the binary operation of an inverse semigroup S by $+$, the inverse of an element $a\in S$ will be denoted by a'. In this case we have $a=a+a'+a$, $a'=$ $=a'+a+a'$. Moreover,

(i) $(a')'=a$,

(ii) $(a+b)'=b'+a'$.

An ordered triple $(S, +, \cdot)$, where S is a non-empty set and '$+$' and '$\cdot$' are two binary operations such that (i) $(S, +)$ is a semigroup, (ii) $(S, \cdot)$ is a semigroup, (iii) '$\cdot$' is both left and right distributive over $+$ is called a *semiring*. A semiring $(S, +, \cdot)$ is called *hemiring* if $+$ is commutative. A semiring (hemiring) $(S, +, \cdot)$ is called *additively inversive semiring* (hemiring) if $(S, +)$ is an inverse semigroup. If a, b are elements of an inversive semiring, then $(ab)'=a'b=ab'$ and $(ab)=a'b'$. If there exists an element $0\in S$ such that 0 is the identity element of $(S, +)$ and $0\cdot a=a\cdot 0=0$ for all $a\in S$, then 0 is called the *zero element* of the semiring $(S, +, \cdot)$. We shall write E_S to denote the set of all additive idempotents of the semiring S. A subset $\mathscr{A}\supseteq E_S$ of an additively inversive semiring S is called a *left ideal* if (i)

1980 *Mathematics Subject Classification.* Primary 16A78; Secondary 16A21.
Key words and phrases. General radicals, additively inversive hemirings.

$a, b \in \mathscr{A} \Rightarrow a+b \in \mathscr{A}$, (ii) $a \in \mathscr{A} \Rightarrow a' \in \mathscr{A}$, (iii) $a \in \mathscr{A}$ and $x \in S \Rightarrow xa \in \mathscr{A}$. If we replace (iii) by (iii)$'$ $a \in \mathscr{A}$ and $x \in S \Rightarrow ax \in \mathscr{A}$, then $\mathscr{A}$ is called a *right ideal*. If both (iii) and (iii)$'$ are satisfied, then $\mathscr{A}$ is called a *two-sided ideal* or simply an *ideal*.

If $(S, +, \cdot)$ is an additively inversive semiring, then E_S is an ideal of S. For if $e_1, e_2 \in E_S$, then $e_1+e_2 \in E_S$ (as idempotents commute in the inverse semigroup $(S, +)$). Again if $a \in E_S S$, $a=es$, where $e \in E_S$ then $a+a=es+es=(e+e)s=es=a$. Thus $a \in E_S$. Hence $E_S S \subseteqq E_S$. Similarly $SE_S \subseteqq E_S$. Thus E_S is an ideal of S. We call E_S a *trivial ideal* of the additively inversive semiring S. A left (right, two-sided) ideal $\mathscr{A}$ such that $\mathscr{A} \supset E_S$ is called a *non-trivial left (right, two-sided) ideal* of S.

Ordinarily we shall use script type $\mathscr{A}$, $\mathscr{B}$ etc. to denote ideals. But when we refer to the complex product of two such ideals we shall use the corresponding latin letters, and write AB. The symbol $\mathscr{A}\mathscr{B}$ will denote the ideal generated by AB.

THEOREM 1.1. *Let S be an additively inversive hemiring. Then the complex sum $\mathscr{A}_1+\mathscr{A}_2$ of two left (right, two-sided ideals) $\mathscr{A}_1$ and $\mathscr{A}_2$ is a left (right, two-sided) ideal of S and is the intersection of all left (right, two-sided) ideals of S containing $\mathscr{A}_1$ and $\mathscr{A}_2$.*

PROOF follows from [2], Theorem 4.

We shall denote by (A) the two-sided ideal generated by a complex A of an additively inversive hemiring S.

The proof of the following theorem follows from [3], Theorem 1.

THEOREM 1.2. *Let $\{\mathscr{A}_\alpha\}_{\alpha \in J}$ be an arbitrary collection of left (right, two-sided) ideals of an additively inversive semiring S. The left (right, two-sided) ideal of S generated by $\{\mathscr{A}\}_{\alpha \in J}$ is the set of all finite sum of elements from $\bigcup_{\alpha \in J} \mathscr{A}_\alpha$. This ideal is denoted by $\sum_{\alpha \in J} \mathscr{A}_\alpha$.*

COROLLARY. *Let $\{\mathscr{A}_1, \dots, \mathscr{A}_n\}$ denote a finite collection of left (right, two-sided) ideals of an additively inversive hemiring S. The left (right, two-sided) ideal of S generated by $\{\mathscr{A}_1, \mathscr{A}_2, \dots, \mathscr{A}_n\}$ is the complex sum $\mathscr{A}_1+\mathscr{A}_2+\dots+\mathscr{A}_n$.*

The concept of homomorphism (isomorphism) for semirings is defined in the usual manner.

A homomorphism ϕ of an additively inversive semiring S into an additively inversive semiring T is called a *principal homomorphism* or a *P-homomorphism* if $\phi | E_S$ is one-to-one mapping of the set E_S onto the set E_T. A P-homomorphism $\phi\colon S \to T$ is called a *non-trivial* homomorphism if $(S)\phi \neq E_T$. If $(S)\phi = E_T$, then ϕ is called a *trivial homomorphism*.

Let ϕ be a homomorphism of an additively inversive semiring S onto an additively inversive semiring T. The set of all element of S which map onto E_T is called the *kernel* of ϕ. Thus $\ker \phi = (E_T)\phi^{-1}$. We have

THEOREM 1.3 ([2], Theorem 8). *Let ϕ be a homomorphism of an additively inversive semiring S onto an additively inversive semiring T. The kernel of ϕ is a two-sided ideal of S and contains E_S.*

THEOREM 1.4 ([2], Theorems 9 and 10). *Let ϕ be a homomorphism of an additively inversive semiring S onto an additively inversive semiring T. If $\mathscr{I}$ is a left (right, two-sided) ideal of S, then $(\mathscr{I})\phi$ is a left (right, two-sided) ideal of T.*

THEOREM 1.5 ([2], Theorem 13). *Let $\mathcal{I}$ be any ideal of an additively inversive hemiring S. Let Σ denote the set of all those additively inversive subhemirings of S which contain $\mathcal{I}$ and $\bar{\Sigma}$ denote the set of all those additively inversive subhemirings of $S/\mathcal{I}$ which contain the set of all idempotents of the additive semigroup of $S/\mathcal{I}$. Denote $S/\mathcal{I}$ by $\bar{S}$. If $T\in\Sigma$ then $T\to T/\mathcal{I}(=\bar{T})$ is a one-to-one mapping of Σ onto $\bar{\Sigma}$. Also if T is an ideal of S containing $\mathcal{I}$ then $\bar{T}$ is an ideal of $\bar{S}$ and $S/T\cong\bar{S}/\bar{T}(=S/\mathcal{I}/T/\mathcal{I})$.*

THEOREM 1.6 ([2], Theorem 14). *Let S be an additively inversive hemiring. Let $\mathcal{I}=\bigcup_{e\in E_S} N_e$ be an ideal of S and T any additively inversive subhemiring of S. If $E_T=E_S\cap T$, then*

$$\Big(\bigcup_{e\in E_T} N_e+T\Big)/\bigcup_{e\in E_T} N_e\cong T/\Big(\bigcup_{e\in E_T} N_e\cap T\Big).$$

2. Radical class

A class **R** of additively inversive hemirings is called a *radical class,* if **R** satisfies the following requirement: $A\in\mathbf{R}$ *if and only if every non-trivial P-homomorphic image of A has a non-trivial* **R**-*ideal.*

We shall write **A** to denote the class of all additively inversive hemirings having their set of additive idempotents isomorphic. Thus if $A_1, A_2\in\mathbf{A}$, then $E_{A_1}\cong E_{A_2}$. We shall write **E** to denote the set of additive idempotents of any member of **A**. It follows that each of **A** and **E** is a *radical class.*

THEOREM 2.1. *A class* **R** *of additively inversive hemirings is a radical class if and only if* **R** *fulfils the following conditions:*

(a) **R** *is P-homomorphically closed.*

(b) *Every additively inversive hemiring $A\in\mathbf{A}$ contains an* **R**-*ideal* $\mathbf{R}(A)$, *which contains every other* **R**-*ideal of A.*

(c) *The additively inversive factor hemiring $A/\mathbf{R}(A)$ does not contain any non-trivial* **R**-*ideal, i.e.* $\mathbf{R}(A/\mathbf{R}(A))=E$.

The proof is essentially the standard one (cf. [4], Theorem 3.2), so we omit it.

DEFINITION. A class **R** of additively inversive hemirings is closed under extension or it has the *extension property,* if it satisfies the condition:

(e) *If B is an* **R**-*ideal of an additively inversive hemiring A such that $C\cong A/B\in\mathbf{R}$, then A is also an* **R**-*additively inversive hemiring.*

THEOREM 2.2. *Every radical class satisfies the extension property* (e). *Further* **R** *is a radical class if and only if* **R** *satisfies conditions* (a), (b) *and* (e).

Using Theorems 1.4 and 1.5 one can use analogous reasoning to that of ring theory (cf., e.g. [4] Theorem 3.3).

DEFINITION (cf. [3]). An element a of an additively inversive hemiring A is called *e-potent,* provided there is a positive integer n such that $a^n=e\in E_A$. A non-empty complex M of an additively inversive hemiring A is called an *e-set* provided every element of M is e-potent. A left (right, two-sided) ideal $\mathcal{A}$ of an additively inversive hemiring $\mathcal{A}$ is called *e-potent* provided there is a positive integer n such that $\mathcal{A}^n\subseteq E$.

REMARK. Any e-potent left (right, two-sided) ideal $\mathcal{A}$ of an additively inversive hemiring S is also an e-set. Any left (right, two-sided) ideal $\mathcal{B}\subseteq\mathcal{A}$ is e-potent.

Let **N** denote the class of all additively inversive hemirings S, where S is an e-set.

LEMMA 2.1. *The class* **N** *has the extension property* (e).

PROOF. Suppose B is an **N**-ideal of A such that $A/B \in \mathbf{N}$. We show that $A \in \mathbf{N}$. Let $a \in A$. Suppose $B = \bigcup_{e \in E} N_e$ and $a \in N_e + a$, where N_e denotes the additive subgroup of B with idempotent $e \in B$ as its identity.

Now $N_e + a \in A/B$ and $A/B \in \mathbf{N}$. Therefore there exists a positive integer n such that

$$(N_e + a)^n \in E.$$

But

$$(N_e + a)^n = N_{e^n} + a^n = N_f + a^n,$$

where $f = e^n \in E$. Thus

$$N_f + a^n \in E = \{N_e\}_{e \in E}.$$

Hence $N_f + a^n = N_g$ for some $g \in E$. Thus $a^n \in N_g \subseteq B$. Since B is **N**-additively inversive hemiring there exists a positive integer k such that $(a^n)^k = e \in E$, i.e. $a^{nk} = (a^n)^k = e \in E$. Thus $a \in A$ is an e-potent. Since a is an arbitrary element of A, A is an e-set. Hence $A \in \mathbf{N}$.

THEOREM 2.3. *The class* **N** *is a radical class.*

PROOF. Obviously, **N** is P-homomorphically closed.

The usual reasoning shows that condition (b) is also satisfied (cf. [4], Theorem 4.2).

By Lemma 2.1, the class **N** satisfies (e). Hence **N** satisfies (a), (b) and (e) so by Theorem 2.2, **N** is a radical class.

DEFINITION. An additively inversive hemiring A is called an *artinian additively inversive hemiring,* if every strictly descending chain of left ideals terminates in a finite number of steps.

LEMMA 2.2. *If A is an artinian additively inversive hemiring, then every* **N**-*left ideal of A is e-potent. In particular, the radical* **N**(A) *of an artinian additively inversive hemiring is e-potent.*

PROOF. We show that every non e-potent left ideal of A has a non e-potent element. Let $\mathscr{L}$ be a non e-potent left ideal of an artinian additively inversive hemiring A and consider the set M of all left ideals $\mathscr{J}$ of A such that $\mathscr{J} \subseteq \mathscr{L}$ and $\mathscr{L}\mathscr{J} \subsetneqq E$. Since $\mathscr{L}$ is not e-potent, $\mathscr{L}\mathscr{L} = \mathscr{L}^2 \subsetneqq E$, so $\mathscr{L} \in M$, i.e. M is nonempty. Choose a left ideal $\mathscr{J}$ which is minimal in M. Since $\mathscr{L}\mathscr{J} \subsetneqq E$, therefore there exists an element $j \in \mathscr{J}$ such that $\mathscr{L}j \subsetneqq E$. Note $j \notin E$, for otherwise $\mathscr{L}j \subseteq E$. Now $\mathscr{L}j$ is a left ideal of A and $\mathscr{L}j \subseteq \mathscr{J} \subseteq \mathscr{L}$ and $\mathscr{L}\mathscr{L}j \subsetneqq E$. Hence $\mathscr{L}j \in M$. Since $\mathscr{J}$ is a minimal left ideal of the class M, therefore $\mathscr{L}j = \mathscr{J}$. Thus there exists an element $a \in \mathscr{L}$ such that $aj = j$. Multiplying both sides from the left by a, we get $a^2 j = aj = j$ and by induction $a^n j = j$ and $j \notin E$. Thus $a \in \mathscr{L}$ is not e-potent. Thus we have shown that if $\mathscr{L}$ is not e-potent left ideal, then $\mathscr{L} \notin \mathbf{N}$. Since radical $\mathbf{N}(A)$ is an **N**-ideal of A, $\mathbf{N}(A)$ is an e-potent.

LEMMA 2.3. *The radical* $\mathbf{N}(A)$ *of an additively inversive hemiring* A *contains all e-potent one-sided ideals.*

PROOF. Suppose $\mathscr{L}$ is an e-potent left ideal of an additive inversive hemiring A. Then there exists a positive integer n such that $\mathscr{L}^n \subseteqq E$. By [3], Theorem 5, $\mathscr{L}+\mathscr{L}A$ is an e-potent two-sided ideal of A. Since the radical $\mathbf{N}(A)$ is a maximal e-potent ideal of A, therefore $\mathscr{L}\subseteq\mathscr{L}+\mathscr{L}A\subseteq\mathbf{N}(A)$. Similarly we can prove $\mathscr{R}\subseteqq\mathbf{N}(A)$ for every e-potent right ideal $\mathscr{R}$ of A.

COROLLARY. *If* A *is an artinian additively inversive hemiring, then its radical* $\mathbf{N}(A)$ *is the union of all e-potent left ideals of* A.

PROOF. By Lemma 2.2 $\mathbf{N}(A)$ is e-potent. Let U_l denote the union of all e-potent left ideals of A. Then $\mathbf{N}(A)\subseteqq U_l$. Let $\mathscr{L}\subseteqq U_l$, i.e. $\mathscr{L}$ is an e-potent left ideal of A. Then by Lemma 2.3, $\mathscr{L}\subseteqq\mathbf{N}(A)$, for all $\mathscr{L}\subseteqq U_l$. That is, $U_l\subseteqq\mathbf{N}(A)$. Hence $\mathbf{N}(A)=U_l$.

THEOREM 2.4. *If* A *is an additively inversive hemiring satisfying the maximal condition for left ideals, then there exists a unique maximal e-potent left ideal of* A*, which is* $\mathbf{N}(A)$.

PROOF. Let C denote the class of all e-potent left ideals of A. By the maximal condition, C contains a maximal member, say M. By Lemma 2.3 $M\subseteqq\mathbf{N}(A)$. Suppose $\mathbf{N}(A)\nsubseteqq M$, and let $m\in\mathbf{N}(A)\setminus M$. Now (m) is an e-potent left ideal and by [3], Theorem 5 the left ideal $(m)+M$ is e-potent. Now $(m)+M\supseteqq M$. As $m\in(m)+M$ and $m\notin M$, $M\subset(m)+M$. This contradicts the maximality of M. Thus $M=\mathbf{N}(A)$.

COROLLARY. *If* A *is an additively inversive hemiring satisfying either the maximal or minimal condition for left ideals, then* $\mathbf{N}(A)$ *is e-potent.*

In the following we shall assume that the set E of the additive idempotents of the considered additively inversive hemiring A, has the following property:

$$(*) \qquad ef = efe = fe \qquad \text{holds for every} \quad e, f\in E.$$

LEMMA 2.4. *Let* $\mathscr{I}$ *be an ideal of an additively inversive hemiring* A *with zero element and* E *having property* $(*)$. *Let* $\mathscr{K}$ *be an ideal of* $\mathscr{I}$ *and* $a\in A$, *further* $\mathscr{I}=\bigcup_{e\in E} L_e$. $\mathscr{K}=\bigcup_{e\in E} N_e$ *and* $a\mathscr{K}=\bigcup_{f\in E} G_f$ *where* $N_e=L_e\cap\mathscr{K}$, $G_f=L_f\cap a\mathscr{K}$ *and* F *is the trivial ideal of* $a\mathscr{K}$. *If* $\mathscr{K}^*=\bigcup_{f\in F} N_f$ *and* $\mathscr{I}^*=\bigcup_{f\in F} L_f$, *then* $\mathscr{I}\mathscr{K}^*+\mathscr{K}^*\mathscr{I}\subseteqq K^*$ *and* $a\mathscr{K}+\mathscr{K}^*$ *is an ideal of* $\mathscr{I}^*$ *and* $(a\mathscr{K}+\mathscr{K}^*)/\mathscr{K}^*$ *is a homomorphic image of* $\mathscr{K}$.

PROOF. Let $a+a'=e$. If $x_1, x_2\in\mathscr{K}^*$, then $x_1\in N_{f_1}$ and $x_2\in N_{f_2}$ for some $f_1, f_2\in F$ and so $x_1+x_2\in N_{f_1}+N_{f_2}=N_{f_1+f_2}\subseteqq\mathscr{K}^*$ as $f_1+f_2\in F$. Hence $\mathscr{K}^*$ as well as $a\mathscr{K}+\mathscr{K}^*$ is closed under addition. Obviously, $k\in a\mathscr{K}+\mathscr{K}^*$ implies $k'\in a\mathscr{K}+\mathscr{K}^*$. Let us consider elements $i\in\mathscr{I}, x_1\in\mathscr{K}$ and $x_2\in\mathscr{K}^*$ such that $i+i'=d, x_1+x_1'=g_1$ and $x_2+x_2'=g_2$. Then

$$i(ax_1+x_2) = (ia)x_1+ix_2 \subseteqq L_{de}N_{g_1}+$$
$$+N_d N_{g_2} \subseteqq N_{deg_1}+N_{dg_2} = N_{edg_1}+N_{g_2 d} \subseteqq \mathscr{K}^*$$

as E is commutative by * and $g_2 \in F$ by $x_2 \in K^*$. Also we have

$$(ax_1+x_2)i = a(x_1 i)+x_2 i \subseteqq aK+N_{g_2} \subseteqq a\mathscr{K}+\mathscr{K}^*.$$

Thus $\mathscr{I}(a\mathscr{K}+\mathscr{K}^*)+(a\mathscr{K}+\mathscr{K}^*)\mathscr{I} \subseteqq a\mathscr{K}+\mathscr{K}^*$ and $\mathscr{I}\mathscr{K}^*+\mathscr{K}^*\mathscr{I} \subseteqq \mathscr{K}^*$ hold. Since F is the common trivial ideal of $a\mathscr{K}+\mathscr{K}^*$ and $\mathscr{I}^*$, $a\mathscr{K}+\mathscr{K}^*$ is an ideal of $\mathscr{I}^*$.

Since $F \subseteqq a\mathscr{K} \cap \mathscr{K}^*$, we may consider the additively inversive factor hemiring $a\mathscr{K}/(a\mathscr{K} \cap \mathscr{K}^*)$. As $a\mathscr{K} = \bigcup_{f \in F} G_f$ and $\mathscr{K}^* = \bigcup_{f \in F} N_f$, we have $a\mathscr{K} \cap \mathscr{K}^* = \bigcup_{f \in F} M_f$ where $M_f = N_f \cap G_f$. Define $\phi\colon \mathscr{K} \to a\mathscr{K}/(a\mathscr{K} \cap \mathscr{K}^*)$ by $(x)\phi = ax+M_{ef}$ where $x \in \mathscr{K}$, $x+x'=f$. If $x, y \in \mathscr{K}, x+x'=f, y+y'=g$, then

$$(x+y)\phi = a(x+y)+M_{e(f+g)} =$$

$$= (ax+M_{ef})+(ay+M_{eg}) = (x)\phi+(y)\phi.$$

Also

$$(xy)\phi = axy+M_{efg} \subseteqq (a\mathscr{K} \cap L_{ef}N_g)+(a\mathscr{K} \cap \mathscr{K}^*) \subseteqq a\mathscr{K} \cap \mathscr{K}^* = \bigcup_{f \in F} M_f.$$

Thus the congruence class $ax+M_{efg}$ coincides with some congruence class M_f, $f \in F$. Let $z \in axy+M_{efg}$. As $z+z'=efg$, $axy+M_{efg}=M_{efg}$. Hence $(xy)\phi=M_{efg}$. Now

$$(x)\phi(y)\phi = (ax+M_{ef})(ay+M_{eg}) =$$

$$= axay+M_{efeg} = axay+M_{efg},$$

as E has property $(*)$. As $axay \subseteqq a\mathscr{K} \cap \mathscr{I}\mathscr{K}^* \subseteqq a\mathscr{K} \cap \mathscr{K}^*$,

$$axay+M_{efg} \subseteqq a\mathscr{K} \cap \mathscr{K}^* = \bigcup_{f \in F} M_f.$$

As above it can be proved that $axay+M_{efg}=M_{efg}$. Thus $(xy)\phi=(x)\phi(y)\phi$ and ϕ is a homomorphism. Now Theorem 1.6 is applicable for $T=a\mathscr{K}$ which gives

$$\phi\colon \mathscr{K} \to a\mathscr{K}/(a\mathscr{K} \cap \mathscr{K}^*) \cong (a\mathscr{K}+\mathscr{K}^*)/\mathscr{K}^*.$$

We finally show that the radical of an ideal of an additively inversive hemiring A is always an ideal of A, provided that the radical class $\mathbf{R}$ has the following properties:

(i) $\mathbf{R}$ is homomorphically closed, that is $A \in \mathbf{R}$ implies $(A)\phi \in \mathbf{R}$ for any homomorphism ϕ of A;

(ii) Let E denote the trivial ideal of A and $F=eE$ where $e \in E$. If

$$A = \bigcup_{e \in E} \mathscr{G}_e \quad \text{and} \quad A^{(e)} = \bigcup_{f \in F} \mathscr{G}_f,$$

then

$$\mathbf{R}(A^{(e)}) = \mathbf{R}(A) \cap A^{(e)}.$$

Note that the previously considered radical class $\mathbf{N}$ possesses both properties (i) and (ii), meanwhile $\mathbf{N}$ has neither (i) nor (ii).

THEOREM 2.5. *Let* $\mathbf{R}$ *be a radical class having properties* (i) *and* (ii). *If* $\mathscr{I}$ *is an ideal of an additively inversive hemiring* A *with zero element and* E *having property* $(*)$, *then* $\mathbf{R}(\mathscr{I})$ *is an ideal of* A.

PROOF. Suppose that $\mathcal{K}=\mathbf{R}(\mathcal{I})$ is not an ideal of A. Since $\mathcal{K}$ is additively closed and also $E\subseteqq\mathcal{K}$, there exists an element $a\in A$ such that $a\mathcal{K}\subsetneqq\mathcal{K}$ or $\mathcal{K}a\subsetneqq a$. Assume that $a\mathcal{K}\subsetneqq\mathcal{K}$. Denoting the trivial ideal of $a\mathcal{K}$ by F, which is exactly $F=eE, a+a'=e$. From $a\mathcal{K}\subsetneqq\mathcal{K}=\bigcup_{e\in E} N_e$ it follows $a\mathcal{K}\subsetneqq\bigcup_{f\in F} N_f=\mathcal{K}^{(e)}$. Hence $(a\mathcal{K}+\mathcal{K}^{(e)})/K^{(e)}$ is a nontrivial P-homomorphic image of $(a\mathcal{K}+\mathcal{K}^{(e)})/\mathcal{K}^{(e)}$. Applying Lemma 2.4 we obtain that $(a\mathcal{K}+\mathcal{K}^{(e)})/\mathcal{K}^{(e)}$ is a homomorphic image of $\mathcal{K}=\mathbf{R}(\mathcal{I})\in\mathbf{R}$. Hence (i) $(a\mathcal{K}+\mathcal{K}^{(e)})/\mathcal{K}^{(e)}\in\mathbf{R}$. Since

$$\mathcal{K}^{(e)}=\bigcup_{f\in F} N_f=\Big(\bigcup_{e\in E} N_e\Big)\cap\Big(\bigcup_{f\in F} L_f\Big)=\mathcal{K}\cap\mathcal{I}^{(e)}$$

where $\mathcal{I}=\bigcup_{e\in E} L_e$, we get

$$\mathcal{I}^{(e)}/\mathcal{K}^{(e)}=\mathcal{I}^{(e)}/\mathcal{K}\cap\mathcal{I}^{(e)}=\mathcal{I}^{(e)}/(\mathbf{R}(\mathcal{I})\cap\mathcal{I}^{(e)})=\mathcal{I}^{(e)}/\mathbf{R}(\mathcal{I}^{(e)}),$$

by property (ii). Thus by Theorem 2.1 (c) $\mathcal{I}^{(e)}/\mathcal{K}^{(e)}$ does not contain any non-trivial $\mathbf{R}$-ideal, contradicting the properties of $(a\mathcal{K}+\mathcal{K}^{(e)})/\mathcal{K}^{(e)}$. Thus $a\mathcal{K}\subseteqq\mathcal{K}$ holds. Similarly we get $\mathcal{K}a\subseteqq\mathcal{K}$ and so $\mathcal{K}=\mathbf{R}(\mathcal{I})$ is an ideal of A.

COROLLARY. *Let* $\mathbf{R}$ *be a radical class having properties* (i) *and* (ii). *If* $\mathcal{I}$ *is an ideal of an additively inversive hemiring* A *with zero element and* E *having property* $(*)$, *then* $\mathbf{R}(\mathcal{I})\subseteqq\mathbf{R}(A)\cap\mathcal{I}$. *In particular,* $\mathbf{R}(A)=E$ *implies* $\mathbf{R}(\mathcal{I})=E$.

REFERENCES

[1] CLIFFORD, A. H. and PRESTON, G. B., *The algebraic theory of semigroups,* Vol. I, Mathematical Surveys, No. 7, American Mathematical Society, Providence, R. I., 1961. *MR* **24** # A2627.

[2] YUSUF, S. M., Ideals in additively inversive semirings, *J. Natur. Sci. and Math.* **5** (1965), 45—56. *MR* **32** #4164.

[3] YUSUF, S. M., The classical radical of an additively inversive semiring, *J. Natur. Sci. and Math.* **5** (1965), 57—69. *MR* **32** #4165.

[4] WIEGANDT, R., *Radical and semisimple classes of rings,* Queen's Papers in Pure and Applied Mathematics, No. 37, Queen's University, Kingston, Ont., 1974. *MR* **50** #2227.

(Received October 24, 1978)

DEPARTMENT OF MATHEMATICS
QUAID-I-AZAM UNIVERSITY
ISLAMABAD
PAKISTAN

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 311—314

ÜBER DIE ANZAHL DER ÄQUIVALENZRELATIONEN DER ENDLICHEN MENGE

von

K. TODOROV

PASCUAL JORDAN

zum 80. Geburtstag

Mit den Stirlingschen Zahlen 2. Art

$$S(n, m) = \frac{1}{m!} \sum_{k=0}^{m} (-1)^{m-k} \binom{m}{k} k^n, \tag{1}$$

die die Anzahl aller Äquivalenzrelationen über einer endlichen Menge von n Elementen mit genau m disjunkten Teilmengen (Klassen) bestimmen, hängt die Lösung einer Reihe von Problemen der Mengentheorie [2], [4], der mathematischen Logik [1], der Halbgruppentheorie [6], der Wahrscheinlichkeitstheorie [3] u.a. zusammen.

Im vorliegenden Artikel sind die Formeln (2) für die Stirlingschen Zahlen $S(n, m)$ gegeben. Davon sind die Sätze 2 und 3 (Formeln (3) und (4)) abgeleitet. Es kommen dabei lediglich algebraische Methoden in Frage.

Hierbei sind für große $m \left(m > \dfrac{n}{2}\right)$ die aus der Formel (2) bestimmten Ausdrücke für $S(n, m)$ relativ einfächer als die aus der Formel (1) bestimmten.

Wir lassen uns möglichst von der in der Literatur üblichen Terminologie leiten. Weiter sei die Menge $X_n = \{a_1, ...; a_n\}$ mit $n > 1$ verschiedenen Elementen gegeben. Beim Rechnen können wir annehmen, daß $X_n = \{1, 2, ..., n\}$ ist, da die konkrete Beschaffenheit der Elemente keine Bedeutung hat.

LEMMA 1. $|f_m(\pi)|$ *sei die Anzahl aller Äquivalenzrelationen* π *der lineargeordneten Menge* $X_n(<) = \{a_1 < a_2 < ... < a_n\}$ *mit m Klassen, deren minimalen Elemente ein und dieselbe Folge*

$$f_m(\pi)\colon a_{i_1}, ..., a_{i_m}$$

bilden und sei

$$X_n = \{a_{i_1}, ..., a_{i_m}\} \cup \{a_{j_1}, ..., a_{j_k}\},$$

wobei

$$1 = i_1 < i_2 < ... < i_m, \; j_1 < j_2 < ... < j_k,$$

und $m + k = n$ *sind. Dann gilt*

$$|f_m(\pi)| = (j_k - k)(j_{k-1} - k + 1) ... (j_1 - 1).$$

Die Forschungsarbeit wurde mit Unterstützung der Alexander von Humboldt-Stiftung erfüllt.
1980 *Mathematics Subject Classification.* Primary 05A17; Secondary 10G15.
Key words and phrases. Equivalence relation, Stirling numbers of 2nd kind.

BEWEIS. Es sei $f_m(\pi)$ eine festgehaltende Folge von m Elementen der Menge X_n. Dann ist die Anzahl der Äquivalenzrelationen π über der Menge X_n, für die $f_m(\pi)$ eine Folge mit den Eigenschaften von Lemma 1,

$$a_{j_1} \equiv a_{i_t} \pmod{\pi} \quad \text{und} \quad i_t < j_1$$

sind, gleich j_1-1. Jeder von diesen j_1-1 Möglichkeiten entsprechen neue j_2-2 Möglichkeiten, für die

$$a_{j_2} \equiv a_{i_k} \pmod{\pi}, \quad i_k < j_2, \quad \text{und} \quad i_k \neq j_1$$

sind. Durch Induktion folgt, daß

$$|f_m(\pi)| = (j_k-k)(j_{k-1}-k+1)\dots(j_1-1)$$

ist.

SATZ 1. *Die Äquivalenzanzahl $S(n,m)$ der Menge von n Elementen mit m Äquivalenzklassen wird nach der Formeln*

$$(2) \qquad S(n,m) = \sum_{j_k=k+1}^{n} \sum_{j_{k-1}=k}^{j_k-1} \dots \sum_{j_2=3}^{j_3-1} \sum_{j_1=2}^{j_2-1} (j_k-k)(j_{k-1}-k+1)\dots(j_1-1)$$

bestimmt, wobei $k=n-m\geqq 1$ und $S(n,n)=1$ sind.

BEWEIS. Es seien $X_n(<)=\{a_1<a_2<\dots<a_n\}$; A_m die Menge aller Äquivalenzrelationen π über der Menge X_n, für die $|X/\pi|=m$ ist. Dann bezeichnen wir für eine bestimmte Äquivalenzrelation $\pi\in A_m$ mit

$$f_m(\pi)\colon a_{i_1}, \dots, a_{i_m}$$

die Folge der minimalen Elemente aller Äquivalenzklassen von π.

Über der Menge A_m definieren wir aufs neue eine binäre Relation θ durch die Regel: $\pi_1\sim\pi_2\Leftrightarrow f_m(\pi_1)=f_m(\pi_2)$. Es ist evident, daß θ eine Äquivalenzrelation über A_m ist, daß jede Äquivalenzrelation π mit m Klassen von X_n eine Folge $f_m(\pi)$ eindeutig definiert und je zwei verschiedene Äquivalenzrelationen von X_n, die zu ein- und derselben Klasse θ gehören, auch verschiedene Äquivalenzrelationen von $X_n(<)$ sind.

Also haben wir nach den Bezeichnungen von Lemma 1, daß

$$|f_m(\pi)| = (j_k-k)(j_{k-1}-k+1)\dots(j_1-1).$$

Ferner folgt aus den Bedingungen

$$1 = i_1 < i_2 < \dots < i_m, \quad j_1 < j_2 < \dots < j_k,$$

daß für alle $r=1, \dots, k$

$$1 = i_1 < j_1 < \dots < j_{r-1} < j_r$$

ist, d. h., es gelten für $r=1,2,\dots,k-1$ die Ungleichungen

$$r < j_r, \quad r < r+1 \leqq j_r \leqq j_{r+1}-1$$

und

$$k+1 \leqq j_k \leqq n.$$

Da die Menge A_m die Summe aller Klassen von A_m/θ ist, ist also die Anzahl $S(n, m)$ der Elemente der Menge A_m nach der Formel

$$S(n, m) = \sum_{j_k=k+1}^{n} \sum_{j_{k-1}=k}^{j_k-1} \dots \sum_{j_2=3}^{j_3-1} \sum_{j_1=2}^{j_2-1} |f_m(\pi)|$$

bestimmt, q.e.d.

Offenbar ist $S(n, n)=1$.

Wie es am Anfang betont würde, sind für große m die aus der Formel (2) bestimmten Ausdrücke für $S(n, m)$ relativ einfacher als die aus der Formel (1) bestimmten. Z. B. haben wir

$$S(n, n-1) = \frac{1}{(n-1)!} \sum_{k=0}^{n-1} (-1)^{n-1-k} \binom{n-1}{k} k^n$$

aus der Formel (1) und

$$S(n, n-1) = \sum_{j_1=2}^{n} (j_1-1) = \binom{n}{2}$$

aus der Formel (2).

SATZ 2. *Es seien U_1 eine bestimmte Untermenge der Menge X_n: $|U_1|=n_1$ und m, r natürliche Zahlen, für die die Ungleichungen $1 \leqq r \leqq m \leqq n = |X_n|$ erfüllt sind. Dann gilt:*

Die Anzahl $S(n_1, n; r, m)$ der Äquivalenzrelationen π über X_n, für welche die Gleichungen $|U_1/\pi|=r$ und $|X_n/\pi|=m$ gelten, ist durch die Formel

$$S(n_1, n; r, m) = \sum_{j_k=k+1}^{n} \dots \sum_{j_{t+1}=n_1+1}^{j_{t+2}-1} \sum_{j_t=t+1}^{n_1} \dots \sum_{j_1=2}^{j_2-1} (j_k-k) \dots (j_1-1), \tag{3}$$

wobei $t=m-r$, $k=n-m$ sind, gegeben.

BEWEIS. Wir können annehmen, daß die Ungleichung $a_u < a_v$ für alle Elemente $a_u \in U_1$ und $a_v \in X_n \setminus U_1$ gilt. Weiter sei π eine Äquivalenzrelation über X_n mit m Klassen, deren minimale Elemente die Folge

$$f_m(\pi)\colon\ 1, a_{i_2}, \dots, a_{i_r}, a_{i_{r+1}}, \dots, a_{i_m},$$

wo

$$a_{i_1} = 1,\ a_{i_2}, \dots, a_{i_r} \in U_1 \quad \text{und} \quad a_{i_{r+1}}, \dots, a_{i_m} \in X_n \setminus U_1,$$

bilden. Wie im Beweis des Satzes 1, bezeichnen wir hier mit $a_{j_1}, \dots, a_{j_k}$ die restlichen Elemente von $f_m(\pi)$ der Menge X_n so, daß $a_{j_{t+1}}, \dots, a_{j_k} \in X_n \setminus U_1$.

Dann enthält die Mannigfaltigkeit aller solchen Äquivalenzrelationen nach Lemma 1 $|f_m(\pi)|$ Elemente, wo

$$|f_m(\pi)| = (j_k-k) \dots (j_{t+1}-t-1)(j_t-t) \dots (j_1-1)$$

ist.

SATZ 3. *Es seien $U_2 \subset U_1 \subset X_n$; $|U_1|=n_1$, $|U_2|=n_2 \geqq r$. Dann wird die Anzahl $S(n_2, n_1, n; r, m)$ der Äquivalenzrelationen über X_n, für die die Gleichungen*

$$|U_2/\pi| = |U_1/\pi| = r, \quad |X_n/\pi| = m$$

erfüllt sind, durch die Formel

$$S(n_2, n_1, n; r, m) = r^{n_1-n_2} S(n_2, n-n_1+n_2; r, m) \tag{4}$$

gegeben.

BEWEIS. Die Behauptung folgt leicht aus den vorhergehenden Beweisen der Sätze 1 und 2, wenn wir beachten, daß die Anzahl der Äquivalenzrelationen π über X_n, für die die Gleichungen

$$|U_2/\pi| = r, \quad |(X_n \setminus (U_1 \cup U_2))/\pi| = |((X_n \setminus U_1) \cup U_2)/\pi| = m$$

gelten, $S(n_2, n-n_1+n_2; r, m)$ ist. Dabei kann man jedes Element der Menge $U_1 \setminus U_2$ mit beliebigen Elementen der Menge U_2 zufügen.

Herrn B. SCHLENDER sei für den Hinweis auf das Buch [6] des Literaturverzeichnisses gedankt.

LITERATURVERZEICHNIS

[1] BAĬRAMOW, R. A., Über den Untergruppenverband der symmetrischen Gruppe, den Unterhalbgruppenverband der symmetrischen Halbgruppe und den Unteralgebrenverband der Postschen Algebra, *Semigroup Forum* **2** (1971), 271—280. *MR* **46** #9177.

[2] ORE, O., Theory of equivalence relations, *Duke Math. J.* **9** (1942), 573—627. *MR* **4**—128.

[3] RIORDAN, J., *An introduction to combinatorial analysis*, Wiley Publications in Mathematical Statistics, John Wiley & Sons, Inc., New York; Chapman & Hall, Ltd., London 1958. *MR* **20** #3077.

[4] SIERPIŃSKI, W., *Algèbre des ensembles*, Monografie Matematyczne. Tom XXIII, Polskie Towarzystwo Matematyczne, Warszawa—Wrocław, 1951. *MR* **13**—541.

[5] TODOROV, K., Über die Unterhalbgruppen der symmetrischen Halbgruppe mit fixiertem Wertbereich und ihre Ordnungen, *C. R. Acad. Bulgare Sci.* **29** (1976), 299—301.

[6] *Handbook of mathematical functions, with formulas, graphs, and mathematical tables*, edited by Milton Abramowitz and Irene A. Stegun, Third printing, with corrections, National Bureau of Standards Applied Mathematics Series, 55. Superintendent of Documents, U. S. Government Printing Office, Washington, D. C., 1965. *MR* **31** #1400.

(Received November 10, 1978)

MATHEMATISCHE FAKULTÄT
DER UNIVERSITÄT
BUL. „A. IVANOV" 5
BG—1126 SOFIA
BULGARIA

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 315—317

A COMMENT ON AUTOMORPHISM GROUPS OF FIELDS

by

E. FRIED

In our paper [2] we proved that every group is isomorphic to the automorphism group of a suitable field. The fields we constructed were very special, e.g., they did not contain any algebraic elements outside their prime fields.

Naturally arises the following

PROBLEM I. Given a field K and a group G, does there exist a field L containing K the automorphism group of which is isomorphic to G?

I conjecture that the answer is affirmative and as in the case of integral domains (see [1]) one can choose L such that the elements of K are fixed by every automorphism.

In fact, I think that fixing K is not a further stronger result but this is "the way" how to prove the original statement. A further relevant question is the following:

Let K be a field and G a group. Does there exist an extension L of K the automorphism group of which is isomorphic to G such that K consists of the fixpoints of the automorphism group?

As easily seen, such an extension is necessarily Galois. If K is the field of rationals and G is finite then this gives a famous unsolved problem of E. NOETHER. I. SAFAREVITCH proved that for solvable groups the answer is positive. If K is the field of complex numbers and G is finite then the answer is, of course, negative. J. KOLLÁR mentioned to me, that if K is the field of rational functions in one indeterminate over the complex numbers then one can prescribe the group of relative automorphisms for finite groups. The proof uses the methods of algebraic geometry [3]. These show that the question in this generality is too difficult to handle.

However, one can weaken this problem in another direction when not the "starting field" but its automorphism group is given. Let us mention that M. ADAMS and J. SICHLER [4] have obtained results on the analogous problem for lattices and posets.

Let Aut K denote the automorphism group of K and Fix K the set of fixpoints of Aut K.

1980 *Mathematics Subject Classification.* Primary 18B15; Secondary 12F99.
Key words and phrases. Fields, automorphism groups, graphs, group universal categories.

PROBLEM II. Let the groups G, H be given. Does there exist a field K such that $\mathrm{Aut}\, K \cong G$ and $\mathrm{Aut\, Fix}\, K \cong H$?

The answer is, in general, negative, for if G has one element only then H must have, also, one element.

CONJECTURE. There is an affirmative answer for Problem II, provided G has more than one element.

The following theorem gives a partial answer to Problem II.

THEOREM. *Given a group H and an infinite group G there exists a field K such that $G \cong \mathrm{Aut}\, K$ and $H \cong \mathrm{Aut\, Fix}\, K$.*

In the proof of this result we shall make use of the following construction given in [2].

Let $\mathcal{G}=(V, E)$ be any (undirected) graph where V denotes the set of vertices and E denotes the set of edges. We may view V as a set of indeterminate over the field Q of rationals. We considered, also, auxiliary indeterminated $t_1, \ldots, t_n, \ldots$ indexed by the natural numbers. Let L be the algebraic closure of $Q(V, V_1; \ldots, t_n, \ldots)$ and we consider a subfield $K(\mathcal{G})$ of L as follows:

We choose a sequence of different odd prime numbers $p, p_1, \ldots, p_n, \ldots$. To each x in V and natural number k we choose exactly one y in L such that $y^{p^k}=x$ and $y^{p^{k-1}}$ is already chosen.

Also, to each (x_i, x_j) in E and to each natural number k we choose a z in L such that $z=x_i+x_j$ for $k=0$ and for positive k we have $z^2=t_k-z'$ where z' was chosen for the same edge and for the number $k-1$.

The field $K(\mathcal{G})=Q(\ldots, y, \ldots, \ldots, z, \ldots)$ has the property that $\mathrm{Aut}\, K(\mathcal{G}) \cong \cong \mathrm{Aut}\,(\mathcal{G})$.

The above construction would suggest the following solution for Problem II. Let $\mathcal{G}_2$ be a subgraph of $\mathcal{G}_1$ such that $\mathrm{Aut}\, \mathcal{G}_1 \cong G$, $\mathrm{Aut}\, \mathcal{G}_2 \cong H$ and the vertices of $\mathcal{G}_2$ are exactly $\mathrm{Fix}\, \mathcal{G}_1$ the fixpoints of $\mathrm{Aut}\, \mathcal{G}_1$. In this case $\mathrm{Fix}\, K(\mathcal{G}_1)=K(\mathcal{G}_2)$ would, clearly, imply an affirmative answer. However, $\mathrm{Fix}\, K(\mathcal{G})=K(\mathrm{Fix}\, \mathcal{G})$ is not true in general. Indeed, suppose G is finite. Then, $K(\mathcal{G})$ is an algebraic extension of $\mathrm{Fix}\, K(\mathcal{G})$, while each vertex not belonging to $\mathrm{Fix}\, \mathcal{G}$ is transcendental over $K(\mathrm{Fix}\, \mathcal{G})$. Therefore our Theorem seems to be the best possible result one can prove using the above construction.

PROOF of the Theorem. It is easy to construct a graph $\mathcal{G}$ such that $\mathrm{Aut}\, \mathcal{G} \cong G$, $\mathrm{Aut\, Fix}\, \mathcal{G} \cong H$ and the orbits of any vertex not belonging to $\mathrm{Fix}\, \mathcal{G}$ is infinite. So, it is enough to prove that $\mathrm{Fix}\, K(\mathcal{G})=K(\mathrm{Fix}\, \mathcal{G})$. The right-hand side is, clearly, contained in the left-hand side. Now, any α of the field-extension belongs to the algebraic closure of some $Q(x_1, \ldots, x_r, t_1, \ldots, t_s)$. Let $\{x_1, \ldots, x_r, t_1, \ldots, t_s\}$ be a minimal set of variables on which α is algebraically dependent. This set is uniquely determined because of the exchange property of algebraic dependence. Now for any automorphism σ, the element $\sigma\alpha$ is algebraically dependent on $\{\sigma x_1, \ldots, \sigma x_r, \sigma t_1, \ldots, \sigma t_s\}$. Therefore if α is a fixed element then the orbits of $x_1, \ldots, x_r$ are finite hence they are fixed elements. This proves the theorem.

REFERENCES

[1] FRIED, E., Automorphism group of integral domains fixing a given subring, *Algebra Universalis* **7** (1977), 373—387. *MR* **56** #2985.

[2] FRIED, E. and KOLLÁR, J., Automorphism groups of fields, *Universal Algebra* (Colloq. Math. Soc. J. Bolyai **29**, Esztergom, 1977), North-Holland Publishing Co., Amsterdam—Oxford—New York, 1982, 293—303.

[3] LANG, S., *Algebra*, Addison-Wesley Publishing Co., Inc., Reading, Mass., 1965 (VIII, § 2, Example 6). *MR* **33** #5416.

[4] ADAMS, M. E. and SICHLER, J., Automorphisms of poset and lattice extensions (preprint).

(Received September 4, 1979)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
ALGEBRA ÉS SZÁMELMÉLET TANSZÉK
MÚZEUM KRT. 6—8.
H—1088 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 319—323

A NOTE ON POINTWISE LIMITS OF FUNCTIONS WHICH ARE APPROXIMATELY CONTINUOUS AND ALMOST EVERYWHERE CONTINUOUS

by

G. V. COX and P. D. HUMKE

1. Introduction

In a recent paper, [G], Z. GRANDE supplies a condition, that we call AP_1 as in [N], which is necessary in order that a function be the pointwise limit of functions which are both approximately continuous and almost everywhere (a.e.) continuous. Subsequently, he asks whether AP_1 is also sufficient provided that the function in question is separately the pointwise limit of each of the above mentioned classes. In [N], T. NISHIURA continued the AP_1 investigation from a topological perspective and used a topology which is between the usual topology and the density topology and for which the continuous functions are exactly those functions which are both approximately continuous and a.e. continuous. This topology, $\mathcal{T}$, was first described by R. O'MALLEY [O'M], where it was called the almost everywhere topology. He then used this topology to define a new and quite technical condition, AP_2, which when coupled with AP_1, is sufficient to conclude that a function which is separately the pointwise limit of approximately continuous functions and the pointwise limit of a.e. continuous functions is also the pointwise limit of functions enjoying both properties. And again, Nishiura asks whether AP_1 alone is sufficient. In this note we show that AP_1 is sufficient for characteristic functions of countable sets and related functions. While the result was known to Nishiura, the arguments given here add insight to the property AP_1, as well as to the topology, $\mathcal{T}$. The method of proof is as follows: First we define a new condition (appropriately designated AP_3) which is easy to handle but proves to be equivalent to AP_1. We then adopt Nishiura's topological perspective and show that AP_3 is entailed by the usual notion of scattered in the appropriate topological space, $(R^n, \mathcal{T})$. Finally we show that scattered sets in this space are A_δ sets, where A denotes the collection of co-zero sets of $\mathcal{T}$-continuous functions and from this are able to conclude that characteristic functions of countable sets which are AP_3 are also AP_2.

2. Preliminaries

Although the basic charge of this discussion concerns functions defined on R^n, Theorem 1 is true in a quite general setting and is treated as such. In addition to the usual topology on R^n, we will also work with O'Malley's almost everywhere topology which Nishiura renamed as the $\mathcal{T}$-topology. The topology, $\mathcal{T}$, consists of sets of the

---

AMS (MOS) subject classifications (1980). Primary 26A15; Secondary 54C30.
Key words and phrases. Baire functions, density topology, almost everywhere continuous, approximately continuous, scattered.

form $G\cup Z$ where G is open in the usual sense, Z is of measure zero, and every point of Z is a point of full density of G. In [N], and [O'M] it is shown that $\mathscr{T}$ is a completely regular, Hausdorff topology which contains the usual topology and is contained in the density topology. Let A denote the collection of co-zero sets of continuous real valued functions on $(R^n, \mathscr{T})$. If $X\subset R^n$, we will denote the usual closure of X by $\overline{X}$ and the $\mathscr{T}$-closure by $\mathrm{cl}\,(X)$; if X is measurable, $\Delta(X)$ will denote the set of points at which X has positive upper density. Finally, X is said to be *scattered* (relative to a specified topology) if $X_0=X$ and

$$X_{\alpha+1}=\{x\in X_\alpha\colon\ x \text{ is a limit point of } X_\alpha\}$$

and

$$X_\alpha=\bigcap_{\beta<\alpha} X_\beta \text{ if } \alpha \text{ is a limit ordinal,}$$

then $X_\alpha=\emptyset$ for some $\alpha<\omega_1$.

The sets of functions concerning us here are $\mathscr{A}$, the set of approximately continuous functions and $\mathscr{P}$, the set of a.e. continuous functions, and $\mathscr{B}_1(\mathscr{C})$, the set of pointwise limits of functions in a specified class $\mathscr{C}$. For consistency we use Nishiura's notation in defining the properties AP_1 and AP_2 below (AP_1 was originally defined in [G]). Property AP_3 belongs to this work and is later shown to be equivalent to AP_1.

A function $f\in\mathrm{AP}_1$ (or has property AP_1) if whenever $a<b$ and U and V are nonempty sets satisfying

i. $U\subset E_a\ (E_a=\{x\colon f(x)<a\})$

ii. $V\subset E^b\ (E^b=\{x\colon f(x)>b\})$

iii. $U\subset\Delta(\overline{U})$ and $V\subset\Delta(\overline{V})$,

then either $U\setminus\overline{V}\neq\emptyset$ or $V\setminus\overline{U}\neq\emptyset$.

A function $f\in\mathrm{AP}_2$ if whenever $a<b$ and F is a closed set with $\mu(F)=\mu(F\cap E_a)<+\infty$ (μ=Lebesgue measure) it follows that the set $W=F\cap(R^n\setminus E_b)$ has the property that $(\overline{S}\setminus\overline{W\cap S})\setminus(R^n\setminus E_b)$ is A_δ where $S=\Delta(\overline{W})$.

A function $f\in\mathrm{AP}_3$ if whenever $a<b$ and U and V are nonempty sets satisfying both $U\subset E_a$ and $V\subset E^b$, then $U\cup V\not\subset\Delta(\overline{U}\cap\overline{V})$.

Now we give several background theorems.

THEOREM N_1 (Theorem 2.4, [N]). *The point x is a $\mathscr{T}$-limit point of X if and only if $x\in\Delta(\overline{X})$.*

(Thus, when considering $\mathscr{T}$-scattered sets, $X_{\alpha+1}=X\cap\Delta(\overline{X}_\alpha)$.)

THEOREM N_2 (Theorem 3.4, [N]). *The family $\mathscr{A}$ is closed under countable unions and is a basis for $\mathscr{T}$.*

THEOREM N_3 (Theorem 5.1, [N]). *If $f\in\mathscr{B}_1(\mathscr{A})\cap\mathscr{B}_1(\mathscr{P})$, then $f\in\mathscr{B}_1(\mathscr{A}\cap\mathscr{P})$ if and only if*

i. *$f\in\mathrm{AP}_1$, and*
ii. *f and $-f$ belong to AP_2.*

3. The main results

We first show that AP_1 and AP_3 determine the same set of functions.

LEMMA 1. *A function* $f \in AP_1$ *if and only if* $f \in AP_3$.

PROOF. Suppose that $f \notin AP_1$. Thus for some $a<b$, there are nonempty sets U and V such that $U \subset E_a$, $V \subset E^b$, $U \subset \Delta(\overline{U})$, $V \subset \Delta(\overline{V})$, and both $U \setminus \overline{V} = \emptyset$ and $V \setminus \overline{U} = \emptyset$. But, then $\overline{U} = \overline{V}$ so that $U \cup V \subset \Delta(\overline{U} \cap \overline{V})$ and it follows that $f \notin AP_3$.

Now, if $f \notin AP_3$ then for some $a<b$ there are nonempty sets U and V contained respectively in E_a and E^b such that $U \cup V \subset \Delta(\overline{U} \cap \overline{V})$. From this last set inequality it follows that each of U and V is contained in both $\Delta(\overline{U})$ and $\Delta(\overline{V})$ and consequently that both $U \setminus \overline{V} = \emptyset$ and $V \setminus \overline{U} = \emptyset$. Thus, $f \notin AP_1$.

LEMMA 2. *Suppose* $f: R^n \to R$ *and that* $S = E^0 \cup E_0$ *is countable. Then* $f \in AP_3$ *if and only if for every* $b>0$, E^b *and* E_{-b} *are* $\mathcal{T}$*-scattered.*

PROOF. Suppose that $f \in AP_3$. Now, if V is any nonempty subset of E^b (or E_{-b}) then we must show that $V \cap \Delta(\overline{V}) \neq V$. But if $V \cap \Delta(\overline{V}) = V$ then U could be chosen to be a dense subset of $\Delta(\overline{V}) \setminus S$. (Note that $\overline{V}$ has positive measure in each of its relative open sets and S is countable.) We would then have (for $b = a/2$) that $U \subset E_a$, $V \subset E^b$, and $U \cup V \subset \Delta(\overline{U} \cap \overline{V}) = \Delta(\overline{V})$. This, of course, contradicts AP_3 and it follows that E^b (or E_{-b}) is $\mathcal{T}$-scattered.

Suppose now that $f \notin AP_3$. Then there are $a<b$ and nonempty sets $U \subset E_a$ and $V \subset E^b$ such that $U \cup V \subset \Delta(\overline{U} \cap \overline{V})$. Either $a<0$ or $b>0$ and for definiteness we suppose the latter. It follows that $V \subset \Delta(\overline{V}) \subset \Delta(E^b)$ and similarly, for every ordinal α, $V \subset (E^b)_\alpha$. This, of course, implies that E^b is not $\mathcal{T}$-scattered and the lemma is proved.

Note that the latter half does not require that S be countable, and that we also have the following

LEMMA 3. *If S is countable, then* χ_S *has* AP_3 *if and only if S is* $\mathcal{T}$*-scattered.*

In what remains, we will investigate the relationships between the properties listed below:

1. S is $\mathcal{T}$-scattered;
2. S is $\mathcal{T}$-G_δ,
3. S is $\mathcal{T}$-A_δ.

THEOREM 1. *If $(X, \mathcal{T})$ is a first countable, regular topological space and $S \subset X$ is countable, then* $1 \Rightarrow 2$. *If, in addition, $(X, \mathcal{T})$ is completely regular, then* $2 \Leftrightarrow 3$ *also holds.*

PROOF. In order to prove that $1 \Rightarrow 2$, we let $S = \{s_1, s_2, \dots\}$ and let $\{U_k^n\}_{n=1}$ be a base of neighborhoods of s_k for every natural number k. Then, for each fixed k there is an ordinal α such that $s_k \in S_\alpha \setminus S_{\alpha+1}$. Let G_k be a neighborhood of s_k such that

[*] $\qquad \overline{G}_k \cap S_\alpha = \{s_k\} \quad (\overline{G}_k \equiv \text{closure of } G_k \text{ in } (X, T)).$

For every pair of naturals k and n, there is an open set G_k^n such that

[**] if $i\in\{1, 2, \dots\}$ and $s_k\notin \bar{G}_i$, then $G_k^n\cap \bar{G}_i\neq\emptyset$,

and

[***] $s_k\in G_k^n\subset G_k\cap U_k^n$.

We prove that $S=\bigcap_{n=1}^{\infty}\bigcup_{k=1}^{\infty} G_k^n$. Obviously, $S\subset\bigcap_{n=1}^{\infty}\bigcup_{k=1}^{\infty} G_k^n$. Suppose that $x\in\bigcap_{n=1}^{\infty}\bigcup_{k=1}^{\infty} G_k^n$, and let α be the first ordinal such that there exists $s_{k_0}\in S_\alpha\setminus S_{\alpha+1}$ with $x\in G_{k_0}$. Let $n\geqq k_0$ and suppose $x\in G_j^n$. Let $s_j\in S_\alpha\setminus S_{\beta+1}$. It follows that $\beta\geqq\alpha$ and thus $s_j\in S_\alpha$. If $i\neq k_0$, then $s_j\notin \bar{G}_{k_0}$ by [*]. Consequently, [**] implies that $G_j^n\cap\bar{G}_{k_0}=\emptyset$ which is a contradiction. Therefore, $x\in G_{k_0}^n$ for every $n>k_0$ and hence by [***], $x\in\bigcap_{n=k_0}^{\infty} G_{k_0}^n\subset\bigcap_{n=k_0}^{\infty} U_{k_0}=\{s_{k_0}\}$.

Now suppose that $(X, \mathcal{T})$ is completely regular. To prove that $2\Rightarrow 3$, we let $S=\{s_1, s_2, \dots\}=\bigcap_{n=1}^{\infty} G_n$, where each G_n is open. Let U_k^n be a cozero set such that $s_k\in U_k^n\subset G_n$, for each pair k, n of natural numbers. Since the class of cozero sets is closed under countable unions, the set $U_n=\bigcup_{k=1}^{\infty} U_k^n$ is a cozero set for every n, and $S=\bigcap_{n=1}^{\infty} U_n$.

The implication that $3\Rightarrow 2$ is immediate, and this completes the proof.

LEMMA 4. *If* $\emptyset\neq S\subset R^n$ *is countable and* $\mathcal{T}-G_\delta$, *then* $S\cap\Delta(\bar{S})\neq S$.

PROOF. As S is $\mathcal{T}-G_\delta$, $S=(G_1\cup Z_1)\cap(G_2\cup Z_2)\cap\dots$ where each G_j is a usual open set having full density at every point of Z_j. Suppose that $S\cap\Delta(\bar{S})=S$. Then $G_j\cap\bar{S}$ is dense in $\bar{S}$ because every neighborhood of a point in $\bar{S}$ contains points of positive upper density of $\bar{S}$, and some of these must be in G_j as Z_j is a null set. But then, $\bigcap_{n=1}^{\infty}(G_j\cap\bar{S})$ is a residual subset of $\bar{S}$, and as such is uncountable. This, however, contradicts the fact that $\bigcap_{n=1}^{\infty}(G_j\cap\bar{S})\subset\bigcap_{n=1}^{\infty} G_j\subset S$ which is countable.

REMARK. The conclusion of the lemmas is still valid if the hypothesis that S is countable is replaced by the hypothesis that S is not residual in $\bar{S}$.

THEOREM 2. *If* $S\subset R^n$ *is countable, then* 1, 2 *and* 3 *are equivalent.*

PROOF. It follows from Theorem 1 that we need only prove $2\Rightarrow 1$. As S is $\mathcal{T}-G_\delta$ and S_β differs from S in an at most countable set, it follows that S_β is $\mathcal{T}-G_\delta$ for every β. Further, it follows from Lemma 4 that $S_{\beta+1}\neq S_\beta$ unless $S_\beta=\emptyset$. Consequently, there is an $\alpha<\omega_1$ such that $S_\alpha=\emptyset$, and so S is $\mathcal{T}$-scattered.

Consider now a function $f\colon R^n\to R$ for which $E^0\cup E_0$ is countable. From a result of D. PREISS [P], we know that $f\in\mathcal{B}_1(\mathcal{A})$ since f is in the usual second class of Baire. Obviously, f is the pointwise limit of functions which are zero except at finitely many points. It follows that $f\in\mathscr{B}_1(\mathscr{P})$. Thus we obtain from our theorem, the following

COROLLARY. *Suppose* $f\colon R^n\to R$ *and that* $E^0\cup E_0$ *is countable. Then* $f\in\mathscr{B}_1(\mathscr{A}\cap\mathscr{P})$ *if and only if* $f\in \mathrm{AP}_1$ *(equivalently* AP_3*).*

PROOF. According to Theorem N_3, we only need to show that f and $-f$ belong to AP_2. If $a<0$, $S=\Delta(\overline{W})$ is empty and so the set in question is A_δ. If $b>0$, then the set $R^n\setminus E_b$ is countable and, by Lemma 2, $\mathscr{T}$-scattered. That is, for $b>0$, let $b'=b/2$ and $R^n\setminus E_b\subset E^{b'}$.

We note that as a special case, we have

COROLLARY. *If S is countable, then* $\chi_S\in\mathscr{B}_1(\mathscr{A}\cap\mathscr{P})$ *if and only if* $\chi_S\in\mathrm{AP}_1$ *(equivalently* AP_3*).*

The authors would like to thank the referee for his useful suggestions including the proof of Theorem 1 given here.

REFERENCES

[D/G] DAVIES, R. and GALVIN, F., Characterizations of countable G_δ subsets of complete metric spaces, *Real Anal. Exchange* **2** (1976), 74—75.

[G] GRANDE, Z., Sur les suites de fonctions approximativement continues et continues presque partout, *Colloq. Math.* **38** (1977/78), 259—262. *MR* **58** #28342.

[M] MAULDIN, R. D., σ-ideals and related Baire systems, *Fund. Math.* **71** (1971), 171—177. *MR* **45** #2107.

[N] NISHIURA, T., The topology of almost everywhere continuous, approximately continuous functions, *Acta Math. Acad. Sci. Hungar.* (to appear).

[O'M] O'MALLEY, R. J., Approximately differentiable functions: the r-topology, *Pacific J. Math.* **72** (1977), 207—222. *MR* **56** #5810.

[P] PREISS, D., Limits of approximately continuous functions, *Czechoslovak Math. J.* **21 (96)** (1971), 371—372. *MR* **44** #4154.

(Received November 9, 1979)

DEPARTMENT OF MATHEMATICS
WESTERN ILLINOIS UNIVERSITY
MACOMB, IL 61455
U.S.A.

Present address:

TRW SYSTEMS
P.O. BOX 1310
SAN BERNADINO, CA 92402
U.S.A.

DEPARTMENT OF MATHEMATICS
ST. OLAF COLLEGE
NORTHFIELD, MN 55057
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 325—333

ON THE OSCILLATION OF BÔCHER'S PAIRS II

by

I. BIHARI

1. Let us consider the half-linear differential equation

$$(1)\qquad (py')'+qf(y, py') = 0 \quad \left('=\frac{d}{dx}\right), \quad x\in I=[a, b], \; -\infty\leqq a<b\leqq\infty$$

under the conditions given in [1], namely

$$p, q\in C(I), \quad p>0, \quad x\in I,$$

$$f(\lambda u, \lambda v) = \lambda f(u, v), \quad uf(u, v)\geqq 0, \quad f(0, v) = 0, \quad \forall \lambda, u, v; \quad f\in C(R^2).$$

(Further suppose the uniqueness of the solution for some given initial condition.) In [1] a first kind of Bôcher's pairs were regarded. Now a second kind of them will be investigated, namely

$$(2)\qquad \begin{aligned} U &= \varphi y_1-\psi py_1', \\ V &= \varphi y_2-\psi py_2' \end{aligned}$$

under the assumptions

$$\left.\begin{aligned} &(3)\qquad \Delta = y_1'y_2-y_2'y_1\neq 0, \; \varphi, \psi\in C_1(I), \\ &(4)\qquad \{\varphi, \psi\} = p(\varphi'\psi-\psi'\varphi)+\varphi^2+\psi pqf(\psi, \varphi)\neq 0^1 \end{aligned}\right\} \; x\in I$$

where y_1 and y_2 are two (linearly independent) solutions of (1).

Now some results of BÔCHER [2] valid for linear equations will be extended here with respect to (1) and a system studied later in Section 2.

Concerning the pair U, V the following assertions hold:

1° U and V have no common zeros. Namely, at such a point by (3) $\varphi=\psi=0$ which is excluded by (4).

[1] I.e. $u=\dfrac{\varphi}{\psi}$ must not satisfy in any point the Riccati-like equation

$$pz'+z^2+pqf(1, z) = 0$$

$\left(\text{fulfilled by } z=\dfrac{py'}{y}\right)$ which corresponds to (1).

1980 *Mathematics Subject Classification.* Primary 34C10; Secondary 34C15.
Key words and phrases. Oscillation, half-linear differential equation, Sturmian theory, Bôcher's pairs.

2° U (and V) has no double zeros. At such a point we would have $U=U'=0$, or in a more detailed form

$$(U=0)\Rightarrow \varphi y_1-\psi p y_1'=0\Rightarrow\begin{cases} p y_1'=\dfrac{\varphi}{\psi}y_1, & \text{if}\quad \psi\neq 0\Rightarrow y_1\neq 0\quad \text{(case 1)}\\[2ex] y_1=\dfrac{\psi}{\varphi}p y_1', & \text{if}\quad \varphi\neq 0\Rightarrow y_1'\neq 0\quad \text{(case 2)}\end{cases}$$

$$(5)\qquad (U'=0)\Rightarrow 0=\varphi' y_1+\varphi y_1'-\psi' p y_1'+\psi q f(y_1, p y_1')=\begin{cases}\dfrac{y_1}{\psi p}\{\varphi,\psi\}, & \text{(case 1)}\\[2ex] \dfrac{y_1'}{\varphi}\{\varphi,\psi\}, & \text{(case 2)},\end{cases}$$

but the last expression cannot vanish.

3° The zeros of U (and V) do not accumulate to a finite point, since this would involve that there $U=U'=0$.

4° The zeros — if any — of U and V separate each other. For, let $x_1<x_2$ be two adjacent zeros of U and suppose $V\neq 0$ for $x_1<x<x_2$. By 1° $V(x_1)\neq 0$, $V(x_2)\neq 0$, i.e. V has constant sign in $[x_1, x_2]$. From (2)

$$U y_2-V y_1=-\psi p\Delta,$$

$$U y_2'-V y_1'=-\varphi\Delta$$

and at a zero of U

$$y_1=\frac{\psi p}{V}\Delta,\quad y_1'=\frac{\varphi}{V}\Delta,$$

thus by (5) we have here

$$U'=\frac{\Delta}{V}\{\varphi,\psi\},$$

consequently

$$\operatorname{sgn} U'|_{x_1}=\operatorname{sgn} U'|_{x_2},$$

which is impossible.

5° Let us define $\alpha=\alpha(x)$ by a branch of $\alpha=\operatorname{arctg}\dfrac{U}{V}$. Then

$$\alpha'=\frac{D}{U^2+V^2},\quad D=U'V-V'U.$$

At a point x_i where $U(x_i)=0$

$$\alpha'=\frac{U'}{V}=\frac{\Delta}{V^2}\{\varphi,\psi\},$$

i.e. α' has the same sign at every zero of U and a point x_i' where $V(x_i')=0$

$$\alpha'=\frac{-V'}{U}=\frac{\Delta}{U^2}\{\varphi,\psi\}$$

holds, thus α' is of the same sign at every zero of V, too. These facts involve the following statement:

If $\Delta\{\varphi, \psi\}>0$, then α passes the values $i\pi$ and $(2i+1)\frac{\pi}{2}$ $(i=0, 1, 2, \dots)$ increasingly (and therefore once) and conversely, provided $\Delta\{\varphi, \psi\}<0$. Consequently, if $\alpha(x_0)=0, x_i<x_i'$, then

$$\alpha(x_i) = i\pi, \quad \alpha(x_i') = (2i+1)\frac{\pi}{2},$$

or $$(i = 0, 1, 2, \dots)$$

$$\alpha(x_i) = -i\pi, \ \alpha(x_i') = -(2i+1)\frac{\pi}{2},$$

respectively.

6° *Theorem of the simultaneous oscillation. If* $\alpha(x)$ *is the function defined in* 5°, $\alpha_0=\operatorname{arctg}\frac{y_1}{y_2}\left(-\frac{\pi}{2}<\alpha_0(x_0)<\frac{\pi}{2}\right)$ *and* $2\varphi\eta-\psi p\eta'$ *does not vanish inside of* $I=(x_0, \infty)$, *then*

$$\operatorname{tg}(\alpha-\alpha_0) = \frac{Uy_2-Vy_1}{Uy_1+Vy_2} = -\frac{2\psi p\Delta}{2\varphi\eta-\psi p\eta'}, \quad (\eta = y_1^2+y_2^2)$$

never becomes infinite at a finite point, involving $|\alpha-\alpha_0|<\frac{\pi}{2}(x\geqq x_0)$ *consequently the two pairs* (y_1, y_2) *and* (U, V) *are oscillatory or non-oscillatory at the same time. Obviously* φ *and* ψ *can be chosen in such a manner that* $\alpha_0(x_0)\neq\pm\frac{\pi}{2}$.[2]

2. Let us consider now the half-linear system

$$\text{(6)} \qquad \begin{aligned} x' &= af_1(x, y)+bg_1(x, y), \\ y' &= cf_2(x, y)+dg_2(x, y) \end{aligned} \quad \left('=\frac{d}{dt}\right), \ t\in[A, B] = I$$

under the conditions stipulated in [1] and repeated here as

$$a, b, c, d\in C(I), \ f_i, g_i\in C(R^2), \ f_i(\lambda u, \lambda v) = \lambda f_i(u, v), \ g_i(\lambda u, \lambda v) = \lambda g_i(u, v), \ \forall\lambda, u, v$$

$$\begin{aligned} \operatorname{sgn} f_i = \operatorname{sgn} u, \ \operatorname{sgn} g_i &= \operatorname{sgn} v, \\ f_i(0, v) = g_i(u, 0) &= 0, \end{aligned} \qquad (i = 1, 2).$$

Let (x_i, y_i) $(i=1, 2)$ be two independent solutions of (6), i.e. $x_1y_2-x_2y_1\neq 0$, $t\in I$ and consider the pair

$$\text{(7)} \qquad \begin{aligned} U &= \varphi x_1-\psi y_1, \\ V &= \varphi x_2-\psi y_2. \end{aligned}$$

Here $\varphi, \psi\in C_1(I)$ and

$$\text{(8)} \quad \{\varphi, \psi\} = \varphi'\psi-\psi'\varphi+\varphi[af_1(\psi, \varphi)+bg_1(\psi, \varphi)]-\psi[cf_2(\psi, \varphi)+dg_2(\psi, \varphi)] \neq 0, \quad t\in I.$$

[2] See the analysis of the condition $2\varphi\eta-\psi p\eta'\neq 0$ in Section 4.

Then the above assertions $1°-6°$ hold also in the present case with the same wording but replacing the above definition of $\{\varphi, \psi\}$ by (8) and y_i, y_i' by x_i, y_i $(i=1,2)$, respectively. (See some details in Section 5.)

3. *Supplement to part I* (i.e. [1]).

Let y be a non-trivial solution of (1), the zeros — if any — of the pair

$$\begin{matrix} \Phi = \varphi_1 y - \varphi_2 p y', \\ \Psi = \psi_1 y - \psi_2 p y' \end{matrix} \qquad \Delta = \varphi_1\psi_2 - \varphi_2\psi_1 \neq 0, \ x \in I$$

— under the conditions $\{\varphi_1, \varphi_2\} \neq 0$, $\{\psi_1, \psi_2\} \neq 0$ — separate each other, and — under the assumption $\{\varphi_1, \varphi_2\}\cdot\{\psi_1, \psi_2\} < 0$, $x \in I$ only one of φ and ψ can vanish, moreover once at most.

Letting now $\psi_1 = 1$, $\psi_2 = 0$ we have $\Psi = y$ and $\{\psi_1, \psi_2\} = 1$. If $\varphi_2 \neq 0$, then $\Delta = -\varphi_2 \neq 0$ and assuming $\{\varphi_1, \varphi_2\} < 0$ the second case occurs which is possible just when y has at most one zero in I, i.e. y is disconjugated on I. On the other hand if y has two zeros $x_1 < x_2$, then $\{\varphi_1, \varphi_2\} < 0$ cannot hold for $x_1 \leqq x \leqq x_2$. E.g., if the conditions

$$pq > 0, \quad -\infty \leqq x \leqq \infty$$

$$\int^{\infty} \frac{dx}{p} = \infty, \quad \int^{\infty} q\,dx = \infty,$$

hold, then (1) (and y) is oscillatory and

$$\{\varphi_1, \varphi_2\} = \left[p\left(\frac{\varphi_1}{\varphi_2}\right)' + \left(\frac{\varphi_1}{\varphi_2}\right)^2 + pqf\left(1, \frac{\varphi_1}{\varphi_2}\right)\right]\varphi_2^2 < 0$$

cannot hold on the whole line $-\infty \leqq x \leqq \infty$.

4. Analysis of the condition $\psi p\eta' - 2\varphi\eta \neq 0$, $x \in I$.

We have to show: φ and ψ can be chosen in such a way that in the formula

$$\text{(9)} \qquad \operatorname{tg}(\alpha - \alpha_0) = \frac{2p\Delta}{\eta\left(z - \dfrac{2\varphi}{\psi}\right)}, \quad z = \frac{p\eta'}{\eta}$$

the denominator does not vanish at a *finite* point. Indeed, $\eta = y_1^2 + y_2^2 \neq 0$ and z remains finite *inside* I, thus $\dfrac{2\varphi}{\psi}$ can be chosen as wanted, e.g. to be larger (smaller) than z. This choice of $\dfrac{2\varphi}{\psi}$ is connected to the given pair (y_1, y_2) of solutions of (1). If (1) is *linear*, i.e. $f(u, v) \equiv u$ then this choice is of universal validity, that means the function $\dfrac{2\varphi}{\psi}$ is suitable to *any* pair $(\tilde{y}_1, \tilde{y}_2)$ of linearly independent solutions, too.

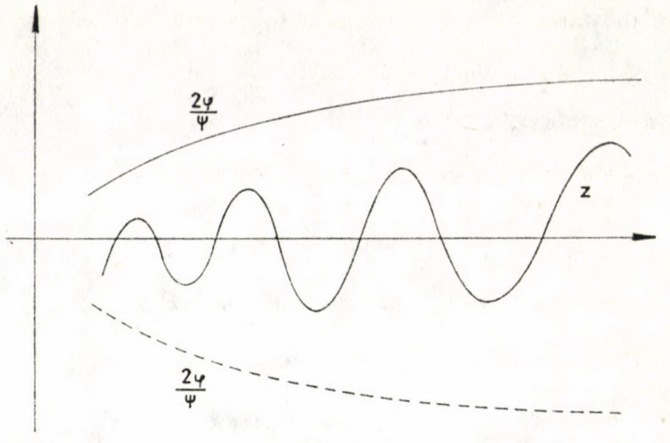


Fig. 1

Indeed, every such pair can be obtained from the pair (y_1, y_2) by a non-degenerated linear transformation which is composed of an orthogonal transformation (rotation),

$$(10)\qquad \begin{aligned} \bar{y}_1 &= y_1 \cos\beta - y_2 \sin\beta \\ \bar{y}_2 &= y_1 \sin\beta + y_2 \cos\beta \end{aligned} \qquad (\beta = \text{const.})$$

which leaves

$$\eta = y_1^2 + y_2^2, \quad \eta' = 2(y_1 y_1' + y_2 y_2'), \quad z = \frac{p\eta'}{\eta}, \quad \Delta = y_1' y_2 - y_2' y_1$$

invariant and a stretching

$$(11)\qquad \tilde{y}_1 = k_1 \bar{y}_1, \; \tilde{y}_2 = k_2 \bar{y}_2, \quad (k_1 \neq 0,\; k_2 \neq 0, \text{ in general } k_1 \neq k_2)$$

where

$$\tilde{\Delta} = \tilde{y}_1' \tilde{y}_2 - \tilde{y}_2' \tilde{y}_1 = k_1 k_2 \bar{\Delta} = k_1 k_2 \Delta \neq 0.$$

Therefore, for (10) $\bar{z} = z$ and the assertion is evident while in the case of transformation (11) we have to consider that the zeros remain invariant (do not move away) consequently if the pairs $(\bar{y}_1, \bar{y}_2)$, $(\bar{U}, \bar{V})$ are oscillatory (non-oscillatory) *simultaneously*, then so are the pairs $(\tilde{y}_1; \tilde{y}_2)$, $(\tilde{U}, \tilde{V})$, too.

However, in the non-oscillatory case we can state more. The quantity z satisfies an equation of Riccati type, namely

$$(12)\qquad z' + \frac{z^2}{p} + 2q = \frac{2pu}{\eta}, \quad u = y_1'^2 + y_2'^2$$

and by the easily verifiable identity

$$4\Delta^2 = 4\eta u - \eta'^2 \quad (p\Delta = c = \text{const.})$$

or

$$\frac{4c^2}{p\eta^2} = \frac{4pu}{\eta} - \frac{z^2}{p},$$

(12) will have the form

$$z'+\frac{z^2}{2p}+2q=\frac{2c^2}{p\eta^2} \tag{13}$$

which involves

$$z<z_0-2\int_{x_0}^{x} q\,dx+2c^2\int_{x_0}^{x}\frac{dx}{p\eta^2}=F(x), \quad z_0=z(x_0). \tag{14}$$

By a theorem of W. E. MILNE (see e.g. [3], p. 354, Th. 6.3) in the *non-oscillatory case* $\int_{x_0}^{\infty}\frac{dx}{p\eta}<\infty$ and one at least of y_1 and y_2 (and also η) does not tend to zero as $x\to\infty$, involving

$$\int_{x_0}^{x}\frac{dx}{p\eta^2}<\infty, \tag{14'}$$

too. If in addition $\int^{\infty} q\,dx>-\infty$ $\left(\text{now} \int^{\infty} q\,dx<\infty \text{ see [7]}\right)$, then by (14) and (14') z is bounded above and the denominator $\eta\left(z-\frac{2\varphi}{\psi}\right)$ — at a suitable choice of $\frac{2\varphi}{\psi}$ — remains above a positive number in absolute value as $x\to\infty$ involving $|\alpha-\alpha_0|<$ $<\frac{\pi}{2}-\varepsilon$ for some $\varepsilon>0$. In the oscillatory case $\int^{\infty}\frac{dx}{p\eta}=\infty$ and just $\int^{\infty} q\,dx=\infty$ is sufficient condition of this behaviour. Thus (14) does not give an upper bound for z. However, the choice $\frac{2\varphi}{\psi}>F(x)$ or simply $\frac{2\varphi}{\psi}>z\,(x\geqq x_0)$ is appropriate in this case to assure $|\alpha-\alpha_0|<\frac{\pi}{2}$ for $x\geqq x_0$.

5. Concerning the system (6) of Section 2 the simultaneous oscillation theorem is the following:

Let t_0 be a zero of U,

$$\alpha=\operatorname{arctg}\frac{U}{V},\ \alpha(t_0)=0,\ \alpha_0=\operatorname{arctg}\frac{y_1}{y_2},\quad \left(\text{or } \operatorname{arctg}\frac{x_1}{x_2}\right),$$

$-\frac{\pi}{2}<\alpha_0(t_0)<\frac{\pi}{2}$, *then φ and ψ can be chosen in such a way that* $|\alpha-\alpha_0|<\frac{\pi}{2}$ *for* $t\geqq t_0$.

The proof proceeds as previously. Namely now

$$\operatorname{tg}(\alpha-\alpha_0)=\frac{Uy_2-Vy_1}{Uy_1+Vy_2}=\frac{\varphi\Delta}{\varphi W-\psi\eta}=\frac{\Delta}{\eta\left(z-\frac{\psi}{\varphi}\right)},\quad z=\frac{W}{\eta}. \tag{15}$$

Here

$$W=x_1y_1+x_2y_2,\quad \eta=y_1^2+y_2^2,\quad \Delta=x_1y_2-x_2y_1.$$

If system (6) is linear

$$(16) \qquad x' = ax+by,\ y' = cx+dy,\ \Delta = k\exp\Big(\int_{x_0}^{x}(a+d)\,dt\Big),\quad k = \text{const},$$

then again an equation can be derived for z, viz. $z' = \dfrac{W'}{\eta} - \dfrac{\eta'}{\eta} z$,

$$W' = (a+d)W + b\eta + c\xi, \qquad \frac{\eta'}{\eta} = 2(cz+d), \qquad (\xi = x_1^2 + x_2^2),$$

involving

$$(17) \qquad z' + 2cz^2 + (d-a)z - b = c\frac{\xi}{\eta}.$$

Furthermore from the identity

$$(x_1y_1+x_2y_2)^2 + (x_1y_2-x_2y_1)^2 = (x_1^2+x_2^2)(y_1^2+y_2^2)$$

we have

$$W^2 + \Delta^2 = \xi\eta \quad \text{or} \quad \frac{\xi}{\eta} = \frac{\Delta^2}{\eta^2} + z^2$$

implying

$$(18) \qquad z' + cz^2 + (d-a)z - b = c\frac{\Delta^2}{\eta^2}.$$

Contrary to the case of equation (1), it cannot be supposed here — not even in the non-oscillatory case — that $x_i > 0$, $y_i > 0$ for $t > t_0$ with some $t_0 > 0$ (which would involve $w > 0$, $z > 0$). If $d = a$, $c > 0$ $(t \geqq t_0)$, then

$$(19) \qquad z(t) < z(t_0) + \int_{t_0}^{t} b\,dt + \int_{t_0}^{t} c\frac{\Delta^2}{\eta^2}\,dt = F(t).$$

In the non-oscillatory case the second integral is convergent if $t \to \infty$ (Milne). Supposing the same behaviour of the first one, z is bounded above, while in the oscillatory case the convergencies of these integrals are not assured. All these facts permit assertions similar to those at the end of Section 4.

6. In order to illucidate the behaviour of the solutions x and y of (16) let us carry out in (16) the transformation

$$(20) \qquad u = xe^{-\int a\,dt}, \quad v = ye^{-\int d\,dt}, \quad f = e^{\int (d-a)\,dt}.$$

We obtain

$$(21) \qquad u' = bfv, \quad v' = \frac{c}{f}u$$

and

$$(22) \qquad \left(\frac{u'}{bf}\right)' - \frac{c}{f}u = 0, \quad \left(\frac{f}{c}v'\right)' - bfv = 0.$$

(21) is a system equivalent to (16) while equations (22) are two independent second order linear equations concerning u and v, respectively, which are satisfied by the solutions of (21) provided $\dfrac{u'}{bf}$ and $\dfrac{f}{c}v'$ are differentiable.

The following statements can be easily proved by means of (21) and (22).

(i) If for $t \geqq t_0$ $bc < 0$, then the zeros — if any — of x and y separate each other.

(ii) If $b > 0$, $\int^{\infty} bf dt = \infty$, $\int^{\infty} \frac{c}{f} dt = -\infty$, then x and y are oscillatory. (Extension of a theorem of WINTNER [7].)

(iii) The simultaneous oscillation theorem can be applied to the pairs (v_1, v_2), (U, V) where $U = \varphi u_1 - \psi v_1$, $V = \varphi u_2 - \psi v_2$. Then equations (18)—(19) have simpler forms (with no linear term), viz.

$$z' + \frac{c}{f} z^2 - bf = \frac{ck}{f\eta^2}, \tag{18'}$$

$$z(t) < z(t_0) + \int_{t_0}^{t} bf dt + k \int_{t_0}^{t} \frac{c}{f\eta^2} dt = F(t), \quad k = \text{const.} \tag{19'}$$

In the non-oscillatory case (19′) gives a bound for z without restriction $d = a$.

(iv) The corresponding of Milne's theorem: x is oscillatory (non-oscillatory) if and only if

$$\left| \int^{\infty} \frac{b\Delta}{\xi} dt \right| = \infty (< \infty), \quad \Delta = e^{\int (a+d) dt}, \quad \xi = x_1^2 + x_2^2.$$

(v) In the non-oscillatory case the extension of a theorem by HARTMAN and WINTNER ([3], p. 355, Th. 6.4) reads as follows:

There is a solution x_1 with the property

$$\left| \int^{\infty} \frac{b\Delta}{x_1^2} dt \right| = \infty$$

and for every solution x_2 which is linearly independent from x_1

$$\left| \int^{\infty} \frac{b\Delta}{x_2^2} dt \right| < \infty.$$

A similar result holds concerning y, too.

(vi) If $\frac{c}{f^2 b}$ decreases and is negative, then the amplitudes of $u = xe^{-\int a dt}$ are decreasing and those of $v = ye^{-\int d dt}$ are increasing. (Extension of a theorem by SONIN and PÓLYA, see [6].)

(vii) If we change a, b, c, d in such a way that $\frac{c}{f}$ or $-bf$ (or both of them) increase, then u and v (i.e. x and y) oscillate more quickly (in a familiar sense).

(viii) Under the conditions of (vi) the areas of the half-waves and quarter-waves of u and v are decreasing and increasing, respectively (see [4] and [5]).

REFERENCES

[1] BIHARI, I., On the oscillations of Bôcher's pairs with respect to half-linear differential equations and systems, *Period. Math. Hungar.* **7** (1976), 153—160. *MR* **58** # 28838.

[2] BÔCHER, M., On certain pairs of transcendental functions whose roots separate each other, *Trans. Amer. Math. Soc.* **26** (1901), 428—436.

[3] HARTMAN, PH., *Ordinary differential equations*, John Wiley & Sons, Inc., New York—London—Sydney, 1964. *MR* **30** # 1270.

[4] MAKAI, E., On a monotonic property of certain Sturm-Liouville functions, *Acta Math. Acad. Sci. Hungar.* **3** (1952), 165—172. *MR* **14**—872.

[5] BIHARI, I., Oscillation and monotonity theorems concerning non-linear differential equations of the second order, *Acta Math. Acad. Sci. Hungar.* **9** (1958), 83—104. *MR* **20** # 1824.

[6] SZEGŐ, G., *Orthogonal polynomials*, American Mathematical Society Colloquium Publications, Vol. 23, Revised ed., American Mathematical Society, Providence, R. I., 1959, (p. 161). *MR* **21** # 5029.

[7] WINTNER, A., A criterion of oscillatory stability, *Quart. Appl. Math.*, **7** (1949), 115—117. *MR* **10**—456.

(Received December 5, 1979)

MTA MATEMATIKAI KUTATÓ INTÉZETE
REÁLTANODA U. 13—15.
H—1053 BUDAPEST
HUNGARY

REFERENCES

[1] [illegible], I., On the oscillations of Bôcher's pairs with respect to half-linear differential equations and systems, *Period. Math. Hungar.* 7 (1976), 153–160. MR 58 # 2863.
[2] Bôcher, M., On certain pairs of transcendental functions whose roots separate each other, *Trans. Amer. Math. Soc.* 26 (1901), 428–436.
[3] Hartman, Ph., *Ordinary differential equations*, John Wiley & Sons, Inc., New York—London—Sydney, 1964. MR 30 # 1270.
[4] [illegible] On a monotonic property of certain Sturm-Liouville functions, *Acta Math. Acad. Sci. Hungar.* 3 (1952), 165–[illegible]; MR 14–872.
[5] [illegible] Oscillation and monotonity theorems concerning non-linear differential equations of the second order, *Acta Math. Acad. Sci. Hungar.* 9 (1958), [illegible]–104. MR 20 # 1824.
[6] [illegible], G., *Orthogonal polynomials*, American Mathematical Society Colloquium Publications, Vol. 23, Revised [illegible], American Mathematical Society, Providence, R. I., 1959. (p. 161). MR 21 # 5029.
[7] [illegible], A., [illegible] of oscillation [illegible], *Quart. Appl. Math.*, 7 (1949), 115–117. MR 10–[illegible].

(Received [illegible], 19[illegible])

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 335—340

LONG PATHS IN RANDOM GRAPHS

by

W. FERNANDEZ DE LA VÉGA

Abstract

A random graph with cn edges, $c>1.39$, contains a path of length $\left(1-\frac{1.39}{c}\right)n$. A random directed graph with cn edges, $c>2.78$, contains a directed path of length $\left(1-\frac{2.78}{c}\right)n$.

§ 1. Introduction

A. We define $G_{n,p}$ as the random graph on n vertices where each edge is present with probability p, independently of the other edges. Similarly $D_{n,p}$ will denote the random directed graph with independent (directed) edges, each one present with probability p, so that the possibility that two vertices be joined in both directions is not excluded.

We further define the random graph $G'_{n,N}$ (resp. the random directed graph $D'_{n,N}$) as the graph obtained by picking at random N elements from the $\frac{n(n-1)}{2}$ possible edges (resp. from the $n(n-1)$ possible directed edges) defined on an n-vertex set. We are going to prove the following theorems, in which "almost surely" means "with a probability tending to 1 as n tends to infinity".

THEOREM 1. *The random undirected graph $G'_{n,\alpha n}$ on n vertices with αn edges, $\alpha>1.39$, almost surely contains a path of length $\left(1-\frac{1.39}{\alpha}\right)n$.*

THEOREM 2. *The random directed graph $D_{n,\beta n}$ on n vertices with βn edges, $\beta>2.78$ almost surely contains a directed path of length $\left(1-\frac{2.78}{\beta}\right)n$.*

B. REMARK. We have learned that M. AJTAI, J. KOMLÓS and E. SZEMERÉDI [1] proved the following theorems:

The random directed graph $D'_{n,\beta n}$ with n vertices and βn directed edges $\beta>1$ almost surely contains a directed path of length cn_3 $c=c(\beta)$.

The random (undirected) graph $G'_{n,\alpha n}$ with n vertices and αn edges, $\alpha>1/2$, almost surely contains a path of length cn, $c=c(\alpha)$.

AMS (MOS) subject classifications (1980). Primary O5C38.
Key words and phrases. Random graph, constant average degree, long path, Markov chain.

Moreover, these authors obtain in each case an exponential rate of convergence to 1 for the probabilities that these graphs contain such paths for suitable $c(\alpha)$ and $c(\beta)$. For large β these authors obtain a path of length $\sim\left(1-\frac{\log\beta}{\beta}\right)n$ so that our Theorem 2 improves slightly their results in this range. On the other hand our Theorem 2 says nothing in the range $1<\beta\leqq 2.78$. Similar remarks apply to the undirected case.

C. Theorems 1 and 2 will be deduced from the following analogous statements concerning $G_{n,p}$ and $D_{n,p}$:

Long paths in the undirected independent case (L.P.U.I.C. for short):

The random undirected graph $G_{n,p}$, $p=\frac{\alpha}{n}$, $\alpha<2.78$, almost surely contains a path of length $\left(1-\frac{2.78}{\alpha}\right)n$.

Long paths in the directed independent case (L.P.D.I.C. for short):

The random directed graph $D_{n,p}$, $p=\frac{\alpha}{n}$, $\beta>2.78$, almost surely contains a path of length $\left(1-\frac{2.78}{\beta}\right)n$.

The proof of the L.P.U.I.C. and L.P.D.I.C. statements is done in §1. Let us now show that the L.P.U.I.C. statement implies Theorem 1. We are going to show that, if $G_{n,p}$ with $p=\frac{c}{n}$, almost surely contains a path of length $\left(1-\frac{b}{c}\right)n$ then so does $G'_{n,N}$, with $N=\frac{cn}{2}$. Clearly, this will imply Theorem 1 if we take $b=2.78$ as given by the L.P.U.I.C. statement. To proceed with the proof let $\mathfrak{X}=\mathfrak{X}(n,N,b)$ denote the subset of the graphs with n vertices and N edges, $N=\frac{cn}{2}$, which do not contain a path of length $\left(1-\frac{b}{c}\right)n$. Let π denote a random permutation of the edges of the complete graph on n vertices and let M denote a binomial r.v. with parameters $\frac{n(n-1)}{2}$ and p, independent of π. Then the N first elements of π are jointly distributed as $G'_{n,N}$, $N=\frac{cn}{2}$, while the M first elements are distributed as $G_{n,p}$. If $H\in\mathfrak{X}$, then very nearly half the times where π begins with H, M will satisfy $M\leqq\frac{cn}{2}$, so that the corresponding element of $G_{n,p}$ will be included in H and will not contain a path of length $\left(1-\frac{b}{c}\right)n$. Hence the $G_{n,p}$ measure of the graphs that do not contain such a path exceeds $\frac{1}{2}-o(1)$ times their $G'_{n,N}$ measure so that the former can tend to zero only if the latter does and that is just what we wished to prove. The same argument applies to show that the L.P.D.I.C. statement implies Theorem 2.

§ 2. The cases with independent edges

In this section we prove the L.P.U.I.C. and L.P.D.I.C. assertions. We give the proof for the undirected case and we shall indicate the rather obvious modification that yields the directed case.

For given p we define r as the positive root of the equation

$$(1) \qquad p = 2r - r^2$$

and we shall consider the random multigraph $H_{n,r}$ defined as the edge union of two independent random graphs with same vertex set, with white and black edges respectively, and both distributed as $G_{n,r}$. It is clear that if we drop the colours of the edges of $H_{n,r}$ and count the double edges only once we obtain a graph distributed as $G_{n,p}$. Hence we just have to prove that $H_{n,r}$, for r equal to the root of (1) corresponding to $p=\frac{\alpha}{n}$, contains almost surely a path of length $\left(1-\frac{2.78}{\alpha}\right)n$. We are going to define an algorithm which, given a multigraph on n vertices with white and black edges, constructs a sequence of paths of this graph and we will show that, when applied to a graph distributed as $H_{n,r}$, this algorithm gives almost surely a path of the required length.

Each of the constructed paths will be considered in the obvious way as a sequence of vertices, some of which can be marked (which vertex is the first and which is the last will be obvious from the construction). We will denote the last vertex of the current path by e. We will consider single vertices as paths. Empty paths do also occur, so that e is not always defined. At each step of the construction we refer to the set of vertices which have not been used before this step as the set of "free vertices" and denote it by V. We denote by Q the current path, by v the cardinality of V.

The algorithm starts with the path $Q=Q_0$ consisting of the first vertex as unique and unmarked vertex. It consists in the iteration of a "BASIC STEP" and ends when the set of free vertices becomes empty. We do not mention in the description of the basic step the updating of the set V which is done of course by immediately suppressing from it each vertex which is added to Q.

BASIC STEP. The basic step takes exactly one of three possible courses of action.

Case 1. *Q is empty.* Then keep it so with probability $q=(1-r)^v$ or (with probability $1-q$) make $Q=$ the first free vertex.

Case 2. *e is not marked.* Then, if there is some free vertex joined to e by a white edge, add the first such vertex to Q. If not, mark e.

Case 3. *e is marked.* Then, if there is some free vertex joined to e by a black edge add the first such vertex to Q. If not, then

— if all the vertices of Q are marked make $Q=$the empty path;

— if there are unmarked vertices in Q mark the rightmost of them and suppress from Q all vertices following it.

This ends the description of the algorithm for the undirected case. For the directed case the only modification needed consists in specifying that the added edges be conveniently directed. Note that the basic step adds at most one vertex to Q (and therefore suppresses at most one vertex from V) at each call. Note also that, by the last stipulation in the description of the basic step, each vertex can be used at most once as unmarked extremity and at most once as marked extremity of current paths.

We shall denote by

Q_k the state of Q just after the k first executions of the basic step,
e_k the extremity of Q_k,
$l(k)$ the length of Q_k,
V_k the corresponding state of the set of free vertices,
v_k the cardinality of V_k.

The following inequality is crucial for our proof:

$$l(k) \geqq 2(n-v_k)-k-2. \tag{2}$$

For the proof of (2) observe that v_{i+1} is equal to v_i-1 exactly when the basic step adds one vertex to Q and then no vertex is marked. When $v_{i+1}=v_i$ the basic step marks at most one vertex. Since only marked vertices are eventually suppressed we can write

$$l(k) \geqq \sum_{i=0}^{k-1} 1_{\{v_{i+1}=v_i-1\}} - \sum_{i=0}^{k-1} 1_{\{v_{i+1}=v_i\}},$$

$$l(k) \geqq n-1-v_k-[k-(n-1-v_k)],$$

which is exactly (2).

If we apply this algorithm to a random graph distributed as $H_{n,r}$ then the v_k's become well-defined random variables. We claim that then the sequence $(v_k)_{k=0,1,\ldots}$ is a Markov chain with state space $\{0,1, \ldots, n-1\}$, initial distribution concentrated at $n-1$ and with the stationary transition probabilities

$$\mathsf{P}[v_{k+1}=j\,|\,v_k=j] = (1-r)^j,$$

and

$$\mathsf{P}[v_{k+1}=j-1\,|\,v_k=j] = 1-(1-r)^j \qquad 0 \leqq j \leqq n-1.$$

Let $i_0, i_1, \ldots, i_k$ be integers satisfying the conditions $i_0=n-1$ and $i_h-1 \leqq i_{h+1} \leqq i_h$, $h=0, 1, \ldots, k-1$. We have to prove that for each such set of integers we have

$$\mathsf{P}[v_{k+1} = i_k\,|\,v_0 = i_0, \ldots, v_k = i_k] = (1-r)^{i_k} \tag{3}$$

which will of course entail

$$\mathsf{P}[v_{k+1} = i_k-1\,|\,v_0 = i_0, \ldots, v_k = i_k] = 1-(1-r)^{i_k}. \tag{4}$$

Let us define the events

$$A_1 = \bigwedge_{j=0}^{k} \{v_j = i_j\} \wedge \{e_k \text{ is not marked}\},$$

and

$$A_2 = \bigwedge_{j=0}^{k} \{v_j = i_j\} \wedge \{e_k \text{ is marked}\}.$$

The non-empty events of the form $A_i \wedge \{e_k=x\} \wedge \{V_k=S\}$, $i=1, 2$, where x ranges over the vertex set and S ranges over the family of subsets of the vertex set of cardinality i_k, form, together with the event $Q_k=\emptyset$, a partition of the conditioning event in (3) and (4). Therefore, it will be enough to verify that

(5) $$\mathsf{P}[v_{k+1} = i_k | A_i,\ e_k = x, V_k = S] = (1-r)^{i_k}$$

holds for conditioning events of positive probability. But, for $i=1$ (resp. $i=2$) the algorithm has never yet looked at the white (resp. black) edges joining x to the set S. Hence the conditional distribution, under the conditioning event in (5), of these white edges (resp. these black edges) coincides with their original distribution, i.e. these edges are in each case conditionally independent and each one is present with probability r. Therefore (5) holds and this ends the proof of (3), (4) and of the Markovian property of the sequence (v_k).

We denote by s_i the time spent by the chain (v_k) in the state i. By known results concerning Bernoulli sequences we have

$$\mathsf{E}s_i = \frac{1}{1-(1-r)^i}$$

which implies

(6) $$\mathsf{E}s_i \leqq \frac{1}{1-\exp\{-ri\}}$$

and

$$\mathsf{Var}\, s_i = \frac{(1-r)^i}{[1-(1-r)^i]^2}$$

which implies

(7) $$\mathsf{Var}\, s_i \leqq \frac{\exp\{-ri\}}{[1-\exp\{-ri\}]^2}.$$

The hitting time of the state j is given by

$$T_j = \sum_{i=j+1}^{n-1} s_i.$$

By the strong Markov property the s_i's are independent and so

$$\mathsf{Var}\, T_j = \sum_{i=j+1}^{n-1} \mathsf{Var}\, s_i.$$

Using (6) and (7) and bounding the sums by integrals it is easily checked that, for $rj \leqq \log 2$, the following inequalities hold

(8) $$\mathsf{E}T_j \leqq n$$

and

(9) $$\mathsf{Var}\, T_j \leqq \frac{2}{r} \leqq \frac{4n}{\alpha}.$$

Define now j_0 by

$$\log 2 < rj_0 \leqq \log 2 + r$$

so that j_0 satisfies

(10) $$j_0 \leqq \frac{\log 2}{r} + 1 \leqq \frac{2n \log 2}{\alpha} + 1.$$

Tchebycheff inequality implies, using (8) and (9)

$$\text{(11)} \qquad \mathsf{P}\left[T_{j_0} \leqq n+2\lambda\sqrt{\frac{n}{\alpha}}\right] \geqq 1-\frac{1}{\lambda^2}.$$

Now, replacing in (2) v_k by the bound of j_0 given by (10) and k by the bound of T_{j_0} given by (11), we obtain

$$\mathsf{P}\left[l(T_{j_0}) \geqq n\left(1-\frac{4\log 2}{\alpha}\right)-2\lambda\sqrt{\frac{n}{\alpha}}-2\right] \geqq 1-\frac{1}{\lambda^2}.$$

Since 2.78 is greater than $4\log 2$, the last inequality implies

$$\mathsf{P}\left[l(T_{j_0}) \geqq n\left(1-\frac{2.78}{\alpha}\right)\right] \to 1 \quad \text{as} \quad n \to \infty,$$

completing the proof.

REFERENCES

[1] Ajtai, Miklós, Komlós, János and Szemerédi, Endre, The longest path in a random graph, *Combinatorica* **1** (1981), 1—12.

(Received December 14, 1979)

UNIVERSITÉ DE PARIS—SUD
L.R.I., BÂT. 490
F—91405 ORSAY Cedex

Mailing address:

4 bis, RUE WULFRAN PUGET
F—13008 MARSEILLE
FRANCE

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 341—356

SIMULTANEOUS APPROXIMATION OVER A FIELD OF FINITE TRANSCENDENCE TYPE

by

K. SARADHA

§ 1. Introduction

In this paper we deal with the values of the exponential function over a field of transcendence type $\leqq \tau$ (see Definition 3.1 below). In 1966 LANG proved that certain values of the exponential function cannot all be algebraic over a field of transcendence type $\leqq \tau$. In the sequel we take the following common hypothesis.

COMMON HYPOTHESIS 1 (CH 1). Let the sets of complex numbers $\{u_1, \ldots, u_m\}$ and $\{v_1, \ldots, v_n\}$ each be linearly independent over Q, the field of rational numbers. Lang [2] proved the following theorem.

THEOREM 1.1. *Under CH 1 if K is a field of transcendence type $\leqq \tau$ $2 \leqq \tau$ and $\tau(m+n) \leqq mn$, then at least one of the numbers*

$$\exp(u_i v_j) \quad 1 \leqq i \leqq m; \quad 1 \leqq j \leqq n \tag{1}$$

is not algebraic over K.

In his thesis, W. D. BROWNAWELL [1] proved the following two theorems.

THEOREM 1.2. *Under CH 1, if $\tau(m+n) < mn+m+n$ and K is a field of transcendence type $\leqq \tau$, then at least one of the numbers*

$$u_i,\ v_j,\ \exp(u_i v_j), \quad 1 \leqq i \leqq m; \quad 1 \leqq j \leqq n \tag{2}$$

is not algebraic over K.

THEOREM 1.3. *Under CH 1 if $\tau(m+n) \leqq mn+m$ and K is a field of transcendence type $\leqq \tau$, then at least one of the numbers*

$$u_i,\ \exp(u_i v_j), \quad 1 \leqq i \leqq m; \quad 1 \leqq j \leqq n \tag{3}$$

is not algebraic over K.

Theorems (1.2) and (1.3) generalize the classical Hermite—Lindemann and Gelfond—Schneider theorems; since Q is a field of transcendence type $\leqq 1$. These theorems (1.2) and (1.3) were rediscovered by WALDSCHMIDT [6] in a more general setup.

AMS (MOS) subject classifications (1970). Primary 10F10; Secondary 10F05, 10F25, 10F30, 10F35.

Key words and phrases. Transcendental numbers, exponential function, simultaneous approximation, transcendence type, field of finite transcendence type.

§ 2. Main results

In this paper we additionally make two more common hypotheses besides CH 1.

CH 2. K is a purely transcendental extension over Q of finite type (see the remark under Definition 3.2).

L is a finite extension of degree D over K.

CH 3.

$$\min_{\substack{l_i \in \mathbf{Z} \\ |l_i| \leqq q \\ l_i \text{ not all zero}}} |l_1 u_1 + \ldots + l_m u_m| > \exp\{-F(q)\}$$

and

$$\min_{\substack{l_i \in \mathbf{Z} \\ |l_i| \leqq q \\ l_i \text{ not all zero}}} |l_1 v_1 + \ldots + l_n v_n| > \exp\{-F(q)\}$$

where $F(x)$ is an increasing real valued function with $\operatorname{lt}_{x \to \infty} F(x) = \infty$.

Under the above three common hypotheses, we now establish quantitative results concerning the three theorems stated in § 1, that is, we get nontrivial lower bounds when the numbers in (1), (2) and (3) are approximated simultaneously by numbers from L.

To every $\alpha \in L$, we associate a real number $t(\alpha)$, called its size (see Definition 3.3 below).

We now state the main results of this paper.

THEOREM 2.1. *Under the common hypotheses CH 1, CH 2 and CH 3, let* α_{ij} $(1 \leqq i < m;\ 1 \leqq j \leqq n)$ *be* mn *elements of* L *with* $t(\alpha_{ij}) \leqq S$. *Also assume that* $mn > \tau(m+n)$, $\tau \geqq 2$. *Then*

$$\sum_{i,j} |\exp(u_i v_j) - \alpha_{ij}| >$$

$$> \exp\left\{-D_1 \left(\frac{S^{mn}}{(\log S)^{m+n}}\right)^{\frac{\tau}{mn-\tau(m+n)}} \left[\log S + F\left(D_2\left(\frac{S^{\tau}}{\log S}\right)^{\frac{\max(m,n)}{mn-\tau(m+n)}}\right)\right]\right\}. \tag{4}$$

where D_1 *and* D_2 *are constants independent of* S *and are easily computable.*

THEOREM 2.2. *Under the common hypotheses CH 1, CH 2 and CH 3, let* $\alpha_i, \beta_j, \alpha_{ij}$ $(1 \leqq i \leqq m;\ 1 \leqq j \leqq n)$ *be* $mn+m+n$ *elements of* L *with* $t(\alpha_i)$, $t(\beta_j)$ *and* $t(\alpha_{ij}) \leqq S$. *Assume that* $mn+m+n > \tau(m+n)$. *Then*

$$\sum_i |u_i - \alpha_i| + \sum_j |v_j - \beta_j| + \sum_{i,j} |\exp(u_i v_j) - \alpha_{ij}| >$$

$$> \exp\left\{-D_3 S^{\frac{\tau(mn+m+n)}{mn-(\tau-1)(m+n)}} \left[1 + \frac{F\left(D_4 S^{\frac{\max(m,n)\tau}{mn-(\tau-1)(m+n)}}\right)}{\log S}\right]\right\}, \tag{5}$$

where D_3 *and* D_4 *are absolute constants independent of* S *and are easily computable.*

THEOREM 2.3. *Under the common hypotheses CH 1, CH 2 and CH 3, let* α_i, α_{ij} $(1 \leqq i \leqq m;\ 1 \leqq j \leqq n)$ *be* $mn+m$ *elements in* L *with* $t(\alpha_i)$ *and* $t(\alpha_{ij}) \leqq S$. *Assume that*

$mn+m>\tau(m+n)$. Then

$$\sum_{i=1}^{m}|u_i-\alpha_i|+\sum_{i,j}|\exp(u_iv_j)-\alpha_{ij}|> \tag{6}$$
$$>\exp\{-D_5S^{\frac{\tau m(n+1)}{mn-(\tau-1)m-\tau n}}[\log S+F\{D_6\max(S^{\frac{\tau(n+1)}{mn-(\tau-1)m-\tau n}},(\log S)^{\frac{1}{n+1}}S^{\frac{\tau(m-1)}{mn-(\tau-1)m-\tau n}})\}]\}$$

where D_5 and D_6 are absolute constants independent of S and are easily computable.

§ 3. Definitions and notations

For sake of convenient reference we collect in this section some definitions and notations. Let us consider a polynomial ($\not\equiv 0$) in q variables

$$P(x_1, \ldots, x_q)\in\mathbf{Z}[x_1, \ldots, x_q].$$

Let $\deg_{x_i}P=r_i$ and $\deg P=\sum_{i=1}^{q}r_i$. $H(P)$ denotes the height of P, i.e., the maximum of the absolute values of the coefficients of P. We define the size of P denoted by $t(P)$ as

$$t(P)=\max\{\deg P, \log H(P)\}.$$

Let K be a subfield of the field of complex numbers and let τ be a real number $\geqq 1$.

DEFINITION 3.1. K is said to have a transcendence type $\leqq\tau$ over Q if K has a finite transcendence degree over Q and if $(x_1, \ldots, x_q)$ is a transcendence basis of K over Q, then for all $\alpha\in\mathbf{Z}[x_1, \ldots, x_q]$, $\alpha\neq 0$, one has

$$-(t(\alpha))^{\tau}\ll\log|\alpha|.$$

(The constant involved in the sign $\ll$ does not depend on α.)

DEFINITION 3.2. The set of elements $\{x_1, \ldots, x_q, y\}$ is said to form a system of generators for K if the following three conditions are satisfied.

1) $K=Q(x_1, \ldots, x_q, y)$;
2) $x_1, \ldots, x_q$ are algebraically independent over Q;
3) y is integral over $\mathbf{Z}[x_1, \ldots, x_q]$.

REMARK. K is purely transcendental if $K=Q(x_1, \ldots, x_q)$ with (2) above. Any element a of K can be written uniquely as

$$a=\sum_{i=1}^{\delta}\frac{Q_i}{R_i}y^{i-1}$$

where $\delta=[K\colon Q(x_1, \ldots, x_q)]$ is the degree of y over $\mathbf{Z}[x_1, \ldots, x_q]$, and Q_i and R_i are elements of $\mathbf{Z}[x_1, \ldots, x_q]$ without common factors. Let P be the least common multiple of $R_1, \ldots, R_\delta$. This is called the denominator of a. Then

$$Pa=\sum_{i=1}^{\delta}P_iy^{i-1}$$

where $P_i\in\mathbf{Z}[x_1, \ldots, x_q]$, $1\leqq i\leqq\delta$.

DEFINITION 3.3. Size of a, $t(a)$, with respect to the system of generators $(x_1, \ldots, x_q, y)$ is defined as

$$t(a) = \max\bigl(t(P), t(P_1), \ldots, t(P_\delta)\bigr).$$

REMARK 1. When K is a purely transcendental field i.e., $K=Q(x_1, \ldots, x_q)$ then any finite extension L of K has a system of generators $(x_1, \ldots, x_q, y)$. This follows easily from the theorem on primitive elements.

REMARK 2. When $\alpha\in K$ has size $t(\alpha)$ then $|\alpha|\leqq C^{t(\alpha)}$ where C is an absolute constant.

REMARK 3. If $K=Q(x_1, \ldots, x_q)$ has transcendence type $\leqq\tau$ with respect to $(x_1, \ldots, x_q)$ then any finite extension L of K has transcendence type $\leqq\tau$ with respect to $(x_1, \ldots, x_q, y)$ (see Lang [2]).

We make use of the above three remarks as we go on without any mention thereof.

§ 4. Some lemmas

(A) LEMMA 4.1 (see Waldschmidt [7]). *Let K be a finite extension of Q of finite type. Let $(x_1, \ldots, x_q, y)$ be a system of generators of K over Q. If $\alpha_1, \ldots, \alpha_m\in \mathbf{Z}[x_1, \ldots, x_q, y]$ then*

$$t(\alpha_1+\ldots+\alpha_m) \leqq \max_{1\leqq i\leqq m} t(\alpha_i)+\log m. \tag{7}$$

There exists a constant $c>0$, depending only on $(x_1, \ldots, x_q, y)$ such that for all $(a_1, \ldots, a_m)\in K^m$, one has

$$t(a_1+\ldots+a_m) \leqq c\bigl(t(a_1)+\ldots+t(a_m)\bigr), \tag{8}$$

$$t(a_1\ldots a_m) \leqq c\bigl(t(a_1)+\ldots+t(a_m)\bigr). \tag{9}$$

If $\alpha\in\mathbf{Z}[x_1, \ldots, x_q, y]$ and $\beta\in\mathbf{Z}[x_1, \ldots, x_q]$, $\beta\neq 0$, one has

$$t\left(\frac{\alpha}{\beta}\right) \leqq c\max\bigl(t(\alpha), t(\beta)\bigr). \tag{10}$$

LEMMA 4.2 (see Waldschmidt [7]). *Let $K_1\subset K_2$ be two fields in $\mathbf{C}$ of finite type with $\{x_1, \ldots, x_q, y\}$ and $\{\xi_1, \ldots, \xi_\nu, z\}$ as systems of generators, respectively. Let t_1 and t_2 be the respective sizes. Then there exists a constant c depending only on these generators such that*

$$\frac{1}{c}\, t_1(\alpha) \leqq t_2(\alpha) \leqq c t_1(\alpha)$$

for all nonzero $\alpha\in K_1$.

This lemma shows that the size that we have defined in § 3 does not depend on the system of generators.

LEMMA 4.3. *Let $K=Q(x_1, \ldots, x_q)$, $I_K=\mathbf{Z}[x_1, \ldots, x_q]$, $L=Q(x_1, \ldots, x_q, y)$ $I_L=\mathbf{Z}[x_1, \ldots, x_q, y]$ and $D=[L:K]$. Consider*

$$\sum_{j=1}^{m} a_{ij} z_j = 0, \qquad 1\leqq i\leqq n. \tag{11}$$

(i) *If* $a_{ij}\in I_K$, $t(a_{ij})\leqq S$ *and* $m>2^{q+1}n$, *then there exists a nontrivial solution for* $z_j\in I_K$ *such that*

$$t(z_j) \leqq C_1(S+\log m), \qquad 1 \leqq j \leqq m.$$

(ii) *If* $a_{ij}\in I_L$, $t(a_{ij})\leqq S$ *and* $m>2^{q+1}Dn$, *then there exists a nontrivial solution for* $z_j\in I_K$ *such that*

$$t(z_j) \leqq C_2(S+\log m), \quad 1 \leqq j \leqq m.$$

Here C_1 *and* C_2 *are absolute constants.*

This is the analogue of standard Siegel's lemma that deals with solutions of a set of linear equations.

PROOF. (i) Write

$$\left.\begin{aligned} a_{ij} &= \sum_{(\lambda)} a_{ij(\lambda)} x_1^{\lambda_1} \dots x_q^{\lambda_q} \\ z_j &= \sum_{(\mu)} z_{j(\mu)} x_1^{\mu_1} \dots x_q^{\mu_q} \end{aligned}\right. \qquad a_{ij(\lambda)},\, z_{j(\mu)} \in \mathbf{Z},$$

where deg z_j with respect to each x_i is $\leqq S$. Substituting this in (11) we find that (11) is equivalent to a system of equations in which there are $m(S+1)^q$ unknowns in $n(2S+1)^q$ equations with rational integer coefficients. Each coefficient in this new system of linear equations has absolute valu e atmost $e^S(1+S)^q$. Hence by the standard Siegel's lemma over Q

$$|z_{j(\mu)}| \leqq \left(me^S(1+S)^{2q}\right)^{\frac{n(2S+1)^q}{m(S+1)^q-n(2S+1)^q}} \leqq \left(me^S(1+S)^{2q}\right)^{\frac{2^q n}{m-2^q n}}.$$

Hence the result in (i) follows.

(ii) The system of equations in (11) is equivalent to

$$\sum_{j=1}^{m}\left(\sum_{k=1}^{D} a_{ijk}\, y^{k-1}\right) z_j = 0 \qquad 1 \leqq i \leqq n,\ a_{ijk}\in I_K.$$

Since $1, y, \dots, y^{D-1}$ are linearly independent over K, this is equivalent to $\sum_{j=1}^{m} a_{ijk}z_j=0$, $1\leqq i\leqq n$, $1\leqq k\leqq D$, $a_{ijk}\in I_K$. Now the result follows from (i).

(B) We now state two important lemmas due to TIJDEMAN [5] which we utilize in establishing the quantitative results mentioned in § 2.

LEMMA 4.4. *Let* m, s *and* t *be positive integers and set* $r=st$. *Let* $\alpha_0, \alpha_1, \dots, \alpha_{m-1}$ *and* $\beta_0, \beta_1, \dots, \beta_{s-1}$ *be* m *and* s *distinct complex numbers, respectively, and let*

$$a = \max_{0\leqq \nu<m} (|\alpha_\nu|, 1) \qquad b = \max_{0\leqq \sigma<s} (|\beta_\sigma|, 1)$$

$$a_0 = \min_{\substack{0\leqq \mu,\ \nu<m \\ \mu\neq\nu}} (|\alpha_\mu-\alpha_\nu|, 1) \quad b_0 = \min_{\substack{0\leqq \varrho,\ \sigma<s \\ \varrho\neq\sigma}} (|\beta_\varrho-\beta_\sigma|, 1).$$

For arbitrary complex numbers A_ν, *let*

$$E(z) = \sum_{\nu=0}^{m-1} A_\nu e^{\alpha_\nu z},$$

$$A = \max_{0\leqq \nu<m} |A_\nu| \quad and \quad E = \max_{\substack{0\leqq \varrho<t \\ 0\leqq \sigma<s}} |E^{(\varrho)}(\beta_\sigma)|.$$

Assume $r \geqq 2m+13ab$. *Then*

$$A \leqq s\sqrt{m!}\,e^{7ab}\left(\frac{1}{2a_0 b}\right)^{m-1}\left(\frac{72b}{b_0\sqrt{s}}\right)^{r} E. \tag{12}$$

This is Theorem 3 in Tijdeman [5].

LEMMA 4.5. *Let n, s, t and u be positive integers and set $m=nu$ and $r=st$. Let $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ and $\beta_0, \ldots, \beta_{s-1}$ be n and s distinct complex numbers, respectively,*

$$a = \max_{0 \leqq \nu < n}\{|\alpha_\nu|, 1\} \qquad b = \max_{0 \leqq \sigma < s}\{|\beta_\sigma|, 1\}$$

$$a_0 = \min_{\substack{0 \leqq \mu,\, \nu < n \\ \mu \neq \nu}}\{|\alpha_\mu - \alpha_\nu|, 1\} \quad b_0 = \min_{\substack{0 \leqq \varrho,\, \sigma < s \\ \varrho \neq \sigma}}\{|\beta_\varrho - \beta_\sigma|, 1\}.$$

Denote by $A_{\mu\nu}$ $(0 \leqq \mu < u;\ 0 \leqq \nu < n)$ any set of mn complex numbers. Further put

$$E(z) = \sum_{\nu=0}^{n-1}\sum_{\mu=0}^{u-1} A_{\mu\nu} z^{\mu} e^{\alpha_\nu z},$$

$$A = \max_{\substack{0 \leqq u < u \\ 0 \leqq \nu < n}} |A_{\mu\nu}| \quad \text{and} \quad E = \max_{\substack{0 \leqq \varrho < t \\ 0 \leqq \sigma < s}} |E^{(\varrho)}(\beta_\sigma)|.$$

Assume finally that $r \geqq 2m+13ab$. *Then*

$$A \leqq s\left(\frac{6a}{a_0\sqrt{n}}\max\left(6, \frac{m}{b}\right)\right)^{m}\left(\frac{72b}{b_0\sqrt{s}}\right)^{r} E. \tag{13}$$

This is Theorem 2 in Tijdeman [5].

§ 5. Proofs

This section gives the proofs of the theorems stated in § 2. The main pattern of the proof is classical. See [3] and [4]. For the first theorem we construct an auxiliary function which is small at certain lattice points while for the second theorem the auxiliary function, together with some of its derivatives is small at certain lattice points. The third theorem can be approached eitherway. See § 5 (c).

Let us now on adopt the following simple notations. $\underline{\lambda}=(\lambda_1, \ldots, \lambda_m)$; $\mu=(\mu_1, \ldots, \mu_n)$ where each λ_i, μ_j take only non-negative integral values. Also $\underline{\lambda u}=\lambda_1 u_1+\ldots+\lambda_m u_m$, $\underline{\mu v}=\mu_1 v_1+\ldots+\mu_n v_n$. The constants $c, c_1, c_2, \ldots$ in Theorem 2.1, $d, d_1, d_2, \ldots$ in Theorem 2.2 and $e, e_1, e_2, \ldots$ in Theorem 2.3 are numbers depending only on K, L, u_i and v_j and independent of S. These are easily computable.

(A) PROOF of Theorem 2.1. Let $\exp(u_i v_j)-\alpha_{ij}=\varepsilon_{ij}$ and $|\varepsilon|=\max(|\varepsilon_{ij}|)$. Without loss of generality assume that $|\varepsilon|<1$. Thus a lower bound for $|\varepsilon|$ will yield the necessary result (4). Let

$$\varphi(z) = \sum_{\underline{\lambda}} p(\underline{\lambda}) \exp(\underline{\lambda u})\, z \tag{14}$$

where $0 \leqq \lambda_i \leqq X$, $1 \leqq i \leqq m$ and $p(\underline{\lambda}) \in I_K$ are unknowns to be determined as follows.

Set

$$\psi_{\underline{\mu}} = \sum_{\underline{\lambda}} p(\underline{\lambda}) \prod_{i=1}^{m} \prod_{j=1}^{n} \alpha_{ij}^{\lambda_i \mu_j} = 0 \quad \text{for} \quad 1 \leqq \mu_j \leqq Y,\ 1 \leqq j \leqq n. \tag{15}$$

This gives $(X+1)^m$ unknowns in Y^n equations with coefficients in L. To make them elements of I_L, multiply by a polynomial $P \in I_K$ with $t(P) \leqq C_1 XYS$. Still the size of the new coefficients is $\leqq c_2 XYS$. Choose

$$X = [2^{(q+1)/m} D^{1/m} Y^{n/m}]. \tag{16}$$

Then by Lemma 4.3, (ii) there exists non-trivial solution $p(\underline{\lambda}) \in I_K$ such that

$$t(p(\underline{\lambda})) \leqq C_3(m \log (X+1) + XYS) \leqq C_4 XYS. \tag{17}$$

Take

$$Y^* = [2^{(m+2)/n} X^{m/n}] + 1. \tag{18}$$

Note that $Y^* > Y$.

Claim. There exists $\underline{\mu}^*$ with $\max \underline{\mu}^* = Y^*$ such that $\psi_{\underline{\mu}^*} \neq 0$ for $|\varepsilon|$ sufficiently small. Suppose $\psi_{\underline{\mu}} = 0$ for all $\underline{\mu}$ with $\max \underline{\mu} \leqq Y^*$. Consider

$$\begin{aligned}
|\varphi(\underline{\mu}\underline{v})| &= |\varphi(\underline{\mu}\underline{v}) - \psi_{\underline{\mu}}| = \\
&= \Big|\sum_{\underline{\lambda}} p(\underline{\lambda}) \Big\{\prod_{i,j} (\alpha_{ij} + \varepsilon_{ij})^{\lambda_i \mu_j} - \prod_{i,j} \alpha_{ij}^{\lambda_i \mu_j}\Big\}\Big| = \\
&= \left|\sum_{\underline{\lambda}} p(\underline{\lambda}) \prod_{i,j} \sum_{\substack{k_{ij}=0 \\ k_{ij} \text{ not all zero}}}^{\lambda_i \mu_j} \binom{\lambda_i \mu_j}{k_{ij}} \alpha_{ij}^{\lambda_i \mu_j - k_{ij}} \varepsilon_{ij}^{k_{ij}}\right| \leqq \qquad (19) \\
&\leqq |\varepsilon| (X+1)^m c^{C_4 XYS} c_1^{C_5 XY^* S} 2^{mnXY^*} \leqq \\
&\leqq |\varepsilon| \exp (C_6 XY^* S).
\end{aligned}$$

Now use Lemma 4.4 with

$$\text{‘}\{\alpha_v\}\text{’} = \{\underline{\lambda}\underline{u}\}_{\substack{0 \leqq \lambda_i \leqq X \\ 1 \leqq i \leqq m}}$$

hence ‘m’$=(X+1)^m$ and

$$\text{‘}\{\beta_\sigma\}\text{’} = \{\underline{\mu}\underline{v}\}_{\substack{1 \leqq \mu_j \leqq Y^* \\ 1 \leqq j \leqq n}}$$

hence ‘s’$=Y^{*n}$. Let ‘t’$=1$, hence ‘r’$=Y^{*n}$. Also ‘a’$\leqq C_7 X$; ‘b’$\leqq C_8 Y^*$. By CH 3, ‘a_0’$\geqq \exp \{-C_9 F(X)\}$ and ,b_0’$\geqq \exp \{-C_{10} F(Y^*)\}$. Take ‘$E(z)$’$=\varphi(z)$, ‘$E$’$\leqq$ $\leqq |\varepsilon| \exp(C_6 XY^* S)$. The condition ,$r \geqq 2m + 13ab$’ can be easily verified for large X. Hence

$$\begin{aligned}
&|p(\underline{\lambda})| \leqq \\
&\leqq Y^{*n} \sqrt{(X+1)^m!}\, e^{7C_7 C_8 XY^*} \left(\frac{1}{2C_8 Y^* \exp(-C_9 F(X))}\right)^{(X+1)^m} \left(\frac{72 C_{10} Y^*}{Y^{*n/2} \exp(-C_{10} F(Y^*))}\right)^{Y^{*n}} \times \\
&\times |\varepsilon| \exp (C_6 XY^* S) \leqq \\
&\leqq |\varepsilon| \exp \left[C_{11} \{X^m (F(X) + \log X) + Y^{*n} (F(Y^*) + \log Y^*) + XY^* S\}\right].
\end{aligned}$$

Note that

$$|p(\underline{\lambda})| \geqq \exp\{-C_{12}(XYS)^{\tau}\}$$

for $p(\underline{\lambda})\neq 0$, and by construction there exists at least one $p(\underline{\lambda})\neq 0$.

If

(20) $|\varepsilon| < \exp\left[-C_{11}\{X^m(F(X)+\log X)+Y^{*n}(F(Y^*)+\log Y^*)+XY^*S\}-C_{12}(XYS)^{\tau}\right]$
then $|p(\underline{\lambda})| < \exp\{-C_{12}(XYS)^{\tau}\}$ for all $\underline{\lambda}$ which is a contradiction. This proves the claim.

Let Y_1 be the least such Y^*. Hence there exists $\underline{\mu}^*$ with $\max\underline{\mu}^*=Y_1$ such that $\psi_{\underline{\mu}^*}\neq 0$ and for all $\underline{\mu}$ with $\max\underline{\mu}<Y_1$, $\psi_{\underline{\mu}}=0$. Note that $Y<Y_1\leqq Y^*$.

Firstly, we estimate $|\varphi(\underline{\mu}^*\underline{\nu})|$ from below. Consider

$$P\psi_{\underline{\mu}^*} = \sum_{\underline{\lambda}} p(\underline{\lambda}) P \prod_{i,j} \alpha_{ij}^{\lambda_i \mu_j^*}$$

where $P\in I_K$ and $P\psi_{\underline{\mu}^*}\in I_L$. We know that $t(P)\leqq C_1XY_1S$. Using (7) and (10) of Lemma 4.1

$$t(\psi_{\underline{\mu}^*}) \leqq C_{13}XY_1S.$$

Hence

$$|\psi_{\underline{\mu}^*}| \geqq \exp\{-C_{14}(XY_1S)^{\tau}\}.$$

From (19) it follows that

$$|\varphi(\underline{\mu}^*\underline{\nu})| \geqq \exp\{-C_{14}(XY_1S)^{\tau}\}-|\varepsilon|\exp(C_{15}XY_1S). \tag{21}$$

Secondly, we estimate $|\varphi(\underline{\mu}^*\underline{\nu})|$ from above. For that consider the integral

$$\frac{1}{2\pi i}\int_C \frac{\varphi(z)}{z-\underline{\mu}^*\nu}\,\frac{1}{\prod\limits_{\substack{\underline{\mu}\\ \max\underline{\mu}<Y_1}}(z-\underline{\mu}\underline{\nu})}\,dz$$

where $C\colon |z|=C_{16}Y_1S$ where C_{16} is taken suitably so that the lattice points $\underline{\mu}\underline{\nu}$ with $\max\underline{\mu}<Y_1$ all lie inside C. Thus we get

$$\varphi(\underline{\mu}^*\underline{\nu}) = \frac{1}{2\pi i}\int_C \frac{\varphi(z)}{z-\underline{\mu}^*\underline{\nu}} \prod_{\substack{\underline{\mu}\\ \max\underline{\mu}<Y_1}} \left(\frac{\underline{\mu}^*\underline{\nu}-\underline{\mu}\underline{\nu}}{z-\underline{\mu}\underline{\nu}}\right) dz -$$

$$-\sum_{\substack{\underline{\mu}'\\ \max\underline{\mu}'<Y_1}} \frac{\varphi(\underline{\mu}'\underline{\nu})}{(\underline{\mu}'\underline{\nu}-\underline{\mu}^*\underline{\nu})}\,\frac{\prod\limits_{\substack{\underline{\mu}\\ \max\underline{\mu}<Y_1}}(\underline{\mu}^*\underline{\nu}-\underline{\mu}\underline{\nu})}{\prod\limits_{\substack{\underline{\mu}\\ \max\underline{\mu}<Y_1\\ \underline{\mu}'\neq\underline{\mu}}}(\underline{\mu}'\underline{\nu}-\underline{\mu}\underline{\nu})}.$$

It is easy to see that

$$\left|\int\right| \leqq \exp(C_{17}XY_1S)\,S^{-(Y_1-1)^n}.$$

To estimate Σ, use (19) and the CH 3. Then we get

$$|\Sigma| \leqq |\varepsilon| \exp\{C_{18} Y_1^n(F(Y_1)+\log Y_1)+XY_1S\}.$$

Thus

$$(22)\quad |\varphi(\underline{\mu}^*\underline{\nu})| \leqq \exp(C_{17}XY_1S)S^{-Y_1^n/2^n}+|\varepsilon|\exp\{C_{18}[Y_1^n(F(Y_1)+\log Y_1)+XY_1S]\}.$$

Now the lower bound (1.b) and the upper bound (u.b) found are inconsistent if

$$\exp\{-C_{14}(XY_1S)^\tau\}-|\varepsilon|\exp\{C_{15}XY_1S\}>$$

$$>\exp\{C_{17}XY_1S\}S^{-Y_1^n/2^n}+|\varepsilon|\exp\{C_{18}[Y_1^n(F(Y_1)+\log Y_1)+XY_1S]\},$$

i.e., if

$$1>\exp\{C_{19}(XY_1S)^\tau\}S^{-Y_1^n/2^n}+|\varepsilon|\exp\{C_{20}Y_1^n(F(Y_1)+\log Y_1)+C_{19}(XY_1S)^\tau\},$$

i.e., if

$$S^{Y_1^n/2^n}>\exp\{C_{19}(XY_1S)^\tau\}\{1+|\varepsilon|\exp[C_{20}Y_1^n(F(Y_1)+\log Y_1)]S^{Y_1^n/2^n}\}.$$

So now if

$$(23)\quad |\varepsilon| \leqq \exp(-C_{20}Y_1^n[F(Y_1)+\log Y_1])S^{-Y_1^n/2^n}$$

then the above inequality is valid if

$$S^{Y_1^n/2^n}>2\exp C_{19}\{(XY_1S)^\tau\},$$

i.e., if

$$Y_1^n\log S>C_{21}(XY_1S)^\tau,$$

i.e., if

$$Y_1^{n-\tau}\log S>C_{22}Y_1^{n\tau/m}S^\tau$$

(substituting for X from (16) and observing that $Y_1>Y$). Thus 1.b. and u.b. are inconsistent if

$$Y_1^{mn-\tau(m+n)/m}>C_{22}\frac{S^\tau}{\log S}.$$

Now let $mn-\tau(m+n)=E_1$ (>0 by assumption). Choose

$$(24)\quad Y=\left[C_{22}\left(\frac{S^\tau}{\log S}\right)^{m/E_1}\right].$$

Then the above conditon is satisfied. Thus the 1.b. and u.b. are inconsistent for this choice of Y when (23) holds. So we get from (20) and (23) that either

(25)

$$|\varepsilon| \geqq \exp\{-C_{11}[X^m(F(X)+\log X)+Y^{*n}(F(Y^*)+\log Y^*)+XY^*S]-C_{12}(XYS)^\tau\}$$

or

$$(26)\quad |\varepsilon|>\exp\{-C_{23}[Y_1^n(F(Y_1)+\log Y_1+\log S)]\}.$$

Note that

$$Y_1 \leqq Y^* < C_{24}\left(\frac{S^\tau}{\log S}\right)^{m/E_1};\quad X \leqq C_{25}\left(\frac{S^\tau}{\log S}\right)^{n/E_1}.$$

On substituting these values in (26) we find

$$|\varepsilon| > \exp\left[-C_{26}\left(\frac{S^\tau}{\log S}\right)^{mn/E_1}\left\{F\left(C_{24}\left(\frac{S^\tau}{\log S}\right)^{m/E_1}\right)+\log\left(\frac{S^\tau}{\log S}\right)\right\}\right]. \tag{27}$$

(25) becomes

$$\begin{aligned}|\varepsilon| > \exp\Bigg[-C_{27}\left(\frac{S^{mn}}{(\log S)^{m+n}}\right)^{\tau/E_1}\Bigg\{F\left(C_{25}\left(\frac{S^\tau}{\log S}\right)^{n/E_1}\right)+\\ +F\left(C_{24}\left(\frac{S^\tau}{\log S}\right)^{m/E_1}\right)+\log\left(\frac{S^\tau}{\log S}\right)\Bigg\}\Bigg].\end{aligned} \tag{28}$$

Thus the required l.b. for $|\varepsilon|$ is of the form

$$\exp\left[-C_{28}\left(\frac{S^{mn}}{(\log S)^{m+n}}\right)^{\tau/E_1}\left\{F\left(C_{29}\left(\frac{S^\tau}{\log S}\right)^{(\max(m,n))/E_1}\right)+\log S\right\}\right].$$

This proves the theorem.

(B) PROOF of Theorem 2.2. Let $u_i-\alpha_i=\varepsilon_j'$, $v_j-\beta_j=\varepsilon_j'$ and $\exp(u_iv_j)-\alpha_{ij}=$ $=\varepsilon_{ij}$ for $1\leqq i\leqq m$, $1\leqq j\leqq n$. Let $|\varepsilon|=\max(|\varepsilon_i|, |\varepsilon_j'|, |\varepsilon_{ij}|)$. Take the auxiliary function as

$$\varphi(z)=\sum_{\lambda_0=0}^{p(x)-1}\sum_{\underline{\lambda}} p(\lambda_0,\underline{\lambda})z^{\lambda_0}\exp(\underline{\lambda}\underline{u})z \tag{29}$$

for $0\leqq\lambda_i\leqq X$, $1\leqq i\leqq m$ and

$$p(X)=\left[\frac{X^{1+m/n}}{\log X}\right].$$

Then

$$\varphi^{(k)}(z)=\sum_{\lambda_0=0}^{p(x)-1}\sum_{\underline{\lambda}} p(\lambda_0,\underline{\lambda})\sum_{r=0}^{k}\binom{k}{r}\frac{\lambda_0!}{(\lambda_0-r)!}z^{\lambda_0-r}(\underline{\lambda}\underline{u})^{k-r}\exp(\underline{\lambda}\underline{u})z \tag{30}$$

for $0\leqq k\leqq p(X)$ where $p(\lambda_0;\underline{\lambda})\in I_K$ are unknowns to be determined as follows. Set

$$\psi_{\underline{\mu},k}=\sum_{\lambda_0=0}^{p(x)-1}\sum_{\underline{\lambda}} p(\lambda_0,\underline{\lambda})\sum_{r=0}^{k}\binom{k}{r}\frac{\lambda_0!}{(\lambda_0-r)!}(\underline{\lambda}\underline{\alpha})^{k-r}(\underline{\mu}\underline{\beta})^{\lambda_0-r}\prod_{i,j}\alpha_{ij}^{\lambda_i\mu_j}=0 \tag{31}$$

for $1\leqq\mu_i\leqq Y$, $1\leqq i\leqq n$; $0\leqq k<p(X)$. (Here $\underline{\lambda}\underline{\alpha}=\lambda_1\alpha_1+\ldots+\lambda_m\alpha_m$ and $\underline{\mu}\underline{\beta}=\mu_1\beta_1+$ $+\ldots+\mu_n\beta_n$.) This gives $p(X)Y^n$ equations in $(X+1)^m p(X)$ unknowns with coefficients in L. A denominator P for the coefficients has size $\leqq d_1(X^{1+(m/n)}+XY)S$. Multiplying by P the above equations, the coefficients go to I_L with size $\leqq d_2(X^{1+(m/n)}+$ $+XY)S$. Choose

$$X=[2^{(q+1)/m}D^{1/m}Y^{n/m}]. \tag{32}$$

Hence by (ii) of Lemma 4.3, there exists $p(\lambda_0,\underline{\lambda})\in I_K$, not all zero with

$$t(p(\lambda_0,\underline{\lambda}))\leqq d_3(X^{1+(m/n)}+XY)S. \tag{33}$$

Set

$$Y^*=[2^{(m+2)/n}X^{m/n}]+1. \tag{34}$$

CLAIM. There exists $\underline{\mu}^*$ with $\max \underline{\mu}^* = Y^*$ such that $\psi_{\underline{\mu}^*, k} \neq 0$ for some k with $0 \leqq k < p(X)$.

Suppose $\psi_{\underline{\mu}^*, k} = 0$ for all $\underline{\mu}^*$ with $\max \underline{\mu}^* = Y^*$ and for all k, $0 \leqq k < p(X)$. Then

$$|\varphi^{(k)}(\underline{\mu}^* \underline{\nu})| = |\varphi^{(k)}(\underline{\mu}^* \underline{\nu}) - \psi_{\underline{\mu}^*, k}| \leqq$$

$$\leqq (X+1)^m p(X) 2^{3p(X)+mnXY^*} p(X)^{p(X)} |\varepsilon| \, d^{d_4(X^{1+(m/n)}+XY^*)S} \leqq \tag{35}$$

$$\leqq |\varepsilon| \exp d_5(X^{1+(m/n)} + XY^*) S$$

(on using the expressions in (30) and (31)).

Now apply Lemma 4.5. Using the notation in that lemma, set

$$\text{`}E(z)\text{'} = \varphi(z);$$

$$\text{`}\{\alpha_\nu\}\text{'} = \{\underline{\lambda u}\}_{0 \leqq \lambda_i \leqq X},$$

$$\text{`}n\text{'} = (X+1)^m;$$

$$\text{`}\{\beta_\sigma\}\text{'} = \{\underline{\mu}\, \underline{\nu}\}_{1 \leqq \mu_j \leqq Y^*},$$

$$\text{`}S\text{'} = Y^{*n}, \quad \text{`}u\text{'} = p(X) = \text{`}t\text{'}, \quad \text{`}m\text{'} = (X+1)^m p(X)$$

$$\text{`}r\text{'} = Y^{*n} p(X), \quad \text{`}a\text{'} \leqq d_6 X, \quad \text{`}b\text{'} \leqq d_7 Y^*,$$

$$\text{`}a_0\text{'} \geqq \exp\{-d_8 F(X)\}, \quad \text{`}b_0\text{'} \geqq \exp\{-d_9 F(Y^*)\},$$

$$\text{`}E\text{'} \leqq |\varepsilon| \exp\{d_5 X^{1+(m/n)} + XY^*) S\}.$$

The condition $r \geqq 2m + 13ab'$ can be easily checked for large X. Then

$$|p(\lambda_0, \underline{\lambda})| \leqq |\varepsilon| \exp\left\{d_{10}\left[X^{1+m+(m/n)}\left(1 + \frac{F(X)}{\log X}\right) + \right.\right.$$

$$\left.\left. + \frac{Y^{*n} X^{1+(m/n)}}{\log X}(F(Y^*) + \log Y^*) + (X^{1+(m/n)} + XY^*) S\right]\right\}.$$

Note that

$$|p(\lambda_0, \underline{\lambda})| \geqq \exp\{-d_{11}(X^{1+(m/n)} + XY)^\tau S^\tau\}$$

for $p(\lambda_0, \underline{\lambda}) \neq 0$ and by construction there always exists a non-zero $p(\lambda_0, \underline{\lambda})$.

If

$$|\varepsilon| < \exp\left[-d_{10}\left\{X^{1+m+(m/n)}\left(1 + \frac{F(X)}{\log X}\right) + \right.\right. \tag{36}$$

$$\left.\left. + \frac{Y^{*n} X^{1+(m/n)}}{\log X}(F(Y^*) + \log Y^*) + (X^{1+(m/n)} + XY^*) S\right\} - d_{11}(X^{1+(m/n)} + XY^*)^\tau S^\tau\right],$$

then

$$|p(\lambda_0, \underline{\lambda})| < \exp\{-d_{11}(X^{1+(m/n)} + XY)^\tau S^\tau\}.$$

This is a contradiction. Hence the claim.

Let Y_1 be the least such Y^*. Thus $\psi_{\underline{\mu},k}=0$ for $\max \underline{\mu}<Y_1$ and for all $0\leqq k<$ $<p(X)$ and there exists $(\underline{\mu}^*, k^*)$ such that $\psi_{\underline{\mu}^*,k^*}\neq 0$ when $\max \underline{\mu}^*=Y_1$ for some $0\leqq k^*<p(X)$. Let us take such pair $(\underline{\mu}^*, k^*)$ with least k^*. Then using familiar argument we find that

$$t(\psi_{\underline{\mu}^*,k^*}) \leqq d_{12}(X^{1+(m/n)}+XY_1)S.$$

Hence

$$|\psi_{\underline{\mu}^*,k^*}| \geqq \exp\{-d_{13}(X^{1+(m/n)}+XY_1)^\tau S^\tau\}.$$

Therefore using (35)

(37) $$|\varphi^{(k^*)}(\underline{\mu}^*\underline{v})| \geqq \exp\{-d_{13}(X^{1+(m/n)}+XY_1)^\tau S^\tau\}-|\varepsilon|\exp\{d_{14}(X^{1+(m/n)}+XY_1)\}.$$

For estimating $\varphi^{(k^*)}(\underline{\mu}^*\underline{v})$ from above, consider the integral

$$\frac{1}{2\pi i}\int_C \frac{\varphi(z)}{(z-\underline{\mu}^*\underline{v})^{k^*+1}} \prod_{\substack{\underline{\mu} \\ \max\underline{\mu}<Y_1}} \frac{1}{(z-\underline{\mu}\underline{v})^{p(x)}}\,dz$$

where $C\colon |z|=d_{15}Y_1^{1+(1/m)}$ containing all lattice points $\underline{\mu}\underline{v}$ with $\max\underline{\mu}<Y_1$. Hence

$$\varphi^{(k^*)}(\underline{\mu}^*\underline{v}) = \frac{k^*!}{2\pi i}\int_C \frac{\varphi(z)\,h(z)}{(z-\underline{\mu}^*\underline{v})^{k^*+1}}\,dz-$$

(38) $$-\sum_{r=0}^{k^*-1}\frac{k^*!}{r!}\varphi^{(r)}(\underline{\mu}^*\underline{v})\frac{1}{2\pi i}\int_{|z-\underline{\mu}^*\underline{v}|=\frac{1}{2}\exp(-F(Y_1))}\frac{h(z)}{(z-\underline{\mu}^*\underline{v})^{k^*-r+1}}\,dz-$$

$$-\sum_{\substack{\underline{\mu}' \\ \max\underline{\mu}'<Y_1}}\sum_{r=0}^{p(X)-1}\frac{k^*!}{r!}\varphi^{(r)}(\underline{\mu}'\underline{v})\frac{1}{2\pi i}\int_{|z-\underline{\mu}'\underline{v}|=\frac{1}{2}\exp(-F(Y_1))}\frac{(z-\underline{\mu}'\underline{v})^r h(z)}{(z-\underline{\mu}^*\underline{v})^{k^*+1}}\,dz$$

where

$$h(z) = \prod_{\substack{\underline{\mu} \\ \max\underline{\mu}<Y_1}}\left(\frac{\underline{\mu}^*\underline{v}-\underline{\mu}\underline{v}}{z-\underline{\mu}\underline{v}}\right)^{p(X)}.$$

Thus

$$|\varphi^{(k^*)}(\underline{\mu}^*\underline{v})| \leqq \Big|\int_C\Big|+|\Sigma|.$$

It is easy to see that

$$\Big|\int_C\Big| \leqq \exp\{d_{16}[(X^{1+(m/n)}+XY)S+XY_1^{1+(1/m)}]\}Y_1^{-(Y_1-1)^n p(X)}$$

and

$$|\Sigma| \leqq |\varepsilon|\exp\{d_{17}p(X)Y_1^n[F(Y_1)+\log Y_1]+d_{18}(X^{1+(m/n)}+XY_1)S\}.$$

Comparing l.b. (37) and u.b. (38), we find that they are inconsistent if

(39)

$$Y_1^{(Y_1^n/2^n)\,p(X)} > \exp\{d_{19}(X^{1+(m/n)}+XY_1)^\tau S^\tau\}\{1+|\varepsilon|\exp[d_{20}p(X)Y_1^n(F(Y_1)+\log Y_1)]\}.$$

So if

$$|\varepsilon| < \exp\{-d_{20}p(X)Y_1^n(F(Y_1)+\log Y_1)\}$$

then the above inequality holds if

$$Y_1^{(Y_1^n/2^n)p(X)} > 2\exp\{d_{19}(X^{1+(m/n)}+XY_1)^\tau S^\tau\},$$

i.e., if

$$Y_1^n \log Y_1 > d_{21} X^{(1+m/n)(\tau-1)} S^\tau \log X,$$

i.e., if

$$Y_1^n > d_{21} S^\tau Y_1^{((m+n)/m)(\tau-1)}$$

on using (32).

Let now $E_2 = mn-(\tau-1)(m+n)$ (>0 by assumption). Choose

$$Y = [d_{21} S^{m\tau/E_2}]. \tag{40}$$

This makes the above inequality hold. Thus for this choice of Y and from (36) and (39) we get either

$$|\varepsilon| \geqq \exp\left\{-d_{10}\left[X^{1+m+(m/n)}\left(1+\frac{F(X)}{\log X}\right)+Y^{*n}\frac{X^{1+(m/n)}}{\log X}(F(Y^*)+\log Y^*)+\right.\right. \tag{41}$$

$$\left.+(X^{1+(m/n)}+XY^*)S]-d_{11}(X^{1+(m/n)}+XY)^\tau S^\tau\right\}$$

or

$$|\varepsilon| \geqq \exp\left\{-d_{20}\frac{X^{1+(m/n)}}{\log X} Y_1^n (F(Y_1)+\log Y_1)\right\}. \tag{42}$$

Now $Y_1 \leqq Y^* \leqq d_{22} S^{m\tau/E_2}$ and $X \leqq d_{23} S^{n\tau/E_2}$ and

$$p(X) \leqq d_{24}\frac{S^{(m+n)\tau/E_2}}{\log S}.$$

Substituting these values in (42), we get

$$|\varepsilon| \geqq \exp\left\{-d_{25} S^{\tau(mn+m+n)/E_2}\left(1+\frac{F(d_{22}S^{m\tau/E_2})}{\log S}\right)\right\}.$$

(41) becomes

$$|\varepsilon| \geqq \exp\left\{-d_{26} S^{\tau(mn+m+n)/E_2}\left(1+\frac{F(d_{22}S^{m\tau/E_2})}{\log S}+\frac{F(d_{23}S^{n\tau/E_2})}{\log S}\right)\right\}.$$

Hence

$$|\varepsilon| > \exp\left\{-d_{27} S^{\tau(mn+m+n)/E_2}\left(1+\frac{F(d_{28}S^{\max(m,n)\tau/E_2})}{\log S}\right)\right\}.$$

This proves the result in (5).

(C) PROOF of Theorem 3. We can prove this theorem either with the method that was used in proving Theorem 2.1 or with that used in Theorem 2.2. While the method of Theorem 2.1 yields a bound only when $mn > \tau(m-1)+\tau n+\tau$, the method of Theorem 2.2 gives a bound for all $mn > \tau(m-1)+\tau n$. Since the ideas are exactly the same as before we shall merely indicate the main steps using the second method, only.

Take the auxiliary function as

$$\varphi(z) = \sum_{\underline{\lambda}} p(\underline{\lambda}) \exp(\underline{\lambda}\underline{u}) z \tag{43}$$

for $0 \leqq \lambda_i \leqq X, 1 \leqq i \leqq m$. Then

$$\varphi^{(k)}(z) = \sum_{\underline{\lambda}} p(\underline{\lambda})(\underline{\lambda}\underline{u})^k \exp(\underline{\lambda}\underline{u}) z$$

for $0 \leqq k < p(X)$ where

$$p(X) = \left[\frac{X^{(m+n)/n+1}}{(\log X)^{n/n+1}}\right]. \tag{44}$$

Let now

$$\psi_{\underline{\mu},k} = \sum_{\underline{\lambda}} p(\underline{\lambda})(\underline{\lambda}\underline{\alpha})^k \prod_{i,j} \alpha_{ij}^{\lambda_i \mu_j} = 0 \tag{45}$$

for $1 \leqq \mu_j \leqq Y, 1 \leqq j \leqq n$; $0 \leqq k < p(X)$. This yields $(X+1)^m$ unknowns in $Y^n p(X)$ equations. Set

$$Y = [X^{(m-1)/n+1}(\log X)^{1/n+1} 2^{-(q+1)/n} D^{-1/n}]. \tag{46}$$

Then using Siegel's Lemma 4.3, there exists $p(\underline{\lambda}) \in I_K$, not all zero, such that

$$t(p(\underline{\lambda})) \leqq e_1 X^{(m+n)/n+1}(\log X)^{1/n+1} S. \tag{47}$$

Let

$$Y^* = [2^{(m+2)/n} X^{(m-1)/n+1}(\log X)^{1/n+1} 2^{-(q+1)/n} D^{-1/n}] + 1. \tag{48}$$

Now use Lemma 4.4 with suitable parameters to conclude that there exists at least one $\underline{\mu}^*$ with $\max \underline{\mu}^* = Y^*$ such that for some k, with $0 \leqq k < p(X)$, $\psi_{\underline{\mu}^*,k} \neq 0$. This happens if

$$\begin{aligned} |\varepsilon| < \exp\big[&-e_3\{X^m(F(X)+\log X) + Y^{*n} p(X)(F(Y^*)+\log X) + \\ &+(X^{(m+n)/n+1}(\log X)^{1/n+1} + XY^*)S\} - e_2(X^{(m+n)/n+1}(\log X)^{1/n+1} S)^\tau\big]. \end{aligned} \tag{49}$$

In the course of proving the above we also find that

$$|\varphi^{(k)}(\underline{\mu}\underline{\nu}) - \psi_{\underline{\mu},k}| \leqq |\varepsilon| \exp\big[e_4\{X^{(m+n)/n+1}(\log X)^{1/n+1} + XY^*\}S\big] \tag{50}$$

whenever $\max \underline{\mu} \leqq Y^*$.

Let Y_1 be the least such Y^*. Choose the pair $(\underline{\mu}^*, k^*)$ with $\max \underline{\mu}^* = Y_1$ and some $0 \leqq k^* < p(X)$ for which $\psi_{\underline{\mu}^*,k^*} \neq 0$. By the usual argument, we can get a lower bound for $\varphi^{(k^*)}(\underline{\mu}^* \underline{\nu})$ as

$$\begin{aligned} |\varphi^{(k^*)}(\underline{\mu}^* \underline{\nu})| \geqq \exp\{&-e_5(X^{(m+n)/n+1}(\log X)^{1/n+1} + XY_1)^\tau S^\tau\} - \\ &-|\varepsilon| \exp\{e_6(X^{(m+n)/n+1}(\log X)^{1/n+1} + XY_1)S\}. \end{aligned} \tag{51}$$

To get an upper bound integrate the function

$$\frac{\varphi(z)}{(z-\underline{\mu}^*\underline{\nu})^{k^*+1}} \prod_{\substack{\underline{\mu} \\ \max \underline{\mu} < Y_1}} \frac{1}{(z-\underline{\mu}\underline{\nu})^{p(X)}}$$

on the circle $|z|=e_7 X^{m-1/2/n+1}$. Then

$$(52)\quad \begin{aligned} |\varphi^{(k^*)}(\underline{\mu}^*\underline{\nu})| \leqq \exp\{e_8 X^{(m+n+1/2)/n+1}(\log X)^{1/n+1} S\} X^{-(Y_1-1)^n p(X)}+ \\ +|\varepsilon| \exp e_9\{Y_1^n p(X)(F(Y_1)+\log Y_1)+(X^{(m+n)/n+1}(\log X)^{1/n+1}+XY_1)S\}. \end{aligned}$$

Comparing (51) and (52) for the inconsistency of the l.b. and u.b. we get

$$|\varepsilon| < \exp\{-e_{10} Y_1^n p(X)(F(Y_1)+\log Y_1)\} X^{-Y_1^n/2^n} p(X)$$

and

$$Y_1^n p(X) \log X > e_{11}(X^{(m+n)/n+1}(\log X)^{1/n+1})^\tau S^\tau.$$

The choice of $p(X)$ and $n>(\tau-1)$ and (48) yield that the last condition is equivalent to

$$X^{(mn-(\tau-1)m-\tau n)/n+1} > e_{12} S^\tau.$$

Let $E_3 = mn-(\tau-1)m-\tau n$ (>0 by assumption). Choose

$$X = [e_{12} S^{\tau(n+1)/E_3}].$$

Note that

$$Y_1 \leqq Y^* \leqq e_{13} S^{(m-1)\tau/E_3}(\log S)^{1/n+1}.$$

Hence for this choice of X, (49) and (53), respectively, become

$$|\varepsilon| \geqq \exp\{-e_{14} S^{\tau m(n+1)/E_3}[\log S+F(e_{15} S^{\tau(n+1)/E_3})+F(e_{16} S^{\tau(m-1)/E_3}(\log S)^{1/n+1})]\}$$

or

$$|\varepsilon| > \exp\{-e_{17} S^{\tau m(n+1)/E_3}[\log S+F(e_{18} S^{\tau(m-1)/E_3}(\log S)^{1/n+1})]\}.$$

Thus in either case we get

$$|\varepsilon| > \exp\{-e_{19} S^{\tau m(n+1)/E_3}[\log S+F\{e_{20}\max(S^{\tau(n+1)/E_3}, S^{\tau(m-1)/E_3}(\log S)^{1/n+1})\}]\}.$$

This holds for $mn > (\tau-1)m+\tau n$. This completes the proof of Theorem 2.3.

ACKNOWLEDGEMENT. I thank Prof. T. S. BHANU MURTHY for his constant encouragement and guidance.

REFERENCES

[1] Brownawell, W. D., Some transcendence results for the exponential function, *Norske Vid. Selsk. Skr. (Trondheim)* **1972**, *no.* 11, 2 pp. *MR* **46** #3454.
[2] Lang, S., *Introduction to transcendental numbers,* Addison—Wesley Publishing Co., Reading, Mass.—London—Don Mills, Ont., 1966. *MR* **35** #5397.
[3] Mignotte, M. and Waldschmidt, M., Approximation simultanée de valeurs de la fonction exponentielle, *Compositio Math.* **34** (1977), 127—139. *MR* **56** #275.
[4] Shorey, T. N., On the sum $\sum_{k=1}^{3} |2^{\pi^k} - \alpha_k|$, α_k algebraic numbers, *J. Number Theory* **6** (1974), 248—260. *MR* **50** #249.
[5] Tijdeman, R., An auxiliary result in the theory of transcendental numbers, *J. Number Theory* **5** (1973), 80—94. *MR* **47** #4942.
[6] Waldschmidt, M., Amélioration d'un théorème de Lang sur l'indépendence algébrique d'exponentielles, *C. R. Acad Sci. Paris Sér. A—B* **272** (1971), A 413—A 415. *MR.* **43** #6163.
[7] Waldschmidt, M., *Nombres transcendants*, Lecture Notes in Mathematics, Vol. 402, Springer-Verlag, Berlin—New York, 1974. *MR* **50** #12931.

(Received February 21, 1980)

RAMANUJAN INSTITUTE
FOR ADVANCED STUDY IN MATHEMATICS
UNIVERSITY OF MADRAS
IND—600 005 MADRAS
INDIA

Present address:

MRS. SARADHA NATARAJAN
FLAT NO. 23, PLOT NO. 276
SION (EAST)
IND—400 022 BOMBAY
INDIA

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 357—362

ON SOME CONVOLUTION TRANSFORMS

by

B. P. DUGGAL

1. Introduction

Let v, V, f and h be some non-trivial elements of $L^2(0, \infty)$. Then

$$(V \circ f)(u) = \int_0^\infty V(ux^{-1}) f(x)\, dx/x$$

and

$$(v \circ h)(u) = \int_0^\infty v(ux^{-1}) h(x)\, dx/x$$

are well defined as elements of $L^2(0, \infty)$. Now let $g \in L^2(0, \infty)$, and let k_1, k_2 be some Lebesgue measurable functions on $(0, \infty)$. We study conditions on k_1 and k_2 in order that the pair of equations

$$\begin{aligned}(V \circ f)(u) &= \int_0^\infty k_1(ux)\, g(x)\, dx, \\ (v \circ h)(u) &= \int_0^\infty k_2(ux) f(x)\, dx\end{aligned} \tag{1.1}$$

hold. We show in this note that a necessary and sufficient condition for (1.1) to be satisfied is that both $k_1^{\hat{}}$ and $k_2^{\hat{}} \in L^2(-\infty, \infty) \cap L^\infty(-\infty, \infty)$ and that $v^{\hat{}} V^{\hat{}\tilde{}} h^{\hat{}} = k_1^{\hat{}\tilde{}} k_2^{\hat{}} g^{\hat{}}$. Here $f^{\hat{}}$ denotes the Mellin transform of the function f and $f^{\tilde{}}(t) = f(-t)$, $-\infty < t < \infty$. Assuming now that $v^{\hat{}}(s) V^{\hat{}}(1-s) / \big(k_1^{\hat{}}(1-s) k_2^{\hat{}}(s)\big) = 1/P(1-s)$, $s = \frac{1}{2} + it$ $(-\infty < t < \infty) \in E_0$—the so called Laguerre-Pólya class ([6], p.174) — this leads us to the inversion formula

$$\frac{1}{P(1-\theta)} h(x) = g(x), \quad 0 < x < \infty, \ \theta = -x(d/dx).$$

A particular case of this inversion formula has recently been considered by NASIM [2].

Suppose again that the system of equations (1.1) holds. Let $h = g$; let $k_1^{\hat{}} = (V^{\hat{}} Q)/P^{\tilde{}}$ and $k_2^{\hat{}} = (v^{\hat{}} P)/Q^{\tilde{}}$ for some functions P and Q in $L^\infty(-\infty, \infty)$. Then our problem reduces to that of iterations of Laplace transforms, and it is seen

---

AMS (MOS) subject classifications (1970). Primary 42A68, 44A05; Secondary 44A35.

Key words and phrases. Convolution, Mellin transform, multiplication operator, Laguerre—Pólya class.

that a necessary and sufficient condition for the pair of functions f and g to be k-transforms of each other, i.e. for the system of equations

$$(1.2)\qquad \begin{aligned}(V\circ f)(u) &= \int_0^\infty k_1(ux)g(x)\,dx,\\ (v\circ g)(u) &= \int_0^\infty k_1(ux)f(x)\,dx\end{aligned}$$

to hold, is that $V^{\wedge}QQ^{\sim}=v^{\wedge}PP^{\sim}$ and $(v^{\wedge}P)/Q^{\sim}\in L^2(-\infty,\infty)\cap L^\infty(-\infty,\infty)$. A suitable choice of v and V (see [1]) in (1.2) then leads us to the reciprocal formulae of WATSON ([4], Theorem 129) and the unsymmetrical formulae of HARDY and TITCHMARSH ([4], p. 226).

2. Some notations

In the sequel, we denote $L^2(0,\infty)$ by L_2 and $L^r(-\infty,\infty)$, $1\leqq r\leqq\infty$, by L^r. M will denote the Mellin transform operator, and the Mellin transform of a function f, whenever it exists, will be denoted by $f^{\wedge}$. For the properties of Mellin transforms we refer the reader to [4]. The functions $f^{\perp}(x)$, $0<x<\infty$, and $f^{\sim}(t)$, $-\infty<t<\infty$, will be defined by $f^{\perp}(x)=x^{-1}f(x^{-1})$ and $f^{\sim}(t)=f(-t)$. Although we do not explicitly say so in the sequel, almost all our results are to be taken as holding almost everywhere only. Also we assume our functions to be non-trivial.

3. Main results

Let the convolution $(f\circ g)$, whenever it exists, of a pair of functions f and g on $(0,\infty)$ be defined by

$$(f\circ g)(u)=\int_0^\infty f(ux^{-1})g(x)\,dx/x.$$

Then, given f and $g\in L_2$, it follows from the properties of Mellin transforms that

$$(3.1a)\qquad (f\circ g)(u)=M^{-1}(f^{\wedge}g^{\wedge})(u)$$

and

$$(3.1b)\qquad (f\circ g)^{\perp}(u)=M^{-1}(f^{\wedge}g^{\wedge\sim})(u).$$

THEOREM 1. *Let v, V be some functions in L_2. The functions f, g and h, all $\in L_2$, satisfy the system of equations*

$$(3.2)\qquad (V\circ f)=(k_1\circ g^{\perp}),\quad (v\circ h)=(k_2\circ f^{\perp})$$

if and only if $k_1^{\wedge}\in L^2\cap L^\infty$, $k_2^{\wedge}\in L^2\cap L^\infty$ and $v^{\wedge}V^{\wedge\sim}h^{\wedge}=k_1^{\wedge\sim}k_2^{\wedge}g^{\wedge}$.

PROOF. Suppose firstly that (3.2) is satisfied; then the bounded linear transformations (of L_2 into itself)

$$S_1g=(k_1\circ g^{\perp}),\quad S_2f=(k_2\circ f^{\perp})$$

satisfy the functional equation

$$S_i t(a) = a^{-1} t(a^{-1}) S_i, \qquad 0 < a < \infty, \quad i = 1, 2,$$

where $t(a)$ is the operator $\big(t(a)f\big)(x) = f(ax)$. It follows from [3] that there exist functions K_1 and $K_2 \in L^\infty$ such that

$$(S_1 g)\hat{} = K_1 g\hat{}\,\tilde{} \quad \text{and} \quad (S_2 f)\hat{} = K_2 f\hat{}\,\tilde{}.$$

We thus have that

$$V\hat{}\,f\hat{} = K_1 g\hat{}\,\tilde{} \quad \text{and} \quad v\hat{}\,h\hat{} = K_2 f\hat{}\,\tilde{},$$

which implies that

$$K_2 K_1\tilde{}\, g\hat{} = v\hat{}\, V\hat{}\,\tilde{}\, h\hat{}. \tag{3.3}$$

Now set $K_2 g\hat{} = b\hat{}$; then $b\hat{} \in L^2$, and $h\hat{} = (K_1\tilde{}/v\hat{}\,V\hat{}\,\tilde{})\,b\hat{}$. Set $K = K_1\tilde{}/(v\hat{}\,V\hat{}\,\tilde{})$; then the multiplication transform $P_K b\hat{} = K b\hat{}$ is bounded on L^2 into itself, and so it follows that $K \in L^\infty$. Since $v\hat{}\,V\hat{}\,\tilde{} \in L^2$, we now have that $v\hat{}\,V\hat{}\,\tilde{}\,K \in L^2$, i.e. $K_1\tilde{}$, and so also $K_1, \in L^2$. A similar argument shows that K_2 also $\in L^2$. Hence K_i, $i = 1, 2, \in L^2 \cap L^\infty$. From the uniqueness of the Mellin transform on L^2 we now have that $K_i = k_i\hat{}$, $i = 1, 2$. Substituting in (3.3), this implies that

$$k_1\hat{}\,\tilde{}\, k_2\hat{}\, g\hat{} = v\hat{}\, V\hat{}\,\tilde{}\, h\hat{},$$

thereby proving the necessity of the conditions.

To see the sufficiency of the conditions, we note that $v\hat{}\,h\hat{} = (k_1\hat{}\,\tilde{}\,k_2\hat{}/V\hat{}\,\tilde{})\,g\hat{} \in \in L^2$. Set $(k_1\hat{}\,\tilde{}\,g\hat{})/V\hat{}\,\tilde{} = f\hat{}\,\tilde{}$. Since $k_2\hat{}\,g\hat{} \in L^2$, and since the multiplication operator $(k_1\hat{}\,\tilde{}/V\hat{}\,\tilde{})\,b\hat{}$, $b\hat{} = k_2\hat{}\,g\hat{}$, is continuous on L^2 into itself, we have that $(k_1\hat{}\,\tilde{}/V\hat{}\,\tilde{}) \in L^\infty$. This implies that $f\hat{}\,\tilde{} = (k_1\hat{}\,\tilde{}/V\hat{}\,\tilde{})\,g\hat{}$, and so also $f\hat{}, \in L^2$. Thus $V\hat{}\,f\hat{} \in L^2$. An application of (3.1) now shows that (3.2) holds. This completes the proof.

We say that $E(s)$, $s = \frac{1}{2} + it$, belongs to the class E_0, the Laguerre—Pólya class, if

$$E(s) = \exp(bs) \prod_{n=1}^{\infty} (1 - s/a_n) \exp(s/a_n),$$

where b, a_n are some real constants and $\sum\limits_{n=1} a_n^{-2} < \infty$. It is known ([6], p. 176) that if $1/E(1-s) \in E_0$, then $E(s) \in L^2(\frac{1}{2} - i\infty, \frac{1}{2} + i\infty)$. The following theorem, due to Widder [5], will be used in our next result.

Theorem W. *If* $E(s) \in E_0$,

$$\varphi(x) = (1/2\pi i) \int_{-i\infty}^{+i\infty} x^{-s}/E(s)\, ds,$$

and

$$h(x) = \int_0^\infty \varphi(xy^{-1})\, g(y)\, dy/y$$

for almost all $x > 0$, *then* $E(\theta) h(x) = g(x)$, *where* $\theta = -x(d/dx)$.

THEOREM 2. *If there exist functions f, g, h, v and V, all $\in L_2$, such that* (3.2) *is satisfied, and if* $1/P^{\sim}=(v^{\wedge}V^{\wedge\sim})/(k_1^{\wedge\sim}k_2^{\wedge})\in E_0$ *on* $s=\frac{1}{2}+it$, $-\infty<t<\infty$, *then* $\dfrac{1}{P(1-\theta)}h(x)=g(x)$.

PROOF. By Theorem 1, $v^{\wedge}V^{\wedge\sim}h^{\wedge}=k_1^{\wedge\sim}k_2^{\wedge}g^{\wedge}$, i.e., $h^{\wedge}=P^{\sim}g^{\wedge}$. The continuity of the multiplication operator $g^{\wedge}\to P^{\sim}g^{\wedge}$ thus implies that $P^{\sim}$, and so also P, $\in L^{\infty}$. Now set $k_1^{\wedge\sim}g^{\wedge}=b^{\wedge}$. Then $b^{\wedge}\in L^2$, and the continuity of the multiplication transform $b^{\wedge}\to(k_2^{\wedge}/(v^{\wedge}V^{\wedge\sim}))b^{\wedge}$ implies that $k_2^{\wedge}/(v^{\wedge}V^{\wedge\sim})\in L^{\infty}$. In consequence $P^{\sim}$, and so also P, $\in L^2$. Thus $P\in L^2\cap L^{\infty}$. Let $p=(1/\sqrt{2\pi})M^{-1}P$. Then, since $h^{\wedge}=P^{\sim}g^{\wedge}$, we have from the properties of Mellin transforms that

$$h(x)=\int_0^{\infty} p(y)g(xy)\,dy=\int_0^{\infty} g(u)\varphi(xu^{-1})\,du/u,$$

where we have set $\varphi(u)=u^{-1}p(u^{-1})$. Since

$$1/\varphi^{\wedge}(s)=1/p^{\wedge}(1-s)=1/P(1-s)\in E_0, \qquad s=\tfrac{1}{2}+it,\ -\infty<t<\infty,$$

we have, upon taking $E(s)=1/P(s)$ in Theorem W, that $[1/P(1-\theta)]h(x)=g(x)$, as required.

4. Applications

Let $v, V\in L_2$. We say that the pair of functions f and $g\in L_2$ are (k_1, k_2)-transforms of each other (or k-transforms of each other if $k_1=k_2$) if there exist functions k_1 and k_2 such that

$$(V\circ f)=(k_1\circ g^{\perp}) \quad\text{and}\quad (v\circ g)=(k_2\circ f^{\perp}).$$

THEOREM 3. *Let θ, v and V be some functions in L_2, and let P and Q be some functions in L^{∞} such that $p=M^{-1}(V^{\wedge}Q)$ and $q=M^{-1}(v^{\wedge}P)$. Then there exist functions f and $g\in L_2$ such that*

$$(V\circ f)=(p\circ\varphi^{\perp}),\quad (v\circ g)=(q\circ\varphi). \tag{4.1}$$

A necessary and sufficient condition for f and g to be (k_1, k_2)-transforms of each other is that

$$(V^{\wedge}Q/P^{\sim})=k_1^{\wedge}\in L^2\cap L^{\infty} \quad\text{and}\quad (v^{\wedge}P/Q^{\sim})=k_2^{\wedge}\in L^2\cap L^{\infty}.$$

PROOF. Since P and $Q\in L^{\infty}$, $(Q\varphi^{\wedge\sim})$ and $(P\varphi^{\wedge})$ are in L^2, and so there exist functions f and $g\in L_2$ such that $f=M^{-1}(Q\varphi^{\wedge\sim})$ and $g=M^{-1}(P\varphi^{\wedge})$. This implies that

$$V^{\wedge}f^{\wedge}=V^{\wedge}Q\varphi^{\wedge\sim}=p^{\wedge}\varphi^{\wedge\sim} \tag{4.2}$$

and

$$v^{\wedge}g^{\wedge}=v^{\wedge}P\varphi^{\wedge}=q^{\wedge}\varphi^{\wedge}. \tag{4.3}$$

An application of (3.1) now shows that (4.1) holds. Again since, by (4.2) and (4.3),

$$V^{\wedge}f^{\wedge}=(p^{\wedge}v^{\wedge\sim}/q^{\wedge\sim})g^{\wedge\sim}$$

and

$$v\hat{}\, g\hat{} = (q\hat{}\, V\hat{}\tilde{}/p\hat{}\tilde{})f\hat{}\tilde{},$$

we have from Theorem 1 (upon letting $h=g$) that f and g are (k_1, k_2)-transforms of each other if and only if

$$(p\hat{}\, v\hat{}\tilde{}/q\hat{}\tilde{}) = (V\hat{}\, Q/P\tilde{}) \in L^2 \cap L^\infty$$

and

$$(q\hat{}\, V\hat{}\tilde{}/p\hat{}\tilde{}) = (v\hat{}\, P/Q\tilde{}) \in L^2 \cap L^\infty.$$

This completes the proof.

REMARK. Theorem 3 contains Theorem I of [1]; for it is clear from Theorem 3 that a necessary and sufficient condition for the functions f and g of (3.4) to be k-transforms of each other is that $(V\hat{}\, Q/P\tilde{}) = (v\hat{}\, P/Q\hat{}) \in L^2 \cap L^\infty$. The condition that $(V\hat{}\, Q/P\tilde{})$ must belong to L^∞, as well as being an element of L^2, although not explicitly stated in Theorem I of [1] is implicit in the fact that the multiplication transform $g\hat{}\tilde{} \to (V\hat{}\, Q/P\tilde{})g\hat{}\tilde{}$ is continuous on L^2 into itself.

Now let

$$v(x) = V(x) = \begin{cases} 0 & \text{if } 0 < x < 1 \\ x^{-1} & \text{if } 1 < x. \end{cases}$$

Then $v\hat{}(t) = V\hat{}(t) = 1/(\frac{1}{2} - it)$, $-\infty < t < \infty$, and it is seen that (3.2) reduces to

$$\text{(4.4)} \qquad \begin{aligned} \int_0^u f(y)\, dy &= u \int_0^\infty k_1(ux) g(x)\, dx, \\ \int_0^u h(y)\, dy &= u \int_0^\infty k_2(ux) f(x)\, dx. \end{aligned}$$

Now set $yk_1(y) = p_1(y)$ and $yk_2(y) = p_2(y)$. Then (4.4) reduces to

$$\text{(4.5)} \qquad \begin{aligned} \int_0^u f(y)\, dy &= \int_0^\infty p_1(ux) g(x)\, dx/x, \\ \int_0^u h(y)\, dy &= \int_0^\infty p_2(ux) f(x)\, dx/x. \end{aligned}$$

THEOREM 4. *If there exist functions f, g and $h \in L_2$ such that* (4.5) *is satisfied, and if* $(1/(p_1\hat{}\tilde{}\, p\hat{})) = (1/P\tilde{}) \in E_0$ *on* $s = \frac{1}{2} + it$, $-\infty < t < \infty$, *then* $(1/P(1-\theta))h(x) = g(x)$; $\theta = -x(d/dx)$.

PROOF. Since $k_i\hat{}(s) = (p_i\hat{}(s)/(1-s))$, $i = 1, 2$, upon substituting for $v\hat{}$, $V\hat{}$, $k_1\hat{}$ and $k_2\hat{}$ we see that the hypotheses of Theorem 2 are satisfied.

We remark that the preceding theorem properly contains the Main Theorem of Nasim [2] (which, indeed, is the non-integrated version of our result).

REFERENCES

[1] DUGGAL, B. P., On Fourier transform type operators, *Period. Math. Hungar.* **9** (1978), 55—61. *MR* **57** #10352.
[2] NASIM, C., An inversion formula for a class of integral transforms, II, *J. Math. Anal. Appl.* **67** (1979), 163—170. *MR* **80m**: 44 002.
[3] SNOO, H. S. V. DE, A note on Watson transforms, *Proc. Cambridge Philos. Soc.* **73** (1973), 83—85.
[4] TITCHMARSH, E. C., *Introduction to the theory of Fourier integrals*, Clarendon Press, Oxford, 1948.
[5] WIDDER, D. V., Inversion of convolution transforms by use of series, *J. Analyse Math.* **21** (1968), 293—312.
[6] WIDDER, D. V., *An introduction to transform theory*, Academic Press, New York, 1971.

(Received February 25, 1980)

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF NAIROBI
P. O. BOX 30197
NAIROBI
KENYA

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 363—374

ON THE ORDER OF CONVERGENCE OF FINITE ELEMENT METHODS FOR THE THIRD BOUNDARY VALUE PROBLEM

by

L. VEIDINGER

Finite element methods for the so-called third boundary value problem have been investigated by many authors (see, for example [1]–[5]). It is well-known that for this problem the functions admissible in the associated variational problem are not required to satisfy any boundary conditions and hence regions in two or more dimensions of general shape may be treated without any difficulty. In the present paper we shall obtain error bounds for some finite element methods in two dimensions under weak assumptions on the regularity of the boundary of the region.

1. Let R be a bounded open plane region whose boundary C consists of a finite number of piecewise analytic simple closed curves. For the sake of simplicity we shall assume that the boundary C consists of two analytic arcs which meet at the corner $A=(0, 0)$ and form an interior angle $\pi\alpha$ $(0<\alpha<2)$ there. The general case can be treated in the same way.

We consider the boundary value problem

$$(1)\qquad Lu(x, y) = g(x, y), \quad (x, y)\in R,$$

$$E\frac{\partial u(x, y)}{\partial N}+\sigma(x, y)u(x, y) = 0, \quad (x, y)\in C,$$

where

$$Lu \equiv \frac{\partial}{\partial x}\left[a(x, y)\frac{\partial u}{\partial x}\right]+\frac{\partial}{\partial x}\left[b(x, y)\frac{\partial u}{\partial y}\right]+\frac{\partial}{\partial y}\left[b(x, y)\frac{\partial u}{\partial x}\right]+$$

$$+\frac{\partial}{\partial y}\left[c(x, y)\frac{\partial u}{\partial y}\right]-f(x, y)u,$$

$$E=\{[a\cos(n, x)+b\cos(n, y)]^2+[b\cos(n, x)+c\cos(n, y)]^2\}^{1/2}$$

and N is the conormal with direction cosines

$$\cos(N, x)=\frac{1}{E}[a\cos(n, x)+b\cos(n, y)], \quad \cos(N, y)=\frac{1}{E}[b\cos(n, x)+c\cos(n, y)].$$

AMS (MOS) subject classifications (1970). Primary 65N30.
Key words and phrases. Elliptic equations, boundary value problems, finite element methods.

Here n is the outward normal to C. At the corner A we require that

$$E_1\frac{\partial u}{\partial N_1}+\sigma(A)u(A)=E_2\frac{\partial u}{\partial N_2}+\sigma(A)u(A)=0,$$

where

$$E_i=\{[a(A)\cos(n_i,x)+b(A)\cos(n_i,y)]^2+$$
$$+[b(A)x\cos(n_i,x)+c(A)\cos(n_i,y)]^2\}^{1/2}\quad(i=1,2),$$

n_1 and N_1 are the "left-hand" normal and conormal, n_2 and N_2 are the "right-hand" normal and conormal to the boundary C at the corner A.

Let the coefficients $a(x,y)$, $b(x,y)$, $c(x,y)$, $f(x,y)$ and the right-hand side $g(x,y)$ be infinitely differentiable in R. Suppose that at all points of R

(2) $$a\xi^2+2b\xi\eta+c\eta^2\geqq\nu(\xi^2+\eta^2)\quad(\nu=\text{const.}>0)$$

for all real ξ,η. Moreover, we assume that $f(x,y)\geqq 0$ in R. On the boundary C we assume that the function $\sigma(x,y)$ is infinitely differentiable,

(3) $$\sigma\geqq 0 \text{ on } C,\quad \sigma\not\equiv 0,\ \text{ if }\ \min_{(x,y)\in R}f(x,y)=0.$$

Let Ω be an open bounded region in the plane of R and let Γ be the boundary of Ω. We denote by $W_2^{(s)}(\Omega)$ the Hilbert space of all functions which, together with their generalized partial derivatives up to the sth order, belong to $L_2(\Omega)$. The norm is given by

$$\|v\|^2_{s,\Omega}=\sum_{j=0}^{s}|v|^2_{j,\Omega},\quad\text{where}\quad |v|^2_{j,\Omega}=\sum_{|i|=j}\|D^i v\|^2_{L_2(\Omega)}.$$

Here $i=(i_1,i_2)$, $|i|=i_1+i_2$, $D^i v=\dfrac{\partial^{|i|}v}{\partial x^{i_1}\partial y^{i_2}}$. We denote by $C^{(s)}(\bar\Omega)$ the set of functions which have continuous partial derivatives of all orders up to s in the closed region $\bar\Omega=\Omega\cup\Gamma$.

It is well-known that under the above assumptions the solution $u(x,y)$ of the boundary value problem (1) minimizes the functional

(4) $$F(v)=\iint_R\left[a\left(\frac{\partial v}{\partial x}\right)^2+2b\frac{\partial v}{\partial x}\frac{\partial v}{\partial y}+c\left(\frac{\partial v}{\partial y}\right)^2+fv^2+2gv\right]dx\,dy+\int_C\sigma v^2\,dC$$

in the space $W_2^{(1)}(R)$.

LEMMA 1. *Let $u(x,y)$ be the solution of the boundary value problem* (1) *and let*

(5) $$x^*=k_Ax+l_Ay,\quad y^*=m_Ax+n_Ay$$

be a linear transformation which transforms the operator into the normal form at the point $A=(0,0)$. Let $r=\sqrt{x^2+y^2}$. If the transformation (5) *transforms the angle $\pi\alpha$ into an angle $\pi\alpha^*(0<\alpha^*<2)$*[1] *then for $n=0,1,2,\ldots$ we have*

$$u(x,y)=\sum_{0<\frac{m}{\alpha^*}+p\leqq n+1}a_{m,p}(\varphi)r^{\frac{m}{\alpha^*}+p}(\log r)^{q_{m,p}}+w(x,y)$$

[1] It is easy to show that α^* depends only on α, $a(A)$, $b(A)$, $c(A)$ and does not depend on the choice of the transformation (5).

where m is a positive integer, p and $q_{m,p}$ are non-negative integers, the coefficients $a_{m,p}(\varphi)$ are infinitely differentiable functions of the polar coordinate $\varphi=\operatorname{arctg}\frac{y}{x}$, $q_{1,0}=0$ if $\alpha^*\neq\frac{1}{s}$ (s is an integer), $q_{1,0}=1$ if $\alpha^*=1,\frac{1}{2},\dots,w(x,y)\in W_2^{(n+2)}(R)$.

This lemma follows from the results of KONDRAT'EV and WIGLEY (see [6]—[8]).

2. We cover R by a finite number of arbitrary real (non-curved) triangles such that any two triangles are either disjoint or have a common vertex or a common side. We retain only those triangles T for which

$$\iint_{T\cap R} dx\,dy>0,$$

i.e., for which T and R have some common area. Denote by M_h the set of triangles covering R, where h is the largest side of the set of triangles covering R. Let ϑ be the smallest angle of all triangles $T\in M_h$. In the sequel we assume that

$$h<c_1\bar{h},\quad \vartheta\geqq\vartheta_0>0, \tag{6}$$

where $\bar{h}$ is the smallest side of all triangles $T\in M_h$, c_1 and ϑ_0 do not depend on h. Denote by R_h the union of all triangles $T\in M_h$. We emphasize that (just as in the case of the Neumann problem (see [9])), neither curved triangles nor special triangles near the boundary are necessary, the triangles T are real and arbitrary. For example, we may choose a regular mesh consisting of right isosceles triangles; this mesh may be especially advantageous in practical computations.

Let $k\geqq2$, $T\in M_h$ and let P_1, P_2, P_3 be the vertices of T. We choose the following nodes for T:

(a) the vertices of T,

(b) the $k-2$ points on each side of T that divide the side into $k-1$ equal segments and

(c) $0{,}5(k-3)(k-2)$ distinct points in the interior of T, chosen so that if a polynomial of degree $k-4$ vanishes at all of them, then it vanishes identically.

Here (c) applies only to $k\geqq4$ and (b) applies only to $k\geqq3$.

Denote by $P_\mu(\mu=1,2,\dots,3k-3+0{,}5(k-2)(k-3))$ the nodes of T and let $p_\mu=p(P_\mu)$ where $p(P_\mu)$ is any real-valued function defined at the nodes P_μ. Let $p(x,y)$ be the Lagrange interpolation polynomial of degree $\leqq k-1$ determined by its values p_μ at the points P_μ.

Denote by $H_{k-1}(R)$ the set of all functions, defined and continuous on R, which are equal on each triangle $T\in M_h$ to the corresponding polynomial $p(x,y)$. It is clear that $H_{k-1}(R)\subset W_2^{(1)}(R)$. The solution $u(x,y)$ of the boundary value problem (1) is approximated by the function $u_h(x,y)$ which minimizes the functional (4) in the space $H_{k-1}(R)$. The existence and uniqueness of $u_h(x,y)$ follows immediately from our assumptions.

LEMMA 2. *Let* $T \in M_h$ *and let* $v(x, y)$ *be a real-valued function defined on* $\bar{T}$. *Define the interpolate* $p(x, y)$ *of* $v(x, y)$ *by requiring that* $v(x, y)-p(x, y)$ *vanish at all the nodes. If* $v(x, y) \in C^{(k)}(\bar{T})$, *then*

$$\max_{(x,y)\in T} |v(x, y)-p(x, y)| \leqq c_2 h^k \max_{\substack{(x,y)\in T\\ |i|=k}} |D^i v(x, y)| \tag{7}$$

where $i=(i_1, i_2)$, $|i|=i_1+i_2$. *If* $v(x, y) \in W_2^{(k)}(T)$, *then*

$$\|v-p\|_{1,T} \leqq c_3 h^{k-1} |v|_{k,T}. \tag{8}$$

Here c_2 *and* c_3 *are positive constants which do not depend on the triangle* T *and the function* $v(x, y)$.

For a proof, see [5], p. 139 and p. 144, respectively.

LEMMA 3. *Let* $T \in M_h$ *and let* $p(x, y)$ *be the Lagrange interpolation polynomial of degree* $\leqq k-1$ *determined by the conditions* $p(P_\mu)=p_\mu(\mu=1, 2, \dots, 3k-3+ +0,5(k-2)(k-3))$. *Then*

$$\max_{(x,y)\in T} |p(x, y)| \leqq c_4 q, \tag{9}$$

$$|p|_{1,T} \leqq c_5 q \tag{10}$$

and

$$\|p\|_{L_2(T)} \leqq c_6 h q, \tag{11}$$

where $q=\max |p_\mu|$, $\mu = 1, 2, \dots, 3k-3+0,5(k-2)(k-3)$, c_4, c_5 *and* c_6 *are positive constants which do not depend on the triangle* T *and the parameters* p_μ.

For a proof, see [10], p. 404.

THEOREM 1. *Let* $u(x, y)$ *be the solution of the boundary value problem* (1) *and let* $u_h(x, y)$ *be the function which minimizes the functional* (4) *in the space* $H_{k-1}(R)$. *Then*

$$\|u-u_h\|_{1,R} < c_7 h^{\beta^*} \tag{12}$$

and

$$\max_{(x,y)\in \bar{R}} |u(x, y)-u_h(x, y)| < c_8 h^{\beta^*} |\log h|^{1/2}, \tag{13}$$

where

$$\beta^* = \begin{cases} \dfrac{1}{\alpha^*} & \text{if } 1<\alpha^*<2, \\ 1-\varepsilon & \text{if } \alpha^*=1, \\ 1 & \text{if } \dfrac{1}{k} \leqq \alpha^* < 1, \\ k-1 & \text{if } 0<\alpha^*<\dfrac{1}{k} \quad \text{or if there are no corners,} \end{cases}$$

c_7 *and* c_8 *are positive constants which depend only on* k, *the region* R, *the function* $\sigma(x, y)$, *the coefficients of the operator* L *and the right-hand side* $g(x, y)$, ε *is any positive real number.*

PROOF. Let the functional $D(v)$ be defined by

$$D(v)=\iint_R\left[a\left(\frac{\partial v}{\partial x}\right)^2+2b\frac{\partial v}{\partial x}\frac{\partial v}{\partial y}+c\left(\frac{\partial v}{\partial y}\right)^2+fv^2\right]dx\,dy+\int_C \sigma v^2\,dC$$

for all $v(x,y)\in W_2^{(1)}(R)$ and let $z(x,y)\in H_{k-1}(R)$. Then, we have (see [4], p. 398)

$$D(u-u_h)\leqq D(u-z). \tag{14}$$

If $\min\limits_{(x,y)\in\bar R} f(x,y)=f_0>0$, then

$$D(v)\geqq \nu\iint_R\left[\left(\frac{\partial v}{\partial x}\right)^2+\left(\frac{\partial v}{\partial y}\right)^2\right]dx\,dy+f_0\iint_R v^2\,dx\,dy$$

for all $v\in W_2^{(1)}(R)$. Hence by the inequality (14) it follows that

$$\|u-u_h\|_{1,R}^2\leqq c_9\left[\|u-z\|_{1,R}^2+\int_C\sigma(u-z)^2\,dC\right] \tag{15}$$

for all $z\in H_{k-1}(R)$. If $\min\limits_{(x,y)\in R} f(x,y)=0$ then from our assumption (3) it follows that $\sigma\geqq\sigma_0>0$ on a part C_0 of C for σ_0 being sufficiently small. Then

$$D(v)\geqq \nu\iint_R\left[\left(\frac{\partial v}{\partial x}\right)^2+\left(\frac{\partial v}{\partial y}\right)^2\right]dx\,dy+\sigma_0\int_{C_0} v^2\,dC \tag{16}$$

for all $v\in W_2^{(1)}(R)$. By the generalized Friedrichs inequality (see [11], p. 212) we have

$$\iint_R v^2\,dx\,dy\leqq c_{10}\left\{\iint_R\left[\left(\frac{\partial v}{\partial x}\right)^2+\left(\frac{\partial v}{\partial y}\right)^2\right]dx\,dy+\int_{C_0}v^2\,dC\right\}. \tag{17}$$

From (14), (16) and (17) we again get the inequality (15).

By Lemma 1 we have for $n=0,1,\ldots$

$$u(x,y)=\sum_{0<\frac{m}{\alpha^*}+p\leqq n+1} a_{m,p}(\varphi)r^{\frac{m}{\alpha^*}+p}(\log r)^{q_{m,p}}+w(x,y).$$

By the Calderón extension theorem (see, for example, [12], p. 171) there exists a function $w_{\text{ext}}(x,y)\in W_2^{(n+2)}(R_h)$ such that $w_{\text{ext}}(x,y)=w(x,y)$ for all $(x,y)\in R$ and

$$\|w_{\text{ext}}\|_{n+2,R_h}\leqq c_{11}\|w\|_{n+2,R}, \tag{18}$$

where c_{11} is a positive constant which depends only on the region R. Let the function $u_{\text{ext}}(x,y)$ be defined by

$$u_{\text{ext}}(x,y)=\sum_{0<\frac{m}{\alpha^*}+p\leqq n+1} a_{m,p}(\varphi)r^{\frac{m}{\alpha^*}+p}(\log r)^{q_{m,p}}+w_{\text{ext}}(x,y) \tag{19}$$

for all $(x, y)\in R_h$. We define the interpolate $\varrho_h u_{\text{ext}}(x, y)$ of $u_{\text{ext}}(x, y)$ by requiring that $u_{\text{ext}}(x, y)-\varrho_h u_{\text{ext}}(x, y)$ vanish at all the nodes. Then from (15) it follows that

$$(20)\qquad \|u-u_h\|^2_{1,R} \leqq c_9\Big[\|u_{\text{ext}}-\varrho_h u_{\text{ext}}\|^2_{1,R_h}+\int\limits_C \sigma(u-\varrho_h u_{\text{ext}})^2\,dC\Big].$$

Denote by $r(A, T)$ the distance from the corner A to the triangle $T\in M_h$. Let us first consider a triangle $T\in M_h$ such that $r(A, T)>r_1$, where r_1 is a fixed small positive real number. Setting $n=k$ in (19) we obtain from the Sobolev embedding theorem (see, for example, [12], p. 32) that if $T\in M_h$, $r(A, T)>r_1$ then

$$(21)\qquad \max_{(x,y)\in T}|D^i u_{\text{ext}}(x, y)|<c_{12},$$

where $i=(i_1, i_2)$, $|i|=i_1+i_2$ and c_{12} is a positive constant which depends only on the data of the boundary value problem (1). Substituting this into (8) we have

$$(22)\qquad \|u_{\text{ext}}-\varrho_h u_{\text{ext}}\|^2_{1,T}<c_{13}h^{2k-2}\operatorname{mes}T,$$

where c_{13} is a positive constant which depends only on the data of the boundary value problem (1). Summing (22) over all triangles $T\in M_h$ for which $r(A, T)>r_1$ we obtain that

$$(23)\qquad \sum_{T\in M_h,\, r(A,T)>r_1}\|u_{\text{ext}}-\varrho_h u_{\text{ext}}\|^2_{1,T}=O(h^{2k-2}).$$

Substituting (21) into (7) we have for all triangles $T\in M_h$ such that $r(A, T)>r_1$

$$(24)\qquad \max_{(x,y)\in T}|u(x, y)-\varrho_h u(x, y)|<c_{14}h^k,$$

where c_{14} is a positive constant which depends only on the data of the boundary value problem (1). Denote by C_1 the part of C which is covered by triangles $T\in M_h$ such that $r(A, T)>r_1$. Then from (24) it follows that

$$(25)\qquad \int\limits_{C_1}\sigma(u-\varrho_h u_{\text{ext}})^2\,dC=O(h^{2k}).$$

From (23) and (25) we obtain

$$(26)\qquad \sum_{T\in M_h,\, r(A,T)>r_1}\|u_{\text{ext}}-\varrho_h u_{\text{ext}}\|^2_{1,T}+\int\limits_{C_1}\sigma(u-\varrho_h u_{\text{ext}})^2\,dC=O(h^{2k-2}).$$

If $\dfrac{1}{k-1}<\alpha^*<2$ and $\alpha^*\neq 1, \dfrac{1}{2}, \dots, \dfrac{1}{k-2}$, then by (8) we have

$$\sum_{T\in M_h,\, h\leqq r(A,T)\leqq r_1}\|u_{\text{ext}}-\varrho_h u_{\text{ext}}\|^2_{1,T}=$$

$$(27)\qquad =O\Big(h^{2k-2}\sum_{T\in M_h,\, h\leqq r(A,T)\leqq r_1}(\operatorname{mes}T)[r(A, T)]^{\frac{2}{\alpha^*}-2k}\Big)=$$

$$=O\Big(h^{2k-2}\iint\limits_{h\leqq r\leqq r_1} r^{\frac{2}{\alpha^*}-2k}\,dx\,dy\Big)=O\Big(h^{2k-2}\int_h^{r_1} r^{\frac{2}{\alpha^*}-2k+1}\,dr\Big)=O(h^{\frac{2}{\alpha^*}}),$$

where $r=\sqrt{x^2+y^2}$.

If $T \in M_h$, $h \leqq r(A, T) \leqq r_1$ then by (7) we have

$$\max_{(x,y)\in T} |u_{\text{ext}}(x, y) - \varrho_h u_{\text{ext}}(x, y)| < c_{15} h^k [r(A, T)]^{\frac{1}{\alpha^*}-k} < c_{15} h^{\frac{1}{\alpha^*}}. \tag{28}$$

Denote by C_2 the part of C which is covered by triangles $T \in M_h$ such that $h \leqq r(A, T) \leqq r_1$. Then from (28) it follows that

$$\int_{C_2} \sigma(u - \varrho_h u_{\text{ext}})^2 \, dC = O(h^{\frac{2}{\alpha^*}}). \tag{29}$$

(27) and (29) imply

$$\sum_{T\in M_h,\, h\leqq r(A,T)\leqq r_1} \|u_{\text{ext}} - \varrho_h u_{\text{ext}}\|_{1,T}^2 + \int_{C_2} \sigma(u - \varrho_h u_{\text{ext}})^2 \, dC = O(h^{\frac{2}{\alpha^*}}). \tag{30}$$

If $\alpha^* = 1, \frac{1}{2}, \dots, \frac{1}{k-1}$, then in a similar way we get

$$\sum_{T\in M_h,\, h\leqq r(A,T)\leqq r_1} \|u_{\text{ext}} - \varrho_h u_{\text{ext}}\|_{1,T}^2 + \int_{C_2} \sigma(u - \varrho_h u_{\text{ext}})^2 \, dC = O(h^{\frac{2}{\alpha^*}-\varepsilon}), \tag{31}$$

where ε is any positive real number.

If $0 < \alpha^* < \frac{1}{k-1}$, then setting $n = k$ in (19) we find that $u_{\text{ext}}(x, y) \in W_2^{(k)}(G_A)$ where $G_A = \{P \in R_h, 0 < r < r_1\}$ and by (8) we have

$$\sum_{T\in M_h,\, h\leqq r(A,T)\leqq r_1} \|u_{\text{ext}} - \varrho_h u_{\text{ext}}\|_{1,T}^2 = O(h^{2k-2}). \tag{32}$$

If $\frac{1}{k} \leqq \alpha^* < \frac{1}{k-1}$, $T \in M_h$, $h \leqq r(A, T) \leqq r_1$, then from (28) it follows that

$$\max_{(x,y)\in T} |u_{\text{ext}}(x, y) - \varrho_h u_{\text{ext}}(x, y)| < c_{15} h^{\frac{1}{\alpha^*}} < c_{15} h^{k-1}. \tag{33}$$

If $0 < \alpha^* < \frac{1}{k}$ then $u(x, y) \in C^{(k)}(\bar{G}_A)$ and thus by (7) we have for all triangles $T \in M_h$ such that $h \leqq r(A, T) \leqq r_1$

$$\max_{(x,y)\in T} |u_{\text{ext}}(x, y) - \varrho_h u_{\text{ext}}(x, y)| < c_{16} h^k.$$

From the last two inequalities it follows that if $T \in M_h$, $h \leqq r(A, T) \leqq r_1$, $0 < \alpha^* < \frac{1}{k-1}$, then

$$\max_{(x,y)\in T} |u_{\text{ext}}(x, y) - \varrho_h u_{\text{ext}}(x, y)| < c_{17} h^{k-1}. \tag{34}$$

By (32) and (34) we have

$$\sum_{T\in M_h,\, h\leqq r(A,T)\leqq r_1} \|u_{\text{ext}} - \varrho_h u_{\text{ext}}\|_{1,T}^2 + \int_{C_2} \sigma(u - \varrho_h u_{\text{ext}})^2 \, dC = O(h^{2k-2}). \tag{35}$$

(30), (31) and (35) imply that

$$\sum_{T\in M_h,\, h\le r(A,T)\le r_1} \|u_{\text{ext}}-\varrho_h u_{\text{ext}}\|_{1,T}^2+\int_{C_2}\sigma(u-\varrho_h u_{\text{ext}})^2\,dC=O(h^{2\beta^*}). \tag{36}$$

Let us now consider a triangle $T\in M_h$ such that $r(A,T)\le h$. If $1<\alpha^*<2$, then from (19) it follows that

$$|u_{\text{ext}}(x,y)-u(A)|_{1,T}^2=O\Big(\iint_{0\le r\le h} r^{\frac{2}{\alpha^*}-2}\,dx\,dy\Big)=O\Big(\int_0^h r^{\frac{2}{\alpha^*}-1}\,dr\Big)=O(h^{\frac{2}{\alpha^*}}) \tag{37}$$

and

$$\|u_{\text{ext}}(x,y)-u(A)\|_{L_2(T)}^2=O(h^{\frac{2}{\alpha^*}+2}). \tag{38}$$

On the other hand, if we define the interpolate $p(x,y)$ of $u(x,y)$ by requiring that $u(x,y)-p(x,y)$ vanish at all the nodes, then by (10) and (19) we have

$$|u(A)-p(x,y)|_{1,T}=O(h^{\frac{1}{\alpha^*}}). \tag{39}$$

Finally, from (11) and (19) we obtain that

$$\|u(A)-p(x,y)\|_{L_2(T)}=O(h^{\frac{1}{\alpha^*}+1}). \tag{40}$$

(37), (38), (39) and (40) imply that if $T\in M_h$, $r(A,T)\le h$ and $1<\alpha^*<2$, then

$$\|u_{\text{ext}}(x,y)-p(x,y)\|_{1,T}=O(h^{\frac{1}{\alpha^*}}). \tag{41}$$

If $T\in M_h$, $r(A,T)\le h$ and $1<\alpha^*<2$, then by (19) we have

$$\max_{(x,y)\in T}|u_{\text{ext}}(x,y)-u(A)|=O(h^{\frac{1}{\alpha^*}}). \tag{42}$$

On the other hand, from (19) and (9) it follows that

$$\max_{(x,y)\in T}|p(x,y)-u(A)|=O(h^{\frac{1}{\alpha^*}}). \tag{43}$$

Denote by C_3 the part of C which is covered by triangles $T\in M_h$ such that $r(A,T)\le h$.

If $1<\alpha^*<2$ then from (42) and (43) it follows that

$$\int_{C_3}\sigma(u-\varrho_h u_{\text{ext}})^2\,dC=O(h^{\frac{2}{\alpha^*}}). \tag{44}$$

(41) and (44) imply that if $1<\alpha^*<2$, then

$$\|u_{\text{ext}}-\varrho_h u_{\text{ext}}\|_{1,T}^2+\int_{C_3}\sigma(u-\varrho_h u_{\text{ext}})^2\,dC=O(h^{\frac{2}{\alpha^*}}). \tag{45}$$

If $\alpha^*=1$, then in a similar way we get

$$\|u_{\text{ext}}-\varrho_h u_{\text{ext}}\|_{1,T}^2+\int_{C_3}\sigma(u-\varrho_h u_{\text{ext}})^2\,dC=O(h^{2-\varepsilon}). \tag{46}$$

Let us now consider the case when $\frac{1}{k} \leqq \alpha^* \leqq 1$. Then from (19) it follows that the first partial derivatives of $u_{\text{ext}}(x, y)$ are bounded in T and, consequently,

$$|u_{\text{ext}}(x, y) - u(A)|^2_{1,T} = O\Big(\iint\limits_{0 \leqq r \leqq h} dx\,dy\Big) = O(h^2). \tag{47}$$

Further, since $u(A) = w(A)$, from (19) by Taylor's formula we obtain that

$$\|u_{\text{ext}}(x, y) - u(A)\|^2_{L_2(T)} = O(h^{\frac{2}{\alpha^*}+2} + h^4) = O(h^4). \tag{48}$$

On the other hand, from (19), using (10), it follows that

$$|u(A) - p(x, y)|^2_{1,T} = O(h^{\frac{2}{\alpha^*}} + h^2) = O(h^2), \tag{49}$$

where $p(x, y)$ is defined as in (39). Using (11) in place of (10) we get

$$\|u(A) - p(x, y)\|^2_{L_2(T)} = O(h^{\frac{2}{\alpha^*}+2} + h^4) = O(h^4). \tag{50}$$

(47), (48), (49) and (50) imply that

$$\|u_{\text{ext}}(x, y) - p(x, y)\|^2_{1,T} = O(h^2). \tag{51}$$

From (19) by Taylor's formula we obtain that

$$\max_{(x,y)\in T} |u_{\text{ext}}(x, y) - u(A)| = O(h^{\frac{1}{\alpha^*}} + h) = O(h). \tag{52}$$

On the other hand, using (9), we get

$$\max_{(x,y)\in T} |p(x, y) - u(A)| = O(h^{\frac{1}{\alpha^*}} + h) = O(h). \tag{53}$$

From (52) and (53) it follows that

$$\int\limits_{C_3} \sigma(u - \varrho_h u_{\text{ext}})^2\, dC = O(h^2). \tag{54}$$

(51) and (54) imply that

$$\|u_{\text{ext}} - \varrho_h u_{\text{ext}}\|^2_{1,T} + \int\limits_{C_3} \sigma(u - \varrho_h u_{\text{ext}})^2\, dC = O(h^2). \tag{55}$$

Finally, if $0 < \alpha^* < \frac{1}{k}$, then $u_{\text{ext}}(x, y) \in W_2^{(k)}(G_A)$ and by (8) we have

$$\|u_{\text{ext}}(x, y) - p(x, y)\|_{1,T} = O(h^{k-1}). \tag{56}$$

On the other hand, if $0 < \alpha^* < \frac{1}{k}$, then $u_{\text{ext}}(x, y) \in C^{(k)}(\overline{G}_A)$ and thus by (7) we have

$$\max_{(x,y)\in T} |u_{\text{ext}}(x, y) - p(x, y)| < c_{18} h^k,$$

whence

$$\int\limits_{C_3} \sigma(u - \varrho_h u_{\text{ext}})^2\, dC = O(h^{2k}). \tag{57}$$

(56) and (57) imply that

$$\|u_{\text{ext}}-\varrho_h u_{\text{ext}}\|^2_{1,T}+\int_{C_3}\sigma(u-\varrho_h u_{\text{ext}})^2\,dC=O(h^{2k-2}). \tag{58}$$

It follows from (6), (45), (46), (55) and (58) that

$$\sum_{T\in M_h,\, r(A,T)\leqq h}\|u_{\text{ext}}-\varrho_h u_{\text{ext}}\|^2_{1,T}+\int_{C_3}(u-\varrho_h u_{\text{ext}})^2\,dC=O(h^{2\beta^*}). \tag{59}$$

(26), (36) and (59) imply that

$$\|u_{\text{ext}}-\varrho_h u_{\text{ext}}\|^2_{1,R_h}+\int_{C}\sigma(u-\varrho_h u_{\text{ext}})^2\,dC=O(h^{2\beta^*}).$$

Substituting this into (20) we get the inequality (12). The inequality (13) follows directly from (12) and a theorem of V. P. IL'IN (see [13], p. 101). This completes the proof of Theorem 1.

3. If we want to compute also the partial derivatives of $u(x, y)$, then we have to use the Hermite interpolation polynomials introduced by KOUKAL (see [14]). As in the previous section, we cover R by a finite number of arbitrary real triangles. We assume now that the corner A is a vertex of a triangle $T\in M_h$. Let P_1, P_2, P_3 be the vertices of the triangle $T\in M_h$ and let P_0 be the center of gravity of T.

Let $l\geqq 2$, $p^i_\mu=p^i(P_\mu)$ $(\mu=0,1,2,3)$, where $i=(i_1, i_2)$, p^i is any real-valued function defined at the points P_μ and let $p(x, y)$ be the Hermite interpolation polynomial of degree $\leqq 2l-1$ determined by the conditions

$$D^i p(P_j)=p^i_j,\quad |i|\leqq l-1,\quad j=1,2,3,$$
$$D^i p(P_0)=p^i_0,\quad |i|\leqq l-2,$$

where $|i|=i_1+i_2$. Denote by $I_{2l-1}(R)$ the set of all functions, defined on R, which are equal on each triangle $T\in M_h$ to the corresponding polynomial $p(x, y)$. It is easy to show that $I_{2l-1}(R)\subset W_2^{(1)}(R)$. The solution $u(x, y)$ of the boundary value problem (1) is approximated by the function $v_h(x, y)$ which minimizes the functional (4) in the space $I_{2l-1}(R)$. The existence and uniqueness of $v_h(x, y)$ follows immediately from our assumptions.

In the sequel we shall restrict ourselves to the case $l=2$, but our results can be probably generalized to arbitrary l (see [14]). We shall need the following two lemmas.

LEMMA 4. *Let $T\in M_h$ and let $v(x, y)$ be a real-valued function which is defined on T and which is differentiable at the vertices of T. Let $p(x, y)$ be the cubic polynomial determined by the conditions*

$$p(P_\mu)=v(P_\mu)\qquad(\mu=0,1,2,3),$$

$$\frac{\partial p(P_\mu)}{\partial x}=\frac{\partial v(P_\mu)}{\partial x},\ \frac{\partial p(P_\mu)}{\partial y}=\frac{\partial v(P_\mu)}{\partial y}\qquad(\mu=1,2,3).$$

If $v(x, y)\in C^{(4)}(\overline{T})$, then

$$\max_{(x,y)\in T}|v(x, y)-p(x, y)|\leqq c_{19}h^4\max_{\substack{(x,y)\in T\\ |i|=4}}|D^i v(x, y)|.$$

If $v(x, y)\in W_2^{(4)}(T)$, *then*

$$\|v-p\|_{1,T} \leqq c_{20} h^3 |v|_{4,T},$$

where c_{19} *and* c_{20} *are positive constants which do not depend on the triangle* T *and the function* $v(x, y)$.

For a proof, see [5], p. 139 and p. 144, respectively.

LEMMA 5. *Let* $T\in M_h$ *and let* $p_\mu^i = p^i(p_\mu)$ *where* $i=(i_1, i_2)$, $|i|=i_1+i_2\leqq 1$ *and* p^i *is any real-valued function defined at the points* P_μ $(\mu=0, 1, 2, 3)$. *Let* $p(x, y)$ *be the cubic polynomial determined by the conditions*

$$p(P_\mu) = p_\mu^{(0,0)} \quad (\mu = 0, 1, 2, 3),$$

$$\frac{\partial p(P_\mu)}{\partial x} = p_\mu^{(1,0)}, \quad \frac{\partial p(P_\mu)}{\partial y} = p_\mu^{(0,1)} \quad (\mu = 1, 2, 3).$$

Then

$$|p|_{1,T} \leqq c_{21}(q_0 + hq_1 + hq_2),$$

$$\max_{(x,y)\in T} |p(x, y)| \leqq c_{22}(q_0 + hq_1 + hq_2)$$

and

$$\|p\|_{L_2(T)} \leqq c_{23} h(q_0 + hq_1 + hq_2),$$

where

$$q_0 = \max_{\mu=0,1,2,3} |p_\mu^{(0,0)}|, \quad q_1 = \max_{\mu=1,2,3} |p_\mu^{(1,0)}|, \quad q_2 = \max |p_\mu^{(0,1)}|,$$

c_{21}, c_{22} *and* c_{23} *are positive constants which do not depend on the triangle* T *and the parameters* p_μ^i.

For a proof, see [10], p. 404.

Using Lemma 4 and Lemma 5 in place of Lemma 2 and Lemma 3, respectively, and repeating the proof of Theorem 1 (see also [10]) we obtain the following theorem.

THEOREM 2. *Let* $u(x, y)$ *be the solution of the boundary value problem* (1) *and let* $v_h(x, y)$ *be the function which minimizes the functional* (4) *in the space* $I_3(R)$. *Then*

$$\|u-v_h\|_{1,R} < c_{24} h^{\gamma^*}$$

and

$$\max_{(x,y)\in \bar{R}} |u(x, y) - v_h(x, y)| < c_{25} h^{\gamma^*} |\log h|^{1/2},$$

where

$$\gamma^* = \begin{cases} \dfrac{1}{\alpha^*} & \text{if} \quad 1 < \alpha^* < 2, \\ 1-\varepsilon & \text{if} \quad \alpha^* = 1, \\ 1 & \text{if} \quad \dfrac{1}{4} \leqq \alpha^* < 1, \\ 3 & \text{if} \quad 0 < \alpha^* < \dfrac{1}{4} \quad \text{or if there are no corners,} \end{cases}$$

c_{24} *and* c_{25} *are positive constants which depend only on the domain* R, *the function* $\sigma(x, y)$ *the coefficients of the operator* L *and the right-hand side* $g(x, y)$, ε *is any positive real number.*

REFERENCES

[1] OGANESJAN, L. A., Convergence of difference schemes in case of improved approximation of the boundary, *Ž. Vyčisl. Mat. i Mat. Fiz* **6** (1966), 1029—1042 (in Russian). *MR* **34** #7044; erratum **35**, p. 1577.

[2] OGANESJAN, A. and RUHOVEC, L. A., Variational-difference schemes for second order linear elliptic equations in a two-dimensional region with a piecewise-smooth boundary, *Ž. Vyčisl. Mat i Mat. Fiz.* **8** (1968), 97—114 (in Russian). *MR* **38** #1846.

[3] OGANESJAN, L. A. and RUHOVEC, L. A., An investigation of the rate of convergence of variation-difference schemes for second order elliptic equations in a two-dimensional region with smooth boundary, *Ž. Vyčisl. Mat i Mat. Fiz.* **9** (1969), 1102—1120 (in Russian). *MR* **45** #4665.

[4] ZLÁMAL, M., On the finite element method, *Numer. Math.* **12** (1968), 394—409. *MR* **39** #5074.

[5] STRANG, G. and FIX, G., *An analysis of the finite element method,* Prentice-Hall Series in Automatic, Computation, Prentice-Hall Inc., Englewood Cliffs, N. J., 1973. *MR* **56** #1747 (Russian translation: *MR* **56** #16985).

[6] KONDRAT'EV, V. A., Boundary value problems for elliptic equations in domains with conical or angular points, *Trudy Moskov. Mat. Obšč* **16** (1967), 209—292 (in Russian). *MR* **37** #1777.

[7] WIGLEY, N. M., Asymptotic expansions at a corner of solutions of mixed boundary value problems, *J. Math. Mech* **13** (1964), 549—576. *MR* **29** #2516.

[8] WIGLEY, N. M., Mixed boundary value problems in plane domains with corners, *Math. Z.* **115** (1970), 33—52. *MR* **41** #7274.

[9] VEIDINGER, L.: On the order of convergence of finite element methods for the Neumann problem, *Studia Sci. Math. Hungar.* **13** (1978), 411—422.

[10] VEIDINGER, L., On the order of convergence of a finite element scheme, *Acta Math. Acad. Sci. Hungar.* **25** (1974), 401—412. *MR* **50** #9012.

[11] FRIEDRICHS, K., Die Randwert- und Eigenwertprobleme aus der Theorie der elastischen Platten, Anwendung der direkten Methoden der Variationsrechnung, *Math. Ann.* **98** (1928), 205—247.

[12] AGMON, S., *Lectures on elliptic boundary value problems,* Van Nostrand Mathematical Studies, No. 2, D. Van Nostrand Co., Inc., Princeton, N. J.—Toronto, Ont.—London, 1965. *MR* **31** #2504.

[13] IL'IN, V. P., Some inequalities in function spaces and their application to the investigation of the convergence of variational processes, *Trudy Mat. Inst. Steklov* **53** (1959), 64—127 (in Russian). *MR* **22** #9738.

[14] KOUKAL, S., Piecewise polynomial interpolations in the finite element method, *Apl. Mat.* **18** (1973), 146—160. *MR* **47** #9851.

(Received March 6, 1980)

MTA MATEMATIKAI KUTATÓ INTÉZETE
REÁLTANODA U. 13—15.
H—1053 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 375—380

HOMOLOGY AND V-CENTRAL SERIES, I

by

J. FURTADO-COELHO

1. Introduction

The variety of all groups, the variety of abelian groups and the varieties of abelian groups of exponent p (p any prime) are such that any subgroup of any (relatively) free group in one of these varieties is itself (relatively) free. Moreover these are the only subvarieties of the variety of all groups for which this holds. For other subvarieties there are some theorems by P. HALL [8], MOSTOWSKY [12], STALLINGS [15], STAMMBACH [16], ŠMELKIN [14] and others which give sufficient conditions for a subgroup of a given group to be (relatively) free. Under inspection these theorems exhibit a very similar structure. On the other hand KNUS [10] gave analogous theorems for different algebraic systems, namely, rings, algebras and Lie algebras. Again, some theorems by Stallings and Knus, used by them and also by Stammbach in order to derive the aforementioned theorems, do present a very similar structure. Following the line indicated in [7] we shall show in the present paper how it is possible to unify and generalize the results already known. The generalization is twofold: first, most of the authors who dealt with the subject only envisaged the case of ordinary groups, whereas our results hold for any type of group with multiple operators; and secondly, our results hold for any type of variety subfunctor [3], [4], whilst most of the previous authors only considered the ordinary commutator [—, —] of ordinary groups, or subfunctors closely related to it. Moreover, as often happens, by putting the problem in more general terms one is able to concentrate on its essential features and leave out the episodic details, thereby obtaining simpler, more transparent proofs.

Looking at the main theorem (Theorem 4) one clearly sees that we are strongly indebted to Stallings [15]. This theorem and its corollaries are given in Section 2. In Section 3 we give some instances of the wide applicability of the general results.

2. Main result

From now on the term "group" with no further specification will mean Ω-group (i.e., group with multiple operators) for a fixed, but otherwise arbitrary, choice of Ω. We assume the reader is familiar with the concept and fundamental properties of a group with multiple operators [9]. We assume, as well, familiarity with the notions described on pages 252—257 of [2].

AMS (MOS) subject classifications (1979). Primary 70J99; Secondary 20F19.

Key words and phrases. Group with multiple operators (group, ring, module, Lie algebra etc.), variety, variety subfunctor, lower central series, nilpotency, polynilpotency, homology group, free presentation of a group.

Let $\mathscr{C}$ be a variety of groups and let $\mathscr{V}$ be a subvariety of $\mathscr{C}$. Let V be the subfunctor of the identity functor in $\mathscr{C}$ associated with $\mathscr{V}$ and let V_1 be as described in [4] or [6]. Given a group A we put [7]

$$V^0(A) = A;\ V^1(A) = V(A);\ V^{n+1}(A) = V_1(V^n(A)|A), \quad n \geqq 1.$$

These ideals of A form a chain $A \supset V(A) \supset V^2(A) \supset \dots$ which we call the *lower V-central series* of A. It is convenient to write A_n for $V^n(A)$, but we must bear in mind that one is referring to a particular variety subfunctor. We shall also write $A_\infty = \bigcap\limits_{n \geqq 0} A_n$. A group A is said to be *V-nilpotent* if there is some n for which $A_n = 0$. It is easy to see that

LEMMA 1. *A group $A \in \mathscr{C}$ is residually V-nilpotent if and only if the intersection of all terms of its lower V-central series is trivial (i.e., iff $A_\infty = 0$).*

We recall that, given an exact sequence

$$0 \to C \to B \to A \to 0 \tag{1}$$

of groups in $\mathscr{C}$, then the sequence (same notation as in [4])

$$D_1U(B) \to D_1U(A) \to C/V_1(C|B) \to U(B) \to U(A) \to 0$$

is exact and functorial with regard to the category of exact sequences of type (1). We shall apply this to the exact sequence $0 \to A_n \to A \to A/A_n \to 0$ to obtain the exact sequence

$$D_1U(A) \to D_1U(A/A_n) \to A_n/A_{n+1} \to 0,$$

since, by definition, $V_1(A_n|A) = A_{n+1}$ and $U(A/A_n) \cong A/A_1 = A/V(A) = U(A)$.

PROPOSITION 1. *If $D_1U(A) = 0$, then $D_1U(A/A_n) \cong A_n/A_{n+1}$ for every n.*

Let B be another group and $f\colon A \to B$ a homomorphism. One has, then, a commutative diagram with exact rows

$$\begin{tikzcd}
D_1U(A) \arrow[r] \arrow[d] & D_1U(A/A_n) \arrow[r] \arrow[d] & A_n/A_{n+1} \arrow[r] \arrow[d] & 0 \\
D_1U(B) \arrow[r] & D_1U(B/B_n) \arrow[r] & B_n/B_{n+1} \arrow[r] & 0
\end{tikzcd}$$

PROPOSITION 2. *Let $f\colon A \to B$ be a homomorphism which induces a surjection $D_1U(A) \to D_1U(B)$. Then*

(i) *if f induces an isomorphism $A/A_n \cong B/B_n$, then f will induce isomorphisms $A_n/A_{n+1} \cong B_n/B_{n+1}$ and $A/A_{n+1} \cong B/B_{n+1}$;*

(ii) *if f induces an isomorphism $f_n\colon A/A_n \xrightarrow{\sim} B/B_n$ for every $n > 0$, then f will induce an injection $f_\infty\colon A/A_\infty \to B/B_\infty$;*

(iii) *let f be as in* (ii) *and, in addition, surjective; then f_∞ will be an isomorphism.*

As regards (ii) we remark that there are cases in which f_∞ is definitely not surjective [15].

Using this proposition and induction on n one can prove

THEOREM 1. *Let $f\colon A\to B$ be a homomorphism which induces an isomorphism $A/A_1\cong B/B_1$ and a surjection $D_1U(A)\to D_1U(B)$. Then f will induce an isomorphism $f\colon A/A_n\cong B/B_n$ for every n; it will also induce an injection $f_\infty\colon A/A_\infty\to B/B_\infty$; if, in addition, f is surjective, then f_∞ will be an isomorphism.*

The proof is parallel to the proof in [15], except for the fact that, as we have shown, the five term exact sequence actually reduces to a three term one. The same applies to Proposition 2.

COROLLARY 1. *Let $f\colon A\to B$ induce an isomorphism $A/A_1\cong B/B_1$ and a surjection $D_1U(A)\to D_1U(B)$.*

(i) *If $A\neq 0$ is V-nilpotent, i.e., if, say, $A_{n-1}\neq 0$, $A_n=0$, for some n, then $A\cong B/B_n$, $B_n=B_{n+1}=\dots$, f is injective and $D_1U(A)\cong D_1U(B)$.*

(ii) *If $B_{n-1}\neq 0, B_n=0$ for some n, then $A/A_n\cong B$, $A_n=A_{n+1}=\dots$ and f is surjective.*

(iii) *If A and B are both V-nilpotent, then their nilpotency class is the same and f is an isomorphism.*

COROLLARY 2. *Consider two subvarieties $\mathscr{V}$ and $\mathscr{V}'$ of $\mathscr{C}$ and let $A, B\in\mathscr{V}'$. Assume A and B are both V-nilpotent. If $D_1U(B)=0$, then a homomorphism $f\colon A\to B$ will be an isomorphism if and only if so is the induced $f_1\colon A/A_1\to B/B_1$. In particular this is so if B is $\mathscr{V}'$-free.*

PROOF. The result follows from Corollary 1. Notice that, since $A, B\in\mathscr{V}'$, both have $\mathscr{V}'$-free presentations and if B is $\mathscr{V}'$-free, then $(O|B)$ will be one such for B. In terms of this presentation $D_1U(B)=0$ [4], [6].

COROLLARY 3. *Let $\mathscr{V}$ and $\mathscr{V}'$ be two subvarieties of $\mathscr{C}$ and assume the free groups of $\mathscr{V}'$ are residually V-nilpotent. Assume moreover that $A\in\mathscr{V}'$ is such that $D_1U(A)=0$, that A/A_1 is $(\mathscr{V}\cap\mathscr{V}')$-free and that a transitive dependence relation holds in A/A_1. If $\{a_i\}$ is a set of elements of A whose images in A/A_1 (under the quotient map) form an independent set, then $\{a_i\}$ freely generates a subgroup of A.*

PROOF. Let us consider the set $\{\bar a_i\}$ of the images in A/A_1 of the a_i's. This set can be completed [2] to a basis of A/A_1, which we may denote by $\{\bar a_j\}$. Let $X=\{a_j\}$ be a set consisting of the a_i's plus counterimages of the elements of $\{\bar a_j\}-\{\bar a_i\}$, one for each of them. Let F be the $\mathscr{V}'$-free group on the set X: we shall define a homomorphism $f\colon F\to A$ by the correspondence $f\colon a_j\to a_j\cdot F/F_1$ is $(\mathscr{V}\cap\mathscr{V}')$-free on $\{\bar a_j\}$, therefore f induces an isomorphism $f_1\colon F/F_1\xrightarrow{\sim} A/A_1$. Consequently, f induces an injection $f\colon F\to A/A_\infty$ and so $f\colon F\to A$ has to be injective.

COROLLARY 4. *Let $\mathscr{V}$ and $\mathscr{V}'$ be two subvarieties of $\mathscr{C}$ and let $A\in\mathscr{V}'$ be such that $D_1U(A)=0$. If A/A_1 is $(\mathscr{V}\cap\mathscr{V}')$-free, then there exists a $\mathscr{V}'$-free group F and a homomorphism $f\colon F\to A$ which induces an isomorphism $f_n\colon F/F_n\xrightarrow{\sim} A/A_n$ for every $n>0$.*

COROLLARY 5. *Let $\mathscr{V}$ and $\mathscr{V}'$ be two subvarieties of $\mathscr{C}$ and suppose that $\mathscr{V}'$ is a V-nilpotent variety. Let $A\in\mathscr{V}'$ be such that $D_1U(A)=0$. If A/A_1 is $(\mathscr{V}\cap\mathscr{V}')$-free, then A is $\mathscr{V}'$-free.*

3. Applications

Let G be an ordinary group, N a normal subgroup and p a prime or 0. We assume that a choice of p has been made and stick to it. We denote by $N \#_p G = N \# G$ the subgroup of G generated by all elements of the form $g^{-1}n^{-1}gnm^p$, with $g \in G$, $n, m \in N$. If in Theorem 1 we put $V(G) = G \# G$ we shall obtain Stalling's theorem [15]. In fact it is not difficult to see [6] that $V_1(N|G) = N \# G$, hence $V^n(G) = G_n = G_{n-1} \# G$; also $H_0(G; Z_p) = Z_p$, $H_1(G; Z_p) = (G/G') \otimes_Z Z_p = G/(G \# G) = G/G_1$, $H_2(G; Z_p) = D_1U(G)$. Stalling's theorem extends some results due to BAER.

The following theorem will provide an illustration for Corollary 2. The concept of polynilpotency for ordinary groups is well-known: given a finite list $(n_1, n_2, \ldots, n_k)$ of integers equal to or greater than 1, one defines polynilpotency of type $(n_1, n_2, \ldots, n_k)$. Let G be an ordinary group: we shall denote by $G_{n_1, n_2, \ldots, n_j}$, the subgroup of G obtained after performing the commutator operations corresponding to the integers $(n_1, n_2, \ldots, n_j)$. Thus $G_{n_1, n_2 \ldots, n_j, 1}$ will denote $[G_{n_1, \ldots, n_j}, G_{n_1, \ldots, n_j}]$.

THEOREM 2. *Let $\mathscr{C}$ be the variety of ordinary groups and let us assume that $G \in \mathscr{C}$ is free polynilpotent of class $(n_1, n_2, \ldots, n_k)$, i.e., G is the quotient of a free group F by $F_{n_1, \ldots, n_k}$. Then an endomorphism $f: G \to G$ will be an automorphism if and only if it induces surjections $D_1U(G_{n_1, \ldots, n_j}) \to D_1U(G_{n_1, \ldots, n_j})$ and automorphisms in $G_{n_1, \ldots, n_j}/G_{n_1, \ldots, n_j, 1}$, where $0 \leqq j \leqq k-1$, and one assumes that, for $j=0$, $G_{n_1, \ldots, n_j} = G$.*

PROOF. It is known [4], [6] that, for $V(G) = [G, G]$, one has $V_1(N|G) = [N, G]$, so that $V^2(G) = [[G, G], G]$, etc. We consider successively the exact sequences $0 \to G_i \to G \to G/G_i \to 0$, $1 \leqq i \leqq n_1 - 1$. If we assume that f induces an automorphism in G/G_1 and a surjection $D_1U(G) \to D_1U(G)$, we end up with an automorphism of G/G_{n_1}; one then goes over to the exact sequences $0 \to G_{n_1, 1} \to G_{n_1} \to G_{n_1}/G_{n_1, 1} \to 0$, $1 \leqq i \leqq n_2 - 1$, etc. If, in particular, the list $(n_1, \ldots, n_k)$ consists of a single integer, then one has a theorem due to Mostowsky [11].

We illustrate Corollary 3 with two theorems.

THEOREM 3 [15]. *Let G be an ordinary group and p a prime number such that $H_2(G; Z_p) = 0$. Let $\{x_i\}$ be a set of elements of G whose images in $H_1(G; Z_p)$ are linearly independent over Z_p. Then $\{x_i\}$ is a basis of a free subgroup of G.*

This result generalizes a theorem of Hall [8] and Mostowsky [12].

THEOREM 4 [14]. *Let G be an ordinary group free polynilpotent of a given type. A subgroup H of G is itself a free polynilpotent group (of the same type) if and only if it possesses a system of generators $\{x_i\}$ linearly independent modulo the commutator group G'.*

The following result, to be found in [16], illustrates Corollary 4. Let $\mathscr{C}$ be the variety of ordinary groups, let $V(G) = G \#_p G$, for a fixed p either prime or zero, and let $\mathscr{V}$ be the corresponding variety. Given a variety $\mathscr{V}' \subset \mathscr{C}$ and $G \in \mathscr{V}'$, we consider $D_1U(G)$ *evaluated in terms of a $\mathscr{V}'$-free presentation* [6].

THEOREM 5. *Let G be a group in a variety $\mathscr{V}' \subset \mathscr{C}$ such that $D_1U(G) = 0$. In the case $p = 0$ we shall require $U(G) = G/[G, G]$ to be $(\mathscr{V}' \cap \mathfrak{Ab})$-free, where*

$\mathfrak{Ab}$ *denotes the variety of abelian groups. Then there will be a* $\mathcal{V}'$*-free group* F *and a homomorphism* $f: F \to G$ *which induces an isomorphism* $f_n: F/F_n \to G/G_n$ *for every* n*, where* F_n *denotes the* n*-th term of the lower central series.*

As regards Corollary 5 one has

THEOREM 6 [16]. *Let the general situation be as in Theorem 5. Let* $\mathcal{V}'$ *be a nilpotent variety of exponent* 0. *If* G *is a finitely generated group in* $\mathcal{V}'$ *with* $D_1U(G)=0$ *for all primes, then* G *is* $\mathcal{V}'$*-free.*

Finally we consider a somewhat different case. Let Λ be a commutative ring with identity and let $\mathscr{C}$ be the variety of Λ-modules. Let J be an ideal of Λ and let us write $V(M)=JM$. This defines a variety, namely, that of the modules which are annihilated by J. It is easy to see that $V_1(N|M)=JN$. Also [4] $M/M_1=M/JM=$ $=(\Lambda/J)\otimes_\Lambda M$, $D_1U(M)=\mathrm{Tor}_1^\Lambda(\Lambda/J, M)$. One can then apply Theorem 1. In particular one has

THEOREM 7. *Let* $\mathscr{C}$ *be the variety of* Λ*-modules and let* J *be a nilpotent ideal of* Λ. *Then:* (i) *Let* $M\in\mathscr{C}$ *be such that* $\mathrm{Tor}_1^\Lambda(\Lambda/J, M)=0$; *then a* Λ*-endomorphism* $f: M\to M$ *will be an automorphism if and only if so is the induced* f_1 *on* $(\Lambda/J)\otimes_\Lambda M$. (ii) *Let* M/JM *be a free* Λ/J*-module; then* $\mathrm{Tor}_1^\Lambda(\Lambda/J, M)=0$ *if and only if* M *is a free* Λ*-module; in particular let* J *be a maximal nilpotent ideal of* Λ; *then* $\mathrm{Tor}_1^\Lambda(\Lambda/J, M)=0$ *if and only if* M *is a free* Λ*-module.*

4. Acknowledgements

Most of the contents of this paper is part of my Ph. D. Thesis [5]. I worked for the thesis under the supervision of Dr. ABRAHAM S.—T. LUE and it is a pleasure to take the opportunity to thank him for all the help and guidance he gave me both before and after my getting the degree. Both the Instituto Superior Técnico and the Instituto de Alta Cultura (Lisbon) gave me generous financial support during the preparation of the thesis. The Instituto Nacional de Investigação Científica and the Fundação Calouste Gulbenkian gave me their generous financial support for the preparation of this paper. Finally, I would like to acknowledge the fact that, some time after this paper had been written, I learned that Prof. A. R.-GRANDJEAN and DR. L. FRANCO FERNÁNDEZ, drawing inspiration from several sources but mainly from my two previous papers [6] and [7], had arrived at practically the same general results that we present in Section 2.

REFERENCES

[1] BAER, R., Representations of groups as quotient groups, I, II. Minimal central chains of a group, III. Invariants of classes of related representations, *Trans. Amer. Math. Soc.* **58** (1945), 295—347, 348—389, 390—419. *MR* 7—371, 372.

[2] COHN, P. M., *Universal algebra*, Harper & Row, Publishers, New York—London, 1965. *MR* **31** #224; erratum **32**, p. 1754.

[3] FRÖHLICH, A., On groups over a d. g. near-ring, II. Categories and functors, *Quart. J. Math. Oxford Ser.* (2) **11** (1960), 211—228. *MR* **22** #11023.

[4] FRÖHLICH, A., Baer-invariants of algebras, *Trans. Amer. Math. Soc.* **109** (1963), 221—244. *MR* **28** #2142.

[5] FURTADO-COELHO, J., Varieties of Ω-groups and associated functors, Ph. D. thesis, University of London, 1972.

[6] FURTADO-COELHO, J., Homology and generalized Baer invariants, *J. Algebra* **40** (1976), 596—609. *MR* **54** #2835.

[7] FURTADO-COELHO, J., V-center and some related concepts, *Estudos de Geometria, Algebra e Análise,* Publicações do II Centenário da Academia das Ciências de Lisboa, Lisbon, 1978, 161—184.

[8] HALL, P., The splitting properties of relatively free groups, *Proc. London Math., Soc.* (3) **4** (1954), 343—356. *MR* **16**—217.

[9] HIGGINS, P. J., Groups with multiple operators, *Proc. London Math. Soc.* (3) **6** (1956), 366—416. *MR* **18**—559.

[10] KNUS, M.-A., Homology and homomorphisms of rings, *J. Algebra* **9** (1968), 274—284. *MR* **37** #2837.

[11] MOSTOWSKI, A. W., Nilpotent free groups, *Fund. Math.* **49** (1960/61), 259—269. *MR* **26** #3781.

[12] MOSTOWSKI, A. W., On automorphisms of relatively free groups, *Fund. Math.* **50** (1961/62), 403—411. *MR* **25** #1204.

[13] NEUMANN, H., *Varieties of groups,* Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 37, Springer-Verlag, New York, Inc., New York, 1967. *MR* **35** #6734.

[14] ŠMEL'KIN, A. L., Free polynilpotent groups, *Dokl. Akad. Nauk SSSR* **151** (1963), 73—75 (in Russian). *MR* **27** #5820. Translation in *Soviet Math. Dokl.* **4** (1963), 950—953.

[15] STALLINGS, J. R., Homology and central series of groups, *J. Algebra* **2** (1965), 170—181. *MR* **31** #232.

[16] STAMMBACH, U., Homological methods in group varieties, *Comment. Math. Helv.* **45** (1970), 287—298. *MR* **43** #324.

[17] TATE, J., Nilpotent quotient groups, *Topology* **3** Suppl. 1 (1964), 109—112.

(Received March 17, 1980)

FACULDADE DE CIÊNCIAS
MATEMÁTICA PURA
134, AV. 24 DE JULHO—3°
P—1300 LISBOA
PORTUGAL

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 381—408

ON APPROXIMATION OF THE SPECTRAL DENSITY FUNCTION OF A SELFADJOINT OPERATOR

by

KRZYSZTOF MOSZYNSKI

Introduction

The aim of this paper is to define and to study a certain method of approximation of so called spectral density function of a bounded selfadjoint operator A in a Hilbert space X.

For the spectral family $E(\lambda)$ of A we define by $\phi_x(\lambda)=\|E(\lambda)x\|^2$ the spectral density function, which is clearly real, non-negative and increasing function of real variable λ.

The family $\{\phi_x\}_{x\in X}$ contains the whole information about the spectrum $\sigma(A)$ of A. Any particular function ϕ_x yields some information about a certain part of $\sigma(A)$. By studying the intervals of constancy, the points of increase, the points of continuity, and the jumps of the functions ϕ_x, it is relatively easy to situate the spectrum $\sigma(A)$ and even to classify its points. Thus it is to be expected that appropriate approximations of functions ϕ_x should yield the corresponding approximate information about the spectrum of A.

In Section 1 we give preliminary definitions and recall several known facts.

In Section 2 we define approximating sequence ϕ_N for spectral density function, and state two theorems on convergence. This is entirely based on classical material, mainly on one of Helly's theorems.

Section 3 discusses approximation of the spectral family $E(\lambda)$ of A. It is shown that at any point of continuity, $E(\lambda)$ can be strongly approximated by a certain sequence of operators which may be effectively constructed. The result can be found in [3], but here it is obtained by a different method, as a byproduct of Section 2, in modifying slightly some theorems from [10].

Section 4 is devoted to the study of the nature of convergence of the sequence $\{\phi_N\}$. It contains theorems on rate of convergence and on the elimination of the so called "rubbish point effect" in the approximation (Theorems 4.1 and 4.2). We discuss also in a special case the behaviour of the sequence $\{\phi_N\}$ at the points of discontinuity of the approximated function.

Section 5 contains another algorithm for constructing the sequence ϕ_N. This algorithm is more effective than that contained in Section 2. It is shown that ϕ_N is completely determined by an eigenproblem for a certain tridiagonal, symmetric, real matrix. However, this algorithm does not seem to be yet in the computer stage,

---

AMS (MOS) subject classifications (1979). Primary 47B25; Secondary 33A65, 41A10, 65D15, 65D32, 65F15.

Key words and phrases. Spectral family, spectral density function, spectral approximation, spectrum, zeros of orthogonal polynomials, selfadjoint operator.

mainly due to the disadvantages of the Lanczos orthogonalization process applied here. The author hopes to be able to eliminate this inconvenience in the future.

All the theorems contained in this paper are proved by elementary methods, based on properties of polynomials.

§ 1. Preliminaries

Consider a linear, bounded, selfadjoint operator A mapping a Hilbert space X into itself:

$$A\colon X \to X.$$

Then

$$A = \int_a^b \lambda \, dE(\lambda) \tag{1.1}$$

where

$$a = \inf_{\|x\|=1} (Ax, x)$$

$$b = \sup_{\|x\|=1} (Ax, x)$$

and $E\colon [a, b] \to \mathscr{L}(X)$, is the so called spectral family of A. For any $\lambda \in [a, b]$, $E(\lambda)\colon X \to X$ is an orthogonal projection such that $E(\lambda)E(\mu) = E(\mu)E(\lambda)$ for any $\lambda, \mu \in [a, b]$, $E(a) = 0$ and $E(b) = I$ — the identity operator.

Moreover, for any $\lambda \in (a, b]$ we have $E(\lambda) = E(\lambda + 0)$ in the strong sense, and

$$\lambda \leqq \mu \quad \text{implies} \quad E(\lambda) \leqq E(\mu)$$

i.e., E is monotone. The last condition is equivalent to the following:

$$\lambda \leqq \mu \quad \text{implies} \quad E(\lambda)E(\mu) = E(\lambda).$$

We shall use the notation $E([c, d]) = E(d) - E(c)$ for $[c, d] \subset [a, b]$.

Integral (1.1) converges in the uniform operator topology [7].

A point $\lambda \in [a, b]$ is a point of constancy iff there is an $\varepsilon > 0$ such that, $(\lambda - \varepsilon, \lambda + \varepsilon) \subset [a, b]$ and E is constant on $(\lambda - \varepsilon, \lambda + \varepsilon)$.

Any point of $[a, b]$ which is not a point of constancy of E is called a point of increase. The spectrum $\sigma(A)$ is the set of all points of increase of E. The points of discontinuity of E are eigenvalues of A.

Let $x \in X, x \neq 0$; define the spectral density function

$$\phi_x(\lambda) = \big(E(\lambda)x, x\big) = \|E(\lambda)x\|^2 \geqq 0, \quad \lambda \in [a, b]. \tag{1.2}$$

It is easy to show that

$$\begin{aligned} &\phi_x(a) = 0, \\ &\phi_x(\lambda) \leqq \phi_x(\mu) \quad \text{if} \quad \lambda \leqq \mu, \quad \lambda, \mu \in [a, b], \\ &\phi_x(\lambda) = \phi_x(\lambda + 0) \quad \text{if} \quad \lambda \in (a, b], \\ &\phi_x(b) = \|x\|^2. \end{aligned} \tag{1.3}$$

Any point of increase of ϕ_x is a point of increase of E, i.e. it belongs to $\sigma(A)$. Any point of discontinuity of ϕ_x is an eigenvalue of A. Hence the set of all the points of increase of ϕ_x is a subset of $\sigma(A)$ and the set of all the points of discontinuity of ϕ_x is a subset of the set of all eigenvalues of A.

We shall write ϕ instead of ϕ_x, whenever it does not lead to a confusion.

Let us introduce now another function $\phi_{x,y}\colon [a, b]\to \mathbf{C}$

$$\phi_{xy}(\lambda) = (E(\lambda)x, y) \tag{1.4}$$

where $x, y\in X$, $x\neq 0$, $y\neq 0$.

It is easy to verify that

$$\phi_{xy}(\lambda) = \frac{1}{4}\{\phi_{x+y}(\lambda)-\phi_{x-y}(\lambda)+i[\phi_{x+iy}(\lambda)-\phi_{x-iy}(\lambda)]\} \tag{1.5}$$

hence ϕ_{xy} is completely determined by the functions of the form ϕ_z, for $z\in X$.

Notice that

$$\phi_{\sum_{j=1}^{n}\alpha_j x_j}(\lambda) = \left(E(\lambda)\sum_{j=1}^{n}\alpha_j x_j,\ \sum_{k=1}^{n}\alpha_k x_k\right) = \sum_{j=1}^{n}\sum_{k=1}^{n}\alpha_j\bar{\alpha}_k\phi_{x_jx_k}(\lambda) = \alpha^* B_{x_1x_2\dots x_n}(\lambda)\alpha \tag{1.6}$$

where $\alpha=[\alpha_1, \dots, \alpha_n]^T$ and $B_{x_1\dots x_n}(\lambda)=(\phi_{x_ix_j}(\lambda))_{i,j=1,2,\dots,n}$ is the $n\times n$ Hermitian matrix.

Moreover,

$$\phi_x(\lambda)-\phi_y(\lambda) = (E(\lambda)x, x-y)+(E(\lambda)(x-y), y).$$

Since $\|E(\lambda)\|=1$ we have the following estimate

$$|\phi_x(\lambda)-\phi_y(\lambda)| \leqq (\|x\|+\|y\|)\|x-y\|. \tag{1.7}$$

Finally, if $f\colon [a, b]\to \mathbf{C}$ is sufficiently regular, then

$$f(A) = \int_a^b f(\lambda)\, dE(\lambda) \tag{1.8}$$

and

$$(f(A)x, y) = \int_a^b f(\lambda)\, d\phi_{xy}(\lambda), \qquad x, y\in X. \tag{1.9}$$

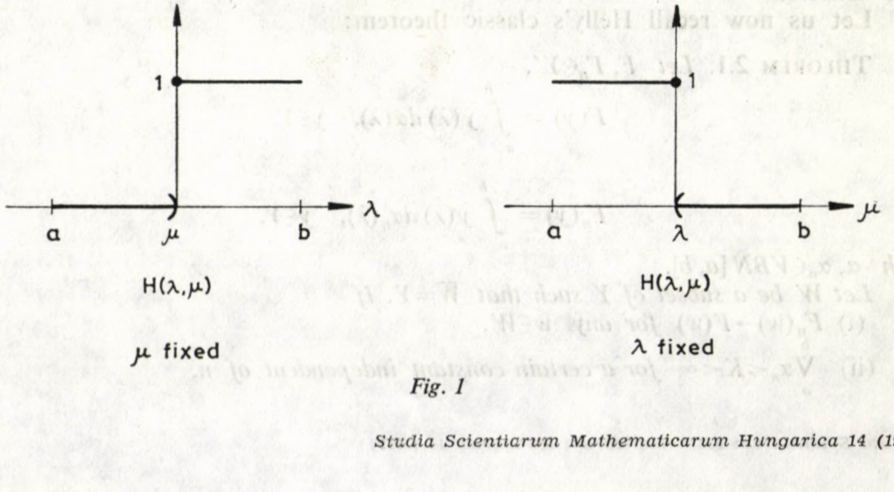


Fig. 1

For any $\lambda, \mu \in [a, b]$ define now the function

$$H(\lambda, \mu) = \begin{cases} 0 & a \leqq \lambda < \mu \\ 1 & \mu \leqq \lambda \leqq b. \end{cases}$$

It is easy to verify that

$$E(\lambda+0) = \int_a^b H(\lambda, \mu)\, dE(\mu) = H(\lambda, A)$$

$(E(\lambda+0)=E(\lambda)$ in $(a, b))$ (see [7]).

§ 2. Approximation of the spectral density function

Consider now the space $Y=C[a, b]$ normed by $\|y\| = \sup_{t\in[a,b]} |y(t)|$ for $y \in Y$. According to the Riesz Theorem any linear bounded functional F belonging to the dual space Y' is of the form

$$F(y) = \int_a^b y(\lambda)\, d\alpha(\lambda) \tag{2.1}$$

where α is a certain function with bounded variation: $\overset{b}{\underset{a}{\mathrm{V}}}\alpha < \infty$.

Correspondence $F \to \alpha$ given by (2.1) is not a function, but we can always choose α such that

$$\alpha(a) = 0$$

and

$$\alpha(\lambda) = \alpha(\lambda+0), \quad \lambda \in (a, b).$$

The space $VBN[a, b]$ of all the functions with bounded variation satisfying these conditions is a B-space normed by $\|\alpha\| = \overset{b}{\underset{a}{\mathrm{V}}}\alpha$.

It is well-known that the correspondence

$$F \mapsto \alpha \in VBN[a, b]$$

is a function.

Let us now recall Helly's classic theorem:

THEOREM 2.1. *Let* $F, F_n \in Y'$,

$$F(y) = \int_a^b y(\lambda)\, d\alpha(\lambda), \quad y \in Y,$$

$$F_n(y) = \int_a^b y(\lambda)\, d\alpha_n(\lambda), \quad y \in Y,$$

with $\alpha, \alpha_n \in VBN[a, b]$.

Let W be a subset of Y such that $\overline{W}=Y$. If

(i) $F_n(w) \to F(w)$ *for any* $w \in W$,

(ii) $\overset{b}{\underset{a}{\mathrm{V}}}\alpha_n < K < \infty$ *for a certain constant independent of* n, *then for any* $y \in Y$

$$F_n(y) \to F(y) \quad when \quad n \to \infty.$$

Deduce now several conclusions. Taking $y(t) \equiv 1$ in Theorem 2.1, we get

(2.2) $$\alpha_n(b) \to \alpha(b) \quad \text{when} \quad n \to \infty.$$

If for any fixed $\lambda \in [a, b]$

$$y(t) = \begin{cases} \lambda - t & \text{for} \quad t \in [a, \lambda] \\ 0 & \text{for} \quad t \in (\lambda, b] \end{cases}$$

then for any $\lambda \in [a, b]$

$$\int_a^\lambda \alpha_n(t)\,dt \to \int_a^\lambda \alpha(t)\,dt, \quad \text{when} \quad n \to \infty.$$

If α_n are increasing (i.e. $t \leqq t'$ implies $\alpha_n(t) \leqq \alpha_n(t')$), then for sufficiently small $h > 0$

$$\frac{1}{h}\int_{\lambda-h}^{\lambda} \alpha_n(\mu)\,d\mu \leqq \alpha_n(\lambda) \leqq \frac{1}{h}\int_{\lambda}^{\lambda+h} \alpha_n(\mu)\,d\mu.$$

Hence

$$\frac{1}{h}\int_{\lambda-h}^{\lambda} \alpha(\mu)\,d\mu \leqq \underline{\lim}\,\alpha_n(\lambda) \leqq \overline{\lim}\,\alpha_n(\lambda) \leqq \frac{1}{h}\int_{\lambda}^{\lambda+h} \alpha(\mu)\,d\mu.$$

Assume now that λ is a point of continuity of α. If $h \to 0$ then

$$\alpha(\lambda) \leqq \underline{\lim}\,\alpha_n(\lambda) \leqq \overline{\lim}\,\alpha_n(\lambda) \leqq \alpha(\lambda)$$

and thus

(2.3) $$\lim \alpha_n(\lambda) = \alpha(\lambda) \quad \text{(see [7])}.$$

If we assume that not only α_n but also α are increasing we obtain similarly

$$\frac{1}{h}\int_{\lambda-h}^{\lambda} \alpha(\mu)\,d\mu \leqq \alpha(\lambda) \leqq \frac{1}{h}\int_{\lambda}^{\lambda+h} \alpha(\mu)\,d\mu.$$

Subtracting the inequalities for α_n and α we get

$$\frac{1}{h}\Big[\int_{\lambda-h}^{\lambda} \alpha(\mu)\,d\mu - \int_{\lambda}^{\lambda+h} \alpha_n(\mu)\,d\mu\Big] \leqq \alpha(\lambda) - \alpha_n(\lambda) \leqq \frac{1}{h}\Big[\int_{\lambda}^{\lambda+h} \alpha(\mu)\,d\mu - \int_{\lambda-h}^{\lambda} \alpha_n(\mu)\,d\mu\Big].$$

Now if $n \to \infty$ then

$$\frac{1}{h}\Big[\int_{\lambda-h}^{\lambda} \alpha(\mu)\,d\mu - \int_{\lambda}^{\lambda+h} \alpha(\mu)\,d\mu\Big] \leqq \underline{\lim}\,(\alpha(\lambda) - \alpha_n(\lambda)) \leqq \overline{\lim}\,(\alpha(\lambda) - \alpha_n(\lambda)) \leqq$$

$$\leqq \frac{1}{h}\Big[\int_{\lambda}^{\lambda+h} \alpha(\mu)\,d\mu - \int_{\lambda-h}^{\lambda} d(\mu)\,d\mu\Big].$$

But, for $\lambda \in (a, b)$

$$\frac{1}{h}\int_{\lambda-h}^{\lambda} \alpha(\mu)\,d\mu \to \alpha(\lambda - 0)$$

and

$$\frac{1}{h}\int_{\lambda}^{\lambda+h} \alpha(\mu)\, d\mu \to \alpha(\lambda+0) = \alpha(\lambda)$$

when $h\to 0$. Hence in this case the following convergence estimate holds for any $\lambda\in(a, b)$:

$$\alpha(\lambda-0)-\alpha(\lambda) \leqq \underline{\lim}\, (\alpha(\lambda)-\alpha_n(\lambda)) \leqq \overline{\lim}\, (\alpha(\lambda)-\alpha_n(\lambda)) \leqq \alpha(\lambda)-\alpha(\lambda-0). \tag{2.4}$$

Let us give another interpretation of the convergence in Theorem 2.1. Put

$$y(t) = \int_0^t \varrho(s)\, ds$$

where ϱ is any integrable function. Integrating by parts we get

$$F(y) = \int_a^b \varrho(t)\alpha(t)\, dt$$

and

$$F_n(y) = \int_a^b \varrho(t)\alpha_n(t)\, dt.$$

Hence Theorem 2.1 implies

$$\int_a^b \varrho(t)(\alpha(t)-\alpha_n(t))\, dt \to 0 \quad \text{when} \quad h\to 0 \tag{2.5}$$

for any integrable ϱ. This is in some sense weighted mean convergence. The last formula can be also understood as $L_2[a, b]$ weak convergence of α_n to α.

Another general theorem can be helpful in studying the nature of convergence of the sequence $\{\alpha_n\}$. We quote here this theorem without proof.

THEOREM 2.2 (see [2] for slightly more general formulation). *Let $\{f_n\}$ be a sequence of real increasing functions converging to a continuous function f on a closed finite interval* $[\alpha, \beta]$. *Then f_n converges uniformly to f on* $[\alpha, \beta]$.

The Lanczos orthogonalization process.

Assume now that the function $\phi(\lambda)=(E(\lambda)x, x)$ defined in § 1 has infinite number of points of increase in $[a, b]$.

Put $x_{-1}=0$, $x_0=\dfrac{x}{\|x\|}\in X$, and define the orthonormal sequence $\{x_k\}$ of points of the Hilbert space X, by the formula

$$Ax_k = a_{k\,k-1}x_{k-1}+a_{kk}x_k+a_{k\,k+1}x_{k+1} \tag{2.6}$$

where

$$\begin{aligned} a_{kj} &= (Ax_k, x_j) \quad \text{for} \quad j=k-1, k \\ a_{k\,k+1} &= (\|Ax_k\|^2-a_{k\,k-1}^2-a_{kk}^2)^{1/2} > 0. \end{aligned} \tag{2.7}$$

The orthogonalization process (2.6), (2.7) is called the Lanczos' process.
In the sequel we shall use the notation

$$(f, g)_\phi = \int_a^b f(\lambda)\bar{g}(\lambda)\, d\phi(\lambda)$$

and $\|\cdot\|_\phi$ for the norm corresponding to the scalar product $(\, , \,)_\phi$.

THEOREM 2.3. *Let $P_0, P_1, \ldots, P_k, \ldots$ be the sequence of orthogonal polynomials with respect to the scalar product $(\, , \,)_\phi$ (j being the degree of P_j). Then the orthogonal sequence $\{x_k\}$ given by* (2.6) *and* (2.7) *is infinite and*

$$x_k = P_k(A)x. \tag{2.8}$$

Moreover, for any $k=0, 1, 2, \ldots$ the following three term recurrence formula holds

$$\lambda P_k(\lambda) = a_{k\,k-1} P_{k-1}(\lambda) + a_{kk} P_k(\lambda) + a_{k\,k+1} P_{k+1}(\lambda). \tag{2.9}$$

PROOF (Induction with respect to k). Put $P_{-1}=0$, $P_0=\dfrac{1}{\|x\|}$ and define the sequence of orthonormal polynomials with respect to the scalar product $(\, , \,)_\phi$. It is well-known that for this sequence the three term recurrence formula holds

$$\lambda P_k(\lambda) = \alpha_{k\,k-1} P_{k-1}(\lambda) + \alpha_{kk} P_k(\lambda) + \alpha_{k\,k+1} P_{k+1}(\lambda)$$

where

$$\alpha_{kj} = \int_a^b \lambda P_k(\lambda) P_j(\lambda)\, d\phi(\lambda) \quad \text{for} \quad j=k-1, k$$

$$\alpha_{k\,k+1} = \Big(\int_a^b \lambda^2 P_k^2(\lambda)\, d\phi(\lambda) - \alpha_{k\,k-1}^2 - \alpha_{kk}^2\Big)^{1/2} > 0$$

are real coefficients.

Observe first that for any k,

$$Q_{k+1}(\lambda) = \lambda P_k(\lambda) - \alpha_{k\,k-1} P_{k-1}(\lambda) - \alpha_{kk} P_k(\lambda)$$

is non-zero polynomial because P_j is exactly of the degree j.

Since Q_{k+1} has only $k+1$ zeros our assumption on ϕ implies

$$\|Q_{k+1}\|_\phi^2 = (Q_{k+1}, Q_{k+1})_\phi > 0.$$

But for any $k=0, 1, 2, \ldots$

$$0 < \|Q_{k+1}\|_\phi^2 = \int_a^b \lambda^2 P_k^2(\lambda)\, d\phi(\lambda) - \alpha_{k\,k-1}^2 - \alpha_{kk}^2 = \alpha_{k\,k+1}^2,$$

thus the sequence of polynomials $\{P_k\}$ is infinite.

Observe that

$$x_0 = P_0(A)x = \frac{x}{\|x\|}$$

and assume that the condition (2.8) holds for $x_0, x_1, \ldots, x_k$.

Then

$$\alpha_{lj} = \int_a^b \lambda P_l(\lambda) P_j(\lambda)\, d\phi(\lambda) = (AP_l(A)P_j(A)x, x) =$$

$$= (AP_l(A)x, P_j(A)x) = (Ax_l, x_j) = a_{lj}$$

for $l=0, 1, \ldots, k; j=l-1, l$, and similarly

$$\alpha_{l\,l+1}^2 = (AP_l(A)x, AP_l(A)x) - a_{l\,l-1}^2 - a_{ll}^2 =$$

$$= \|Ax_l\|^2 - a_{l\,l-1}^2 - a_{ll}^2 = a_{l\,l+1}^2$$

for $l=0, 1, \ldots, k$. In other words, for $l=0, 1, \ldots, k$

$$\lambda P_l(\lambda) = a_{l\,l-1} P_{l-1}(\lambda) + a_{ll} P_l(\lambda) + a_{l\,l+1} P_{l+1}(\lambda).$$

Taking A for λ we get

$$AP_k(A)x = a_{k\,k-1} P_{k-1}(A)x + a_{kk} P_k(A)x + a_{k\,k+1} P_{k+1}(A)x.$$

Applying now the inductive hypothesis we obtain

$$Ax_k = a_{k\,k-1} x_{k-1} + a_{kk} x_k + a_{k\,k+1} x_{k+1}$$

and hence

$$x_{k+1} = P_{k+1}(A)x.$$

Finally,

$$\|x_{k+1}\|^2 = (P_{k+1}(A)x, P_{k+1}(A)x) = (P_{k+1}^2(A)x, x) =$$

$$= \int_a^b P_{k+1}^2(\lambda)\, d\phi(\lambda) = \|P_{k+1}\|_\phi^2 > 0;$$

thus the sequence $\{x_k\}$ is infinite.

Observe that the Lanczos process determines the infinite tridiagonal matrix of coefficients of the three term recurrence formulae (2.6) (2.7). Denote by T_N its N-th principal submatrix:

$$T_N = \begin{bmatrix} a_{00}, & a_{01}, & 0, & 0, & \ldots & \ldots, & 0 \\ a_{10}, & a_{11}, & a_{12}, & 0, & \ldots & \ldots, & 0 \\ 0 & . & . & . & \ldots & & 0 \\ . & . & . & . & \ldots & & . \\ . & . & . & . & \ldots & & . \\ . & . & . & . & \ldots & & \\ 0 & 0 & & & & a_{N-1\,N-2}, & a_{N-1\,N-1} \end{bmatrix}.$$

It is easy to prove that the zeros $s_1^N, s_2^N, \ldots, s_N^N$ of the orthogonal polynomial P_N^N are the eigenvalues of T_N. Since the numbers $s_1^N, s_2^N, \ldots, s_N^N$ are zeros of the orthogonal polynomial P_N, they are all distinct and contained in the open interval (a, b); thus we can assume that

$$a < s_1^N < s_2^N < \ldots \;\ldots < s_N^N < b \quad \text{see [10].}$$

Let H be the function defined in Section 1, and let $l_j^N(\lambda)$ be the basic Lagrange interpolatory polynomials corresponding to the knots $s_1^N, \dots, s_N^N$, i.e.

$$l_j^N(\lambda) = \frac{(\lambda-s_1^N)\dots(\lambda-s_{j-1}^N)(\lambda-s_{j+1}^N)\dots(\lambda-s_N^N)}{(s_j^N-s_1^N)\dots(s_j^N-s_{j-1}^N)(s_j^N-s_{j+1}^N)\dots(s_j^N-s_N^N)} = \frac{P_N(\lambda)}{(\lambda-s_j^N)P_N'(s_j^N)}.$$

Let $L_N(\cdot, f)$ be the Lagrange interpolatory polynomial for f and $s_1^N, \dots, s_N^N$. Define q_j^N and ϕ_N by the formulae:

$$\text{(2.10)} \qquad \begin{aligned} q_j^N &= \int_a^b l_j^N(\lambda)\, d\phi(\lambda) \qquad j=1, 2, \dots, N \\ \phi_N(\lambda) &= \sum_{j=1}^N q_j^N H(\lambda, s_j^N) \quad \text{for} \quad \lambda \in [a, b]. \end{aligned}$$

Then clearly

$$\phi_N(\lambda) = \int_a^b L_N(\mu, H(\lambda, \cdot))\, d\phi(\mu). \ \square$$

THEOREM 2.4 (the convergence theorem).

(i) *For* $j=1, 2, \dots, N$

$$q_j^N = \int_a^b [l_j^N(\lambda)]^2\, d\phi(\lambda) > 0$$

hence ϕ_N *are increasing step functions.*

(ii) *For any* $\lambda \in (a, b)$ *the estimate*

$$\text{(2.11)} \quad \phi(\lambda-0)-\phi(\lambda) \leqq \varliminf (\phi(\lambda)-\phi_N(\lambda)) \leqq \varlimsup (\phi(\lambda)-\phi_N(\lambda)) \leqq \phi(\lambda)-\phi(\lambda-0)$$

holds.

(iii) *For any point of continuity* $\lambda \in (a, b)$ *of the function* ϕ

$$\text{(2.12)} \qquad \phi_N(\lambda) \to \phi(\lambda) \quad \text{when} \quad N \to \infty;$$

and

$$\phi_N(a) = \phi(a) = 0, \quad \phi_N(b) = \phi(b) = \|x\|^2.$$

(iv) *The convergence* (2.12) *is uniform in any closed interval* $[\alpha, \beta] \subset [a, b]$ *such that any* $\lambda \in [\alpha, \beta]$ *is the point of continuity of* ϕ.

(v) *For any* ϱ *integrable in* $[a, b]$

$$\int_a^b \varrho(\lambda)\phi_N(\lambda)\, d\lambda \to \int_a^b \varrho(\lambda)\phi(\lambda)\, d\lambda \quad \text{when} \quad N \to \infty.$$

PROOF.

(i) Let w be a polynomial of degree $m \leqq 2N-1$; then

$$\text{(2.13)} \qquad \int_a^b w(\lambda)\, d\phi(\lambda) = \int_a^b L_N(\lambda, w)\, d\phi(\lambda).$$

Taking $w=(l_j^N)^2$ in (2.13) we get $m=2N-2$ and

$$L_N(\lambda, (l_j^N)^2) = l_j^N(\lambda)$$

hence

$$0 < \int_a^b [l_j^N(\lambda)]^2 \, d\phi(\lambda) = \int_a^b l_j^N(\lambda) \, d\phi(\lambda) = q_j^N.$$

(ii) (iii) Observe first that since the functions ϕ_N are increasing we have

$$\bigvee_a^b \phi_N = \phi_N(b) = \sum_{j=1}^N q_j^N H(b, s_j^N) = \sum_{j=1}^N q_j^N =$$

$$= \int_a^b \sum_{j=1}^N l_j^N(\lambda) \, d\phi(\lambda) = \int_a^b d\phi(\lambda) = \phi(b) = \|x\|^2.$$

Thus ϕ_N satisfy the assumption (ii) of Theorem 2.1.

Let W be the set of all polynomials (clearly $\overline{W}=Y=C[a, b]$) and let $w \in W$ be an arbitrary polynomial of degree m. If $N \geqq \frac{m+1}{2}$ then according to (2.13)

$$\int_a^b w(\lambda) \, d\phi(\lambda) = \int_a^b L_N(\lambda, w) \, d\phi(\lambda) = \int_a^b \sum_{j=1}^N l_j^N(\lambda) w_j(s_j^N) \, d\phi(\lambda) =$$

$$= \sum_{j=1}^N q_j^N w(s_j^N) = \sum_{j=1}^N q_j^N \int_a^b w(\lambda) \, dH(\lambda, s_j^N) = \int_a^b w(\lambda) \, d \sum_{j=1}^N q_j^N H(\lambda, s_j^N) =$$

$$= \int_a^b w(\lambda) \, d\phi_N(\lambda).$$

This means that

$$F_N(w) = \int_a^b w(\lambda) \, d\phi_N(\lambda) = F(w) = \int_a^b w(\lambda) \, d\phi(\lambda)$$

for $w \in W$, i.e. the hypothesis (i) of Theorem 2.1 is satisfied, and hence, for any $f \in Y$, $F_N(f) \to F(f)$ when $N \to \infty$.

From the condition $a < s_j^N < b$, $j=1, 2, \ldots, N$ it follows that $\phi_N(a)=\phi(a)=0$.

Since the functions ϕ_N and ϕ are increasing, we can apply (2.3) and (2.4). This gives (ii) and (iii). Conditions (iv) and (v) follow from Theorem 2.2 and the formula (2.5), respectively. □

As we shall show further, it is possible to compute the values of the function ϕ_N by means of a certain relatively simple algorithm (see § 5).

§ 3. Approximation of the spectral family

For fixed $\lambda \in [a, b]$, let

$$\Pi_N(\lambda, x) = L_N(A, H(\lambda, \cdot), x)$$

where $L_N(\cdot, H(\lambda, \cdot), x)$ is the Lagrange interpolatory polynomial for the function $H(\lambda, \cdot)$ (λ—fixed) with knots $s_1^N, s_2^N, \ldots, s_N^N$. Observe that s_j^N, $j=1, 2, \ldots, N$ depend on x (because $\phi(\lambda)=\|E(\lambda)x\|^2$ depends on x), so do also L_N and Π_N. We shall here exhibit this fact in notation. Notice that for x and λ fixed

$$\Pi_N(\lambda, x)\colon X \to X$$

is a linear selfadjoint operator. However, x can be also understood as the variable. So we have defined a nonlinear operator

$$\Pi_N(\lambda, \cdot)\cdot\colon X \to X.$$

Similarly, for $I=[c, d]\subset[a, b]$ the function

$$\Pi_N(I, x)\colon X \to X$$

defined by

$$\Pi_N(I, x) = L_N(A, H(d, \cdot)-H(c, \cdot), x)$$

is a linear operator, while

$$\Pi_N(I, \cdot)\cdot\colon X \to X$$

is a nonlinear one.

For the function $H(\lambda, \cdot)$ the following version of the ERDŐS—TURÁN theorem holds:

THEOREM 3.1. *If λ is a point of continuity of ϕ, then*

$$\int_a^b [L_N(\mu, H(\lambda, \cdot), x)-H(\lambda, \mu)]^2\, d\phi(\lambda) = \tag{3.1}$$

$$= \|L_N(\cdot, H(\lambda, \cdot)\cdot x)-H(\lambda, \cdot)\|_\phi^2 \to 0 \quad \textit{as} \quad N \to \infty.$$

SKETCH of the proof. We prove first that if λ is the point of continuity of ϕ and $\varrho\colon [a, b]\to \mathbf{R}$ is continuous then

$$Q_N(H(\lambda, \cdot)\cdot \varrho) \to \int_a^b H(\lambda, \mu)\varrho(\mu)\, d\phi(\mu) = \int_a^\lambda \varrho(\mu)\, d\phi(\mu)$$

where

$$Q_N(f) = \sum_{j=1}^N q_j^N f(s_j^N)$$

with

$$q_j^N = \int_a^b l_j^N(\mu)\, d\phi(\mu),$$

is the Gaussian quadrature of the function f.

From this we infer that

$$Q_N([H(\lambda, \cdot)-\varrho]^2) \to \int_a^b [H(\lambda, \mu)-\varrho(\mu)]^2\, d\phi(\mu) \quad \text{as} \quad N\to\infty,$$

and hence

$$0 \leqq \overline{\lim}\, \|L_N(\cdot, H(\lambda, \cdot)-\varrho, x)-(H(\lambda, \cdot)-\varrho)\|_\phi^2 \leqq 4\|H(\lambda, \cdot)-\varrho\|_\phi^2$$

as $N\to\infty$.

Now we can take for $\varepsilon>0$ arbitrary such a polynomial ϱ that $\|H(\lambda, \cdot)-\varrho\|_\phi^2<\varepsilon$, and (3.1) follows easily.

Complete proof can be found in [6]. □

We can rewrite the formula (3.1) in the following way

$$(3.2)\qquad \int_a^b [L_N(\mu, H(\lambda, \cdot), x)-H(\lambda, \mu)]^2\, d\phi_x(\mu) = \|[\Pi_N(\lambda, x)-E(\lambda)]x\|^2 \to 0$$

as $N\to\infty$, if $\lambda\in(a, b)$ is a point of continuity of ϕ_N.

Define now the space $Y(A, x)\subset X$ for any fixed $x\in X$:

$$Y(A, x) = \left\{y\in X \,|\, y = \sum_{j=0}^{\infty} \alpha_j A^j x \text{ and } \sum_{j=0}^{\infty} |\alpha_j|\,\|A\|^j < \infty\right\}.$$

Theorem 3.2.

1° *Let λ be a point of continuity of E. Then the nonlinear operator $\Pi_N(\lambda, \cdot)\cdot$: $:X\to X$ converges strongly to $E(\lambda)$ in X, i.e. for any $x\in X$,*

$$\|[\Pi_N(\lambda, x)-E(\lambda)]x\| \to 0 \quad when \quad N\to\infty.$$

2° *Let λ be a point of continuity of ϕ_x. Then the linear operator $\Pi_N(\lambda, x)\colon X\to X$ converges strongly to $E(\lambda)$ on $Y(A, x)$, i.e. for any $y\in Y(A, x)$,*

$$\|[\Pi_N(\lambda, x)-E(\lambda)]y\| \to 0 \quad when \quad N\to\infty.$$

Proof. 1° follows directly from (3.2). To prove 2° take $y=\sum_{j=0}^{\infty}\alpha_j A_x^j$ with $\sum_{j=0}^{\infty}|\alpha_j|\,\|A\|^j\le K<\infty$. By the commutativity, for any point of continuity $\lambda\in(a, b)$ of ϕ_x:

$$\|[\Pi_N(\lambda, x)-E(\lambda)]y\| = \left\|[\Pi_N(\lambda, x)-E(\lambda)]\sum_{j=0}^{\infty}\alpha_j A^j x\right\| \le$$

$$\le \sum_{j=0}^{\infty}|\alpha_j|\,\|A\|^j\,\|[\Pi_N(\lambda, x)-E(\lambda)]x\| \le K\|[\Pi_N(\lambda, x)-E(\lambda)]x\|$$

and the formula (3.2) implies

$$\|[\Pi_N(\lambda, x)-E(\lambda)]y\| \le K\|[\Pi_N(\lambda, x)-E(\lambda)]x\| \to 0$$

when $N\to\infty$. □

By linearity of the interpolatory polynomial L_N with respect to the interpolated function, we get for any $y\in Y(A, x)$ and $I=[c, d]\subset[a, b]$

$$\|[\Pi_N(I, x)-E(I)]y\| \to 0 \quad \text{when} \quad N\to\infty,$$

provided that c and d are points of continuity of ϕ_x.

§ 4. Rate of the convergence estimates

We shall first investigate the convergence of functions ϕ_N to ϕ in an interval of constancy of ϕ, i.e. an interval consisting of points of constancy of ϕ. Clearly any interval of constancy is open. For fixed N any interval of constancy of ϕ can contain at most one zero of the polynomial P_N, orthogonal with respect to $(\ ,\ )_\phi$ (see [10]). Such zero is called a "rubbish point" whenever, roughly speaking, it does not approximate ends of the interval of constancy when $N\to\infty$ (compare Lemma 4.1). Obviously, the existence or non-existence of "rubbish points" has an influence on the character of convergence.

LEMMA 4.1 (the rubbish points dumping). *Let $[\alpha,\beta]\subset[a,b]$ be a closed interval contained in an interval of constancy of ϕ and let $s\in[\alpha,\beta]$ be zero of the orthogonal polynomial P_N (i.e. s is a 'rubbish point'). Let*

$$q_s^N=\int\limits_a^b \frac{P_N(\lambda)}{(\lambda-s)P_N'(s)}\,d\phi(\lambda)>0.$$

Then for any natural p there is a constant $C(p)$ such that for $N>1+p$

$$0<q_s^N\leqq C(p)\cdot\left(\frac{b-c}{2(N-1)}\right)^p\cdot\omega_p\left(\frac{b-a}{2(2N-2-p)}\right)=o\left(\frac{1}{N^p}\right), \tag{4.1}$$

where ω_p is the modulus of continuity of a certain function from $C^\infty[a,b]$ depending on the interval $[\alpha,\beta]$ (hence $\omega_p(\tau)\to 0$ when $\tau\to 0$).

PROOF. Let (λ_1,λ_2) be an open interval of constancy of ϕ such that λ_1,λ_2 are points of increase of ϕ and $s\in[\alpha,\beta]\subset(\lambda_1,\lambda_2)$.

There exists a function $f_{[\alpha,\beta]}\colon [a,b]\to\mathbf{R}$ such that

$$f_{[\alpha,\beta]}\in C^\infty[a,b]$$

and

$$f_{[\alpha,\beta]}(\lambda)=\begin{cases}0 & \text{for } \lambda\leqq\lambda_1\\ 1 & \text{for } \lambda\in[\alpha,\beta]\\ 0 & \text{for } \lambda\geqq\lambda_2.\end{cases}$$

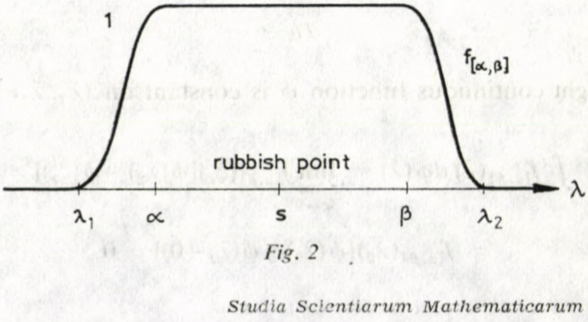


Fig. 2

Let Q_{2N-1} be an arbitrary polynomial of degree $2N-1$. Applying the notation of § 2 we can write

$$\int_a^b f_{[\alpha,\beta]}(\lambda)\,d\phi(\lambda)-\int_a^b L_N(\lambda, f_{[\alpha,\beta]})\,d\phi(\lambda)=$$

$$=\int_a^b [f_{[\alpha,\beta]}(\lambda)-Q_{2N-1}(\lambda)]\,d\phi(\lambda)+\int_a^b L_N(\lambda, Q_{2N-1}-f_{[\alpha,\beta]})\,d\phi(\lambda).$$

On the other hand, since $f_{[\alpha,\beta]}(s_j^N)=0$ for $s_j^N\neq s, j=1,2,\dots,N$ and $f_{[\alpha,\beta]}(s)=1$, we have

$$\int_a^b L_N(\lambda, f_{[\alpha,\beta]})\,d\phi(\lambda)=\int_a^b \frac{P_N(\lambda)}{(\lambda-s)P_N'(s)}\,d\phi(\lambda)=q_s>0. \tag{4.2}$$

Moreover

$$\int_a^b f_{[\alpha,\beta]}(\lambda)\,d\phi(\lambda)=\int_a^{\lambda_1} f_{[\alpha,\beta]}(\lambda)\,d\phi(\lambda)+\int_{\lambda_2}^b f_{[\alpha,\beta]}(\lambda)\,d\phi(\lambda)=0. \tag{4.3}$$

To verify (4.3) observe that

$$\int_{\lambda_1}^{\lambda_2} f_{[\alpha,\beta]}(\lambda)\,d\phi(\lambda)=\lim_{\substack{r\to\infty\\ \Delta r\to 0}}\sum_{k=1}^{r} f(\bar{\xi}_k)[\phi(\xi_{k+1})-\phi(\xi_k)]$$

where

$$\lambda_1=\xi_1<\xi_2<\dots<\xi_{r+1}=\lambda_2,\quad \xi_k\le\bar{\xi}_k\le\xi_{k+1}$$

and

$$\Delta r=\operatorname{Max}(\xi_{k+1}-\xi_k).$$

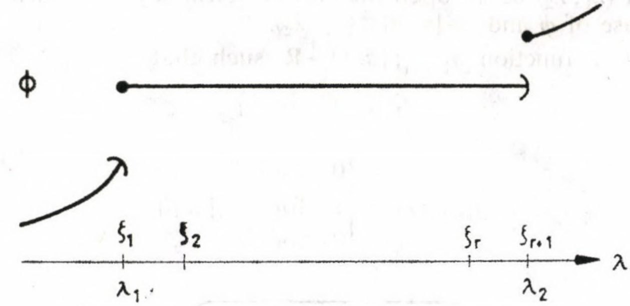


Fig. 3

Since the right continuous function ϕ is constant on (λ_1,λ_2) and $f_{[\alpha,\beta]}(\lambda_2)=0$, hence

$$\int_{\lambda_1}^{\lambda_2} f_{[\alpha,\beta]}(\lambda)\,d\phi(\lambda)=\lim_{\substack{r\to\infty\\ \Delta r\to 0}} f_{[\alpha,\beta]}(\bar{\xi}_r)[\phi(\lambda_2)-\phi(\xi_r)]=$$

$$=f_{[\alpha,\beta]}(\lambda_2)[\phi(\lambda_2)-\phi(\lambda_2-0)]=0.$$

From (4.2) and (4.3) it follows that

$$-q_s^N = \int_a^b [f_{[\alpha,\beta]}(\lambda)-Q_{2N-1}(\lambda)]\,d\phi(\lambda)+\int_a^b L_N(\lambda, Q_{2N-1}-f_{[\alpha,\beta]})\,d\phi(\lambda).$$

Since $q_j^N>0, j=1, 2, \ldots, N$ and $\sum_{j=1}^N q_j^N=\phi(b)=\|x\|^2$ we get the following estimate for q_S:

$$0 \leqq q_S^N \leqq 2\|f_{[\alpha,\beta]}-Q_{2N-1}\|_\infty \|x\|^2 \tag{4.4}$$

where $\|\cdot\|_\infty$ is the supremum norm taken over $[a, b]$.

The estimate (4.4) is valid for any polynomial Q_{2N-1} of degree $\leqq 2N-1$; hence

$$0 \leqq q_S^N \leqq 2\|x\|^2 \inf_{Q_{2N-1}} \|f_{[\alpha,\beta]}-Q_{2N-1}\|_\infty. \tag{4.5}$$

To evaluate the right-hand side of (4.5) we can apply JACKSON's Theorem [8] which states that if $g\in C^p[a, b]$ and $n>p+1$ then

$$\inf_{P_n}\|g-P_n\|_\infty \leqq \frac{6(3e)^p}{p+1}\left(\frac{b-a}{n-1}\right)^p \omega\left(g^{(p)}, \frac{b-a}{2(n-1-p)}\right)$$

where $\omega(f, \tau)$ is the modulus of continuity of f and P_n ranges over the set of all polynomials of degree $\leqq n$.

Let $\omega_p(\tau)=\omega(f_{[\alpha,\beta]}^{(p)}, \tau)$; we get

$$0 \leqq q_S^N \leqq \|x\|^2 \frac{6(3e)^p}{p+1}\left(\frac{b-a}{2N-2}\right)^p \omega_p\left(\frac{b-a}{2(2N-2-p)}\right);$$

Here

$$C(p) = \|x\|^2 \frac{6(3e)^p}{p+1}. \ \square$$

THEOREM 4.1. *Let* $[\alpha, \beta]\subset[a, b]$ *be a closed interval contained in a certain open interval of constancy of the function* ϕ. *Then for any natural* p *there exists a positive constant* $K(p)$ *such that for* $N>1+\frac{p}{2}$

$$|\phi(\lambda)-\phi_N(\lambda)| \leqq K(p)\cdot\left(\frac{b-a}{2(N-1)}\right)^p \omega_p\left(\frac{b-a}{2(2N-2-p)}\right) \tag{4.7}$$

holds uniformly for $\lambda\in[\alpha, \beta]$. *Here* ω_p *has the same meaning as in Lemma* 4.1 $(\omega_p(\tau)\to 0$ *when* $\tau\to 0)$.

If for N *sufficiently great the interval* $[\alpha, \beta]$ *contains no zeros of* P_N *("rubbish points"), then* ϕ_N *is constant in* $[\alpha, \beta]$.

PROOF. Let λ_1, λ_2 be points of increase of ϕ such that (λ_1, λ_2) is the open interval of constancy of ϕ, containing $[\alpha, \beta]$. Define the function $H_{[\alpha,\beta]}\in C^\infty[a, b]$ such that

$$0 \leqq H_{[\alpha,\beta]}(\mu) = \begin{cases} 1 & \text{for } a \leqq \mu \leqq \alpha \\ \leqq 1 & \text{for } \alpha \leqq \mu \leqq \beta \\ 0 & \text{for } \beta \leqq \mu \leqq b. \end{cases}$$

Clearly, such function exists:

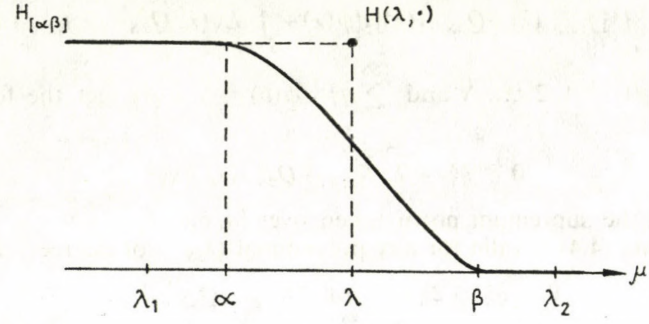


Fig. 4

Let $\lambda\in[\alpha,\beta]$. Then since ϕ is constant on the closed interval $[\alpha,\beta]$, we have

$$\phi(\lambda)=\int_a^b H(\lambda,\mu)\,d\phi(\mu)=\int_a^\alpha d\phi(\mu)+\int_\alpha^\beta H(\lambda,\mu)\,d\phi(\mu)=$$

$$=\phi(\alpha)+\int_\alpha^\beta H_{[\alpha,\beta]}(\mu)\,d\phi(\mu)=\int_a^b H_{[\alpha,\beta]}(\mu)\,d\phi(\mu).$$

In the interval $[\alpha,\beta]$ of constancy of ϕ there is at most one from among the points $s_j^N, j=1,2,\dots,N$, say $s_{j_0}^N$ such that $s_{j_0}^N\in[\alpha,\beta]$. Hence

$$\phi_N(\lambda)=\int_a^b L_N(\mu,H(\lambda,\cdot))\,d\phi(\mu)=\sum_{\substack{j=1\\ j\neq j_0}}^N q_j^N H(\lambda,s_j^N)+q_{j_0}^N H(\lambda,s_{j_0}^N)=$$

$$=\sum_{\substack{j=1\\ j\neq j_0}}^N q_j^N H_{[\alpha,\beta]}(s_j^N)+q_{j_0}^N H_{[\alpha,\beta]}(s_{j_0}^N)+q_{j_0}^N[H(\lambda,s_{j_0}^N)-H_{[\alpha,\beta]}(s_{j_0}^N)]=$$

$$=\int_a^b L_N(\mu,H_{[\alpha,\beta]})\,d\phi(\mu)+q_{j_0}^N[H(\lambda,s_{j_0}^N)-H_{[\alpha,\beta]}(s_{j_0}^N)].$$

This implies

$$\phi(\lambda)-\phi_N(\lambda)=$$

$$=\int_a^b[H_{[\alpha,\beta]}(\mu)-L_N(\mu,H_{[\alpha,\beta]})]\,d\phi(\mu)+q_{j_0}^N[H_{[\alpha,\beta]}(s_{j_0}^N)-H(\lambda,s_{j_0}^N)].$$

Observe that

$$|H_{[\alpha,\beta]}(s_{j_0}^N)-H(\lambda,s_{j_0}^N)|\leqq 1;$$

and by Lemma 4.1

$$0\leqq q_{j_0}^N\leqq C(p)\cdot\left(\frac{b-a}{2(N-1)}\right)^p\cdot\omega_p\left(\frac{b-a}{2(2N-2-p)}\right).$$

It remains to estimate

$$\int_a^b [H_{[\alpha,\beta]}(\mu)-L_N(\mu, H_{[\alpha,\beta]})]\, d\phi(\mu).$$

We can apply the same method as in the proof of Lemma 4.1. Let Q_{2N-1} be an arbitrary polynomial of degree $\leqq 2N-1$. Then

$$\Big|\int_a^b [H_{[\alpha,\beta]}(\mu)-L_N(\mu, H_{[\alpha,\beta]})\, d\phi(\mu)\Big| \leqq$$

$$\leqq \Big|\int_a^b [H_{[\alpha,\beta]}(\mu)-Q_{2N-1}(\mu)]\, d\phi(\mu)\Big| + \Big|\int_a^b L_N(\mu, Q_{2N-1}-H_{[\alpha,\beta]})\, d\phi(\mu)\Big| \leqq$$

$$\leqq \|H_{[\alpha,\beta]}-Q_{2N-1}\|_\infty \phi(b) + \sum_{j=1}^N q_j^N \|H_{[\alpha,\beta]}-Q_{2N-1}\|_\infty = 2\|x\|^2 \|H_{[\alpha,\beta]}-Q_{2N-1}\|_\infty.$$

Hence

$$\Big|\int_a^b [H_{[\alpha,\beta]}(\mu)-L_N(\mu, H_{[\alpha,\beta]})]\, d\phi(\mu)\Big| \leqq 2\|x\|^2 \inf_{Q_{2N-1}} \|H_{[\alpha,\beta]}-Q_{2N-1}\|_\infty$$

where Q_{2N-1} ranges over the set of all polynomials of degree $\leqq 2N-1$.

We can apply Jackson's theorem to evaluate the last inequality

$$\Big|\int_a^b [H_{[\alpha,\beta]}(\mu)-L_N(\mu, H_{[\alpha,\beta]})]\, d\phi(\mu)\Big| \leqq$$

$$\leqq 2\|x\|^2 C(p)\cdot\left(\frac{b-a}{2(N-1)}\right)^p \cdot \omega_p\left(\frac{b-a}{2(2N-2-p)}\right).$$

Finally we get

$$|\phi(\lambda)-\phi_N(\lambda)| \leqq (1+2\|x\|^2)C(p)\cdot\left(\frac{b-a}{2(N-1)}\right)^p \cdot \omega_p\left(\frac{b-a}{2(2N-2-p)}\right)$$

uniformly on $[\alpha, \beta]$. □

Let us pass to an estimate for intervals of continuity of ϕ.

LEMMA 4.2. (This is a slight modification of the so called "separation theorem". Proof is the same as in 3.4.11 of [10] or in [6].)

Let P_N be a polynomial of degree N, orthogonal with respect to the scalar product $(\ ,\ )_\phi$ and let $s_1^N, \ldots, s_N^N$ be its zeros. We assume here that a and b are points of increase of ϕ.

Put

$$q_j^N = \int_a^b l_j^N(\lambda)\, d\phi(\lambda) \qquad j = 1, 2, \ldots, N$$

with

$$l_j^N(\lambda) = \frac{P_N(\lambda)}{(\lambda-s_j^N)P_N'(s_j^N)}$$

and

$$z_j = q_1^N + q_2^N + \ldots + q_j^N \qquad j = 1, 2, \ldots, N.$$

Then there exists numbers

$$a < y_1^N < y_2^N \ldots < y_{N-1}^N < b = y_N^N$$

such that

$$\phi(y_j^N - 0) \leqq z_j^N \leqq \phi(y_j^N) \quad j = 1, 2, \ldots, N \tag{4.8}$$

$$\phi(s_j^N) < z_j^N < \phi(s_{j+1}^N - 0) \quad j = 1, 2, \ldots, N-1 \tag{4.9}$$

$$s_j^N < y_j^N < s_{j+1}^N \quad j = 1, 2, \ldots, N-1. \tag{4.10}$$

LEMMA 4.3. (An extension of 6.11.1 from [10].)

Let P_N be an N-th degree polynomial orthogonal with respect to $(\,,\,)_\phi$, let $s_1^N < < s_2^N < \ldots < s_N^N$ be its zeros. Take natural numbers n, m satisfying the condition $m(2n-1) \leqq 2N-1$.

Let

$$\psi(\vartheta) = \frac{a+b}{2} + \frac{a-b}{2}\cos\vartheta \quad \textit{for} \quad 0 \leqq \vartheta \leqq \pi.$$

$$\psi(\vartheta_j^N) = s_j^N \quad j = 1, 2, \ldots, N; \; 0 \leqq \vartheta_j^N \leqq \pi.$$

Consider a $\gamma \in [0, \pi]$ $\gamma \neq \vartheta_j^N$ for $j = 1, 2, \ldots, N$ and let ϑ_S^N be the nearest to γ among all $\vartheta_j^N, j = 1, 2, \ldots, N$. For n sufficiently great define

$$a_n = \begin{cases} \psi\left(\gamma - \dfrac{\pi}{2n}\right) & \textit{when} \quad \gamma \in (0, \pi) \\ a & \textit{when} \quad \gamma = 0, \end{cases}$$

$$b_n = \begin{cases} \psi\left(\gamma + \dfrac{\pi}{2n}\right) & \textit{when} \quad \gamma \in [0, \pi) \\ b & \textit{when} \quad \gamma = \pi. \end{cases}$$

Then

$$\left|\sin\frac{\gamma - \vartheta_s^N}{2}\right| \leqq \frac{\pi}{4n}\left(\frac{2\phi(b)}{\phi(b_n) - \phi(a_n)}\right)^{\frac{1}{2m}}. \tag{4.11}$$

PROOF. Let for a fixed γ

$$\varrho_{m,n}(\lambda) = \left[\frac{\sin n(\gamma - \vartheta)}{\sin\dfrac{\gamma - \vartheta}{2}}\right]^{2m} + \left[\frac{\sin n(\gamma + \vartheta)}{\sin\dfrac{\gamma + \vartheta}{2}}\right]^{2m}$$

where $\lambda = \psi(\vartheta)$, $0 \leqq \vartheta \leqq \pi$, and $a \leqq \lambda \leqq b$.

It is easy to show that $\varrho_{m,n}(\lambda)$ is in $[a, b]$ the polynomial of degree $m(2n-1)$ with respect to λ [10]. Put $\varphi = \gamma - \vartheta$ and

$$F_n(\varphi) = \frac{\sin n\varphi}{\sin\dfrac{\varphi}{2}}.$$

Applying the formula

$$F_n(\varphi) = F_{n-1}(\varphi)\cos\varphi + 2\cos(n-1)\varphi\cos\frac{\varphi}{2}$$

we prove by easy induction that F_n increases in the interval $\left[-\frac{\pi}{n}, 0\right]$ and decreases in the interval $\left[0, \frac{\pi}{n}\right]$. Moreover,

$$F_n(0) = \lim_{\varphi\to 0}\frac{\sin n\varphi}{\sin\frac{\varphi}{2}} = 2n$$

$$F_n\left(\frac{\pi}{2n}\right) = F_n\left(-\frac{\pi}{2n}\right) = \frac{1}{\sin\frac{\pi}{4n}} \geqq \frac{4}{\pi}n.$$

Hence for $\mu=\psi(\gamma)$

$$\varrho_{m,n}(\mu) \geqq (F_n(0))^{2m} = (2n)^{2m},$$

and for $\lambda\in[a_n, b_n]$

$$\varrho_{m,n}(\lambda) \geqq \left[F_n\left(\frac{\pi}{2n}\right)\right]^{2m} \geqq \left(\frac{4}{\pi}n\right)^{2m}. \tag{4.12}$$

On the other hand, if $m(2n-1)\leqq 2N-1$, applying (2.13) and (4.12) we get

$$\begin{aligned}\int_a^b \varrho_{m,n}(\lambda)\,d\phi(\lambda) = \sum_{j=1}^N \varrho_{m,n}(s_j^N)q_j^N \geqq \int_{a_n}^{b_n} \varrho_{m,n}(\lambda)\,d\phi(\lambda) \geqq \\ \geqq \left(\frac{4}{\pi n}\right)^{2m}[\phi(b_n)-\phi(a_n)].\end{aligned} \tag{4.13}$$

Using the same method as in [4] we can easily prove that

$$\left|\sin\frac{\gamma-\vartheta_s^N}{2}\right| \leqq \begin{cases}\left|\sin\frac{\gamma-\vartheta_j^N}{2}\right| \\ \left|\sin\frac{\gamma+\vartheta}{2}\right|\end{cases} \quad \text{for } j=1,2,\ldots,N$$

hence

$$\varrho_{m,n}(s_j^N) \leqq \frac{2}{\left(\sin\frac{\gamma-\vartheta_s^N}{2}\right)^{2m}}$$

and

$$\begin{aligned}\int_a^b \varrho_{m,n}(\lambda)\,d\phi(\lambda) = \sum_{j=1}^N \varrho_{m,n}(s_j^N)q_j^N \leqq \frac{2}{\left(\sin\frac{\gamma-\vartheta_s^N}{2}\right)^{2m}}\sum_{j=1}^N q_j^N = \\ = \frac{2}{\left(\sin\frac{\gamma-\vartheta_s^N}{2}\right)^{2m}}\phi(b).\end{aligned} \tag{4.14}$$

Combining (4.13) and (4.14) we get

$$\left(\frac{4}{\pi n}\right)^{2m}[\phi(b_n)-\phi(a_n)] \leqq \frac{2}{\left(\sin\frac{\gamma-\vartheta_s^N}{2}\right)^{2m}}\phi(b),$$

i.e.

$$\left|\sin\frac{\gamma-\vartheta_s^N}{2}\right|^{2m} \leqq \left(\frac{4}{\pi n}\right)^{2m}\frac{2\phi(b)}{\phi(b_n)-\phi(a_n)}.$$

THEOREM 4.2. *Assume that there exists a subinterval* $[\alpha, \beta]\subset[a, b]$ *such that for any pair* $\lambda_1, \lambda_2\in[\alpha, \beta]$ *satisfying the condition*

$$0 \leqq \lambda_2-\lambda_1 \leqq 1$$

the following inequality holds

$$L_2(\lambda_2-\lambda_1)^{\alpha_2} \leqq \phi(\lambda_2)-\phi(\lambda_1) \leqq L_1(\lambda_2-\lambda_1)^{\alpha_1} \tag{4.15}$$

where positive constants $L_1, L_2, \alpha_1, \alpha_2$ *depend only on* $[\alpha, \beta]$ *and* ϕ.

Then there exists a constant K depending on $[\alpha, \beta]$ *and* ϕ *such that for any* $m>\frac{\alpha_2}{2}$, $\lambda\in(\alpha, \beta)$ *and N sufficiently great the following estimate holds*

$$|\phi(\lambda)-\phi_N(\lambda)| \leqq \frac{K\left(\frac{m(m+3)}{2}\right)^{\alpha_1-\frac{\alpha_1\alpha_2}{2m}}}{N^{\alpha_1-\frac{\alpha_1\alpha_2}{2m}}}. \tag{4.16}$$

PROOF. We shall use the notation introduced in Lemma 4.3.

For arbitrary fixed positive integers m, n take $N\geqq\frac{m(2n-1)+1}{2}$. We shall first estimate the value $|s_j^N-s_{j+1}^N|$. For any fixed $j=1, 2, \dots, N$ take $\gamma=\frac{\vartheta_j^N+\vartheta_{j+1}^N}{2}$. Applying Lemma 4.3 we get

$$\left|\sin\frac{\gamma-\vartheta_j^N}{2}\right| = \left|\sin\frac{\gamma-\vartheta_{j+1}^N}{2}\right| \leqq \frac{\pi}{4n}\left(\frac{2\phi(b)}{\phi(b_n)-\phi(a_n)}\right)^{\frac{1}{2m}}. \tag{4.17}$$

Since

$$\psi(\gamma_1)-\psi(\gamma_2) = \frac{a-b}{2}[\cos\gamma_1-\cos\gamma_2] = (b-a)\sin\frac{\gamma_1+\gamma_2}{2}\sin\frac{\gamma_1-\gamma_2}{2} =$$

$$= (b-a)\sin\frac{2\gamma_1-(\gamma_1-\gamma_2)}{2}\sin\frac{\gamma_1-\gamma_2}{2} =$$

$$= (b-a)\left[\sin\alpha_1\cos\frac{\gamma_1-\gamma_2}{2}\sin\frac{\gamma_1-\gamma_2}{2}-\cos\gamma_1\left(\sin\frac{\gamma_1-\gamma_2}{2}\right)^2\right],$$

taking $\mu=\psi(\gamma)$, we obtain

$$\mu-s_k^N=(b-a)\left[\sin\gamma\cos\frac{\gamma-\vartheta_k^N}{2}-\cos\gamma\left(\sin\frac{\gamma-\vartheta_k^N}{2}\right)^2\right]$$

for $k=j$ and $j+1$ and hence

$$|s_{j+1}^N-s_j^N|=|\mu-s_{j+1}^N-\mu+s_j^N|=2(b-a)\left|\sin\gamma\cos\frac{\gamma-\vartheta_j^N}{2}\sin\frac{\gamma-\vartheta_j^N}{2}\right|,$$

because $\vartheta_{j+1}^N-\gamma=\gamma-\vartheta_j^N$. Applying now (4.17) we get

$$|s_{j+1}^N-s_j^N|\leqq\frac{(b-a)\pi}{2n}\left(\frac{2\phi(b)}{\phi(b_n)-\phi(a_n)}\right)^{\frac{1}{2m}} \tag{4.18}$$

for $j=1,2,\dots,N$, independently of j.

From the formula (4.15) it follows that any point of (α,β) is a point of increase of ϕ. Since near any point of increase of ϕ there is always a zero of P_N for N sufficiently great (see [4], [10]) hence for an arbitrary $\lambda_0\in(\alpha,\beta)$ we can find j, $1\leqq j\leqq N$ such that

$$\alpha<s_j^N\leqq\lambda_0<s_{j+1}^N<\beta\qquad(s_j^N<s_{j+1}^N)$$

(provided that N is sufficiently great).

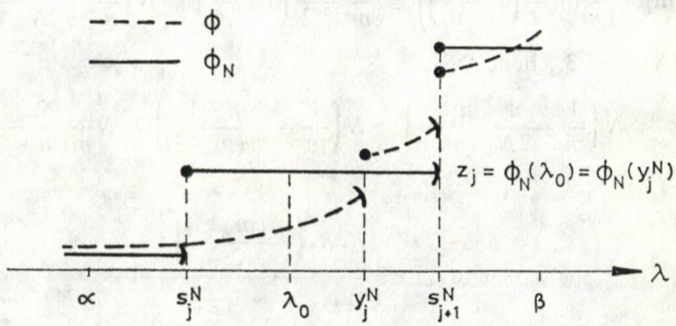


Fig. 5

Applying Lemma 4.2 (see formulae (4.9) and (4.10)) and the definition of ϕ_N, we get

$$\phi(s_j^N)<z_j=\phi_N(\lambda_0)<\phi(s_{j+1}^N-0)\leqq\phi(s_{j+1}^N).$$

On the other hand, since ϕ is increasing, we have

$$\phi(s_j^N)\leqq\phi(\lambda_0)<\phi(s_{j+1}^N).$$

Now, applying (4.15) we obtain

$$|\phi(\lambda_0)-\phi_N(\lambda_0)|<\phi(s_{j+1}^N)-\phi(s_j^N)\leqq\frac{K_1}{n^{\alpha_1}}\frac{1}{(b_n-a_n)^{\frac{\alpha_1\alpha_2}{2m}}}$$

where K_1 is some constant.

However,

$$b_n - a_n = \frac{a-b}{2}\left[\cos\left(\gamma + \frac{\pi}{2n}\right) - \cos\left(\gamma - \frac{\pi}{2n}\right)\right] =$$

$$= (b-a)\sin\gamma \sin\frac{\pi}{2n} \geqq \frac{\pi}{4n}(b-a)\sin\gamma$$

provided that $n \geqq 1$. Hence

$$\frac{1}{(b_n-a_n)^{\frac{\alpha_1\alpha_2}{2m}}} \leqq K_2 n^{\frac{\alpha_1\alpha_2}{2m}}$$

with another constant K_2. Therefore

$$|\phi(\lambda_0) - \phi_N(\lambda_0)| \leqq \frac{K_1 K_2}{n^{\alpha_1 - \frac{\alpha_1\alpha_2}{2m}}}.$$

Till now the only assumption on n is $n \leqq \frac{N}{m} + \frac{1}{2}\left(1 - \frac{1}{m}\right)$. Assume additionally that $N \geqq \frac{m+1}{2} + 1$ and put

$$n = \text{Entier}\left(\frac{N}{m} + \frac{1}{2}\left(1 - \frac{1}{m}\right)\right) \geqq \frac{N}{m} - \frac{1}{2}\left(1 + \frac{1}{m}\right) = N\left(\frac{1}{m} - \frac{1}{2N}\frac{m+1}{m}\right).$$

Since $2N \geqq m+3$, hence

$$n \geqq N\left(\frac{1}{m} - \frac{1}{2N}\frac{m+1}{m}\right) \geqq N\left(\frac{1}{m} - \frac{m+1}{m(m+3)}\right) = N\frac{2}{m(m+3)}$$

and finally

$$|\phi(\lambda_0) - \phi_N(\lambda_0)| \leqq \frac{K_1 K_2 \left(\frac{m(m+3)}{2}\right)^{\alpha_1 - \frac{\alpha_1\alpha_2}{2m}}}{N^{\alpha_1 - \frac{\alpha_1\alpha_2}{2m}}}. \square$$

Behaviour of the sequence $\{\phi_N\}$ *in points of discontinuity of* ϕ.

A very rough but general information on the behaviour of ϕ_N at the point of discontinuity of ϕ is furnished by Theorem 2.4 (see (2.11)). Here we shall investigate only a special case of the operator $A = A^*: X \to X$ such that $\sigma(A)$ is countable with the unique accumulation point $\mu \in \sigma(A)$. This class contains all the compact operators and there is no "rubbish points" in the approximation (see [4]).

Let

$$\sigma(A) = \{\lambda_\nu\}_{\nu=-\infty}^{\infty} \cup \{\mu\}$$

where

$$\mu = \lim_{|\nu| \to \infty} \lambda_\nu$$

and

$$\lambda_\nu > \lambda_{\nu+1} > \mu \quad \text{for} \quad \nu \geqq 0,$$

$$\lambda_\nu < \lambda_{\nu-1} < \mu \quad \text{for} \quad \nu < 0.$$

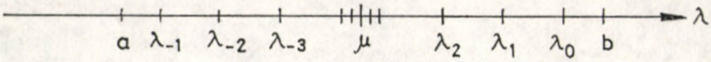


Fig. 6

Let

$$E_\nu = E(\lambda_\nu) - E(\lambda_{\nu-1}) \quad \text{for} \quad \nu \geqq 0,$$

$$E_\nu = E(\lambda_{\nu-1}) - E(\lambda_\nu) \quad \text{for} \quad \nu < 0,$$

then for any $x \in X$

$$x = \int_a^b dE(\lambda)x = \sum_{\nu=-\infty}^{\infty} E_\nu x.$$

For arbitrary fixed $x \neq 0$ let

$$\varphi_\nu = E_\nu x / \|E_\nu x\|$$

then φ_ν is an orthogonal sequence and

$$x = \sum_{\nu=-\infty}^{\infty} \varphi_\nu (x, \varphi_\nu) = \sum_{\nu=1}^{\infty} A_\nu \varphi_\nu$$

where

$$A_\nu = (x, \varphi_\nu) \qquad \nu = \ldots -1, 0, 1, \ldots .$$

Moreover

$$\|x\|^2 = \sum_{\nu=-\infty}^{\infty} A_\nu^2.$$

This enables us to apply here the theory of orthogonal polynomials on discrete set of points discussed in [4]. In this case ϕ as well as ϕ_N are piecewise constant, increasing functions on $[a, b]$.

From Theorem 3.3.7 of [4] it follows that if we consider for arbitrary $\delta > 0$ the points of discontinuity λ_ν of ϕ, lying outside of the interval $[\mu-\delta, \mu+\delta]$ (let us order them: $\lambda_1 < \lambda_2 < \lambda_3 < \ldots < \lambda_p$), then λ's less then μ are approximated by corresponding s_j from right-hand side, while λ's greater then μ are approximated from left-hand side. Every $\lambda_j, j=1, 2, \ldots, p$ is approximated by exactly one s_j^N. All the other s_j^N are contained in $[\mu-\delta, \mu+\delta]$. Applying now Lemma 4.2 and Theorem 2.4 we see that for N great enough the situation is like on the diagram below (given for $p=6$).

On this diagram arrows indicate the direction of convergence of s_j^N to λ_j and of ϕ_N to ϕ in the points of continuity of ϕ. Dots indicate the right limits contained in the graph. (—— or ——) indicate that the corresponding one sided limit does not belong to this part of the graph.

More exactly we can state the following

THEOREM 4.3. *Let $\lambda_1 < \lambda_2$ be two consecutive points of discontinuity of the function ϕ, i.e. such points of discontinuity that between them there is no other point of discontinuity of ϕ.*

1° *Assume $a < \lambda_1 < \lambda_2 < \mu$. Then*

(i) *for any $\xi \in (\lambda_1, \lambda_2)$, $\phi_N(\xi) > \phi(\xi)$ holds for N great enough and $\phi_N(\xi) \to \phi(\xi)$ when $N \to \infty$ uniformly in any closed subinterval of (λ_1, λ_2);*

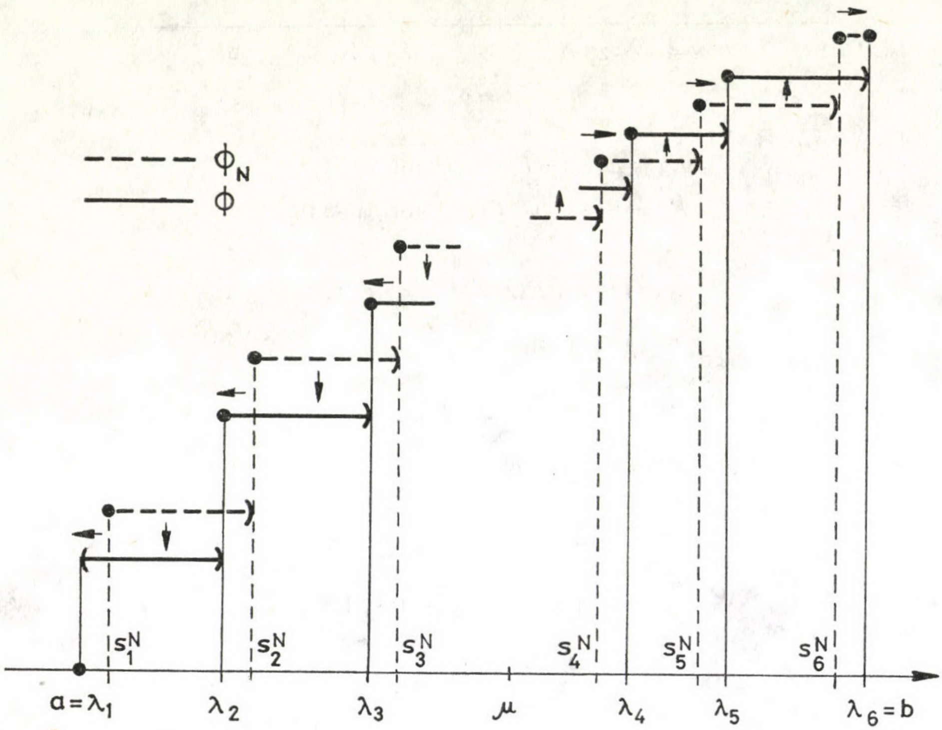


Fig. 7

(ii) $\Phi_N(\lambda_i) \leqq \phi(\lambda_i)$ *and* $\phi_N(\lambda_i) \to \phi(\lambda_i - 0)$ *when* $N \to \infty$ *for* $i=1, 2$.

2° *Assume* $\mu < \lambda_1 < \lambda_2 < b$. *Then*

(i) *for any* $\xi \in (\lambda_1, \lambda_2)$, $\phi_N(\xi) < \phi(\xi)$ *holds for N great enough and* $\phi_N(\xi) \to \phi(\xi)$ *when* $N \to \infty$, *uniformly in any closed subinterval of* (λ_1, λ_2);

(ii) $\phi_N(\lambda_i) \leqq \phi(\lambda_i)$ *and* $\phi_N(\lambda_i) \to \phi(\lambda_i)$ *when* $N \to \infty$ *for* $i=1, 2$.

3° $\phi_N(a) = 0$ *and* $\phi_N(b) = \phi(b)$.

PROOF. 1° (i) Since ξ belongs to the open interval (λ_1, λ_2) for N great enough $\xi \in (s_1^N, s_2^N)$ and hence $\phi_N(\xi) > \phi(\xi)$ follows from Lemma 4.2.

Since ξ is the point of continuity of ϕ the convergence $\phi_N(\xi) \to \phi(\xi)$ follows from Theorem 2.4.

(ii) Let λ_0 be the nearest to λ_1 point of discontinuity of ϕ less then λ_1; then

$$a < \lambda_0 < \lambda_1 < \lambda_2 < \mu.$$

Since, for N great enough $\dfrac{\lambda_{i-1} + \lambda_i}{2} \in (s_{i-1}^N, s_i^N)$, $i=1, 2$; hence applying Theorem 2.4, we get

$$\phi(\lambda_i) > \phi_N(\lambda_i) = \phi_N\left(\frac{\lambda_{i-1}+\lambda_i}{2}\right) \to \phi\left(\frac{\lambda_{i-1}+\lambda_i}{2}\right) = \phi(\lambda_i - 0).$$

Proof of 2° is similar; 3° is obvious because of Theorem 2.4. □

REMARK. Let now A be compact. Denote by $s^N_{i^N_j}$ the nearest number from the set $\{s^N_\nu\}^N_{\nu=1}$ to the number $\lambda_i, i=1, 2, ..., p$.

The easy extension of the result contained in [9] to the case of a compact self-adjoint operator enables us to state the following rate of convergence estimate:

$$|\lambda_i - s^N_{i^N_j}| \leqq Cq^{-2N}$$

where $q>1$ is some constant depending on the distance between λ_i and the remaining part of the spectrum, and independent of N.

§ 5. An algorithm

The definition of the function ϕ_N given in Section 2 (see (2.10)) seems to be rather uncomfortable for practical application.

Here we present another method of computing $\phi_N(\lambda)$.

Denote, as before by T_N the $N\times N$ real, symmetric, tridiagonal matrix determined by the coefficients of Lanczos' process (2.6)(2.7). Its eigenvalues are the numbers $s^N_1 < s^N_2 < ... < s^N_N$ which we shall use in computing the values of the function ϕ_N:

$$\phi_N(\lambda) = \sum_{j=1}^N q^N_j H(\lambda, s^N_j).$$

Since direct application of the formula

$$q^N_j = \int_a^b l^N_j(\lambda)\, d\phi(\lambda) = (l^N_j(A)x, x)$$

seems to be not very convenient (necessity of calculating $l^N_j(A)$!), we give here different expression for q^N_j.

Let u^N_j be the normalized eigenvector of T_N, corresponding to the eigenvalue $s^N_j, j=1, 2, ..., N$, and let $U_N=[u^N_1, u^N_2, ..., u^N_N]$ be $N\times N$ — orthogonal matrix composed of the vectors u^N_j as its columns.

THEOREM 5.1. *Put*

$$u^N_j = [u^N_{j_1}, u^N_{j_2}, ..., u^N_{j_N}]^T \qquad j = 1, 2, ..., N.$$

Then

$$q^N_j = \|x\|^2 (u^N_{j_1})^2, \qquad j = 1, 2, ..., N,$$

and hence

$$\phi_N(\lambda) = \|x\|^2 \sum_{j=1}^N (u^N_{j_1})^2 H(\lambda, s^N_j).$$

PROOF. Put

$$c_j^N = (P_0(s_j^N)^2 + P_1(s_j^N)^2 + \ldots + P_{N-1}(s_j^N)^2)^{1/2} > 0.$$

We prove first that

$$u_j^N = \left[\frac{P_0(s_j^N)}{c_j^N}, \frac{P_1(s_j^N)}{c_j^N}, \ldots, \frac{P_{N-1}(s_j^N)}{c_j^N}\right]^T; \quad j = 1, 2, \ldots, N. \tag{5.1}$$

Indeed, $\|u_j^N\| = 1$ and for $l=0, 1, \ldots, N-1; j=1, 2, \ldots, N$

$$a_{ll-1}P_{l-1}(s_j^N) + (a_{ll} - s_j^N)P_l(s_j^N) + a_{ll+1}P_{l+1}(s_j^N) = 0$$

while

$$P_N(s_j^N) = 0 \qquad j = 1, 2, \ldots, N.$$

The above formulae can be written in matrix form as follows

$$(T_N - s_j^N I)\begin{bmatrix} P_0(s_j^N) \\ \vdots \\ P_{N-1}(s_j^N) \end{bmatrix} \frac{1}{c_j^N} = 0,$$

hence (5.1) holds.

Since the matrix

$$U_N = \begin{bmatrix} \dfrac{P_0(s_1^N)}{c_1^N}, & \ldots, & \dfrac{P_0(s_N^N)}{c_N^N} \\ \vdots & & \vdots \\ \dfrac{P_{N-1}(s_1^N)}{c_1^N}, & \ldots, & \dfrac{P_{N-1}(s_N^N)}{c_N^N} \end{bmatrix}$$

is orthogonal, we have:

$$\begin{gathered} \sum_{j=1}^N \frac{P_0(s_j^N)^2}{(c_j^N)^2} = 1 \\ \sum_{j=1}^N \frac{P_0(s_j^N)P_l(s_j^N)}{(c_j^N)^2} = 0 \qquad l = 1, 2, \ldots, N-1. \end{gathered} \tag{5.2}$$

We can consider (5.2) as the system of N linear algebraic equations with non-singular matrix

$$\begin{bmatrix} P_0(s_1^N), & \ldots, & P_0(s_N^N) \\ \vdots & & \\ P_{N-1}(s_1^N), & \ldots, & P_{N-1}(s_N^N) \end{bmatrix}$$

and with unique solution

$$\left[\frac{P_0(s_1^N)}{(c_1^N)^2}, \ldots, \frac{P_0(s_N^N)}{(c_N^N)^2}\right]^T.$$

On the other hand, for any polynomial W of degree $\leqq 2N-1$ the following condition is satisfied

$$\int_a^b w(\lambda)\, d\phi(\lambda) = \int_a^b L_N(\lambda, w)\, d\phi(\lambda) = \sum_{j=1}^N q_j^N w(s_j^N).$$

Hence, taking $w(\lambda)=P_0(\lambda)P_l(\lambda)$, we get when $l=0$:

$$(5.3)\qquad 1=\int_a^b P_0^2(\lambda)\,d\phi(\lambda)=\sum_{j=1}^N q_j^N(P_0(s_j^N))^2=P_0^2\sum_{j=1}^N q_j^N=P_0^2\|x\|^2$$

because P_0 is constant.

Moreover, for $l=1, 2, \ldots, N-1$:

$$(5.4)\qquad 0=\int_a^b P_0(\lambda)P_l(\lambda)\,d\phi(\lambda)=\sum_{j=1}^N q_j^N P_0(s_j^N)P_l(s_j^N).$$

Notice that, because of the uniqueness of solution of (5.2), the equations (5.3), (5.4) yield

$$q_j^N=\frac{1}{(c_j^N)^2}\qquad j=1, 2, \ldots, N$$

and thus (5.1), (5.3) imply

$$(u_{j1}^N)^2=\frac{(P_0(s_j^N))^2}{(c_j^N)^2}=\frac{1}{\|x\|^2}\cdot\frac{1}{(c_j^N)^2}=\frac{q_j^N}{\|x\|^2}$$

for $j=1, 2, \ldots, N$. $\square$

Assume now that in the space X there is a sequence of points of the form $\{x_j^M\}$ for $j=1, 2, \ldots, r_M$; $M=1, 2, 3, \ldots$ such that for any $x\in X$ and any $\varepsilon>0$ we can find natural M_0 in such a way that for $M>M_0$ the coefficients $\alpha_1^M, \ldots, \alpha_{r_M}^M$ exist and satisfy the inequality

$$(5.5)\qquad \left\|x-\sum_{j=1}^{r_M}\alpha_j^M x_j^M\right\|<\varepsilon.$$

This is equivalent to the property of approximation of the space X by the family of finite dimensional subspaces $\{X_M\}_{M=1,2\ldots}$, where

$$X_M=\operatorname{Span}\{x_1^M, x_2^M, \ldots, x_{r_M}^M\}.$$

Many examples of such sequences are known when, for example, X is a Sobolev space. In this case even the rate of convergence of approximation (5.5) can be determined (see Galerkin and finite element method [1]). Thus in this case we can proceed in the following way:

1° Choose enough good sequence $x_1^M, \ldots, x_{r_M}^M$ to have satisfactory approximation for all elements $x\in X$, for which we want to have ϕ_x.

2° Applying formulae (1.5) and (1.6) from Section 1 as well as Theorem 5.1, form the matrix

$$\tilde{B}_{x_1^M,\ldots,x_{r_M}^M}(\lambda),$$

approximating the matrix $B_{x_1^M,\ldots,x_{r_M}^M}(\lambda)$ from the formula (1.6).

Taking now the values of the quadratic form $\alpha^*\tilde{B}_{x_1^M\ldots x_{r_M}^M}(\lambda)\alpha$ for $\alpha=[\alpha_1, \ldots, \alpha_{r_M}]^T$, we obtain the approximation of the function ϕ_x for any $x\in\operatorname{Span}\{x_1^M, \ldots, x_{r_M}^M\}$.

Rate of approximation of ϕ_x for any $x\in X$ is determined by Theorems from Section 4 and by the formula (1.7) from Section 1.

REFERENCES

[1] *The mathematical foundations of the finite element method with applications to partial differential equations,* Proceedings of a Symposium held at the University of Maryland, Baltimore, Md., June 26—30, 1972, edited by A. K. Aziz, Academic Press, New York—London, 1972. *MR* **49** #11824.

[2] Łojasiewicz, S., *Wstęp do teorii funkcji rzeczywistych,* Biblioteka Matematyczna, Tom 46, Państwowe Wydawnictwo Naukowe, Warszawa, 1973. *MR* **55** #5806.

[3] Kato, T., *Perturbation theory for linear operators,* Die Grundlehren der mathematischen Wissenschaften, Band 132, Springer-Verlag, New York, 1966. *MR* **34** #3324.

[4] Moszyński, K., A theorem on the approximation of the spectrum of a selfadjoint system of ordinary differential equations by the Lánczos process, *Studia Sci. Math. Hungar.* **10** (1975), 255—271. *MR* **58** #1367.

[5] Moszyński, K., Approssimazione spettrale di un operatore autoaggiunto con il metodo dei momenti, *Riv. Mat. Univ. Parma* (4) **6** (1980), 89—102.

[6] Moszyński, K., On approximation of the spectral density function of a selfadjoint operator, Inst. Math. Polish Acad. Sci., Warszawa, Preprint 206, 1980.

[7] Riesz, F. and Sz.-Nagy, B., *Leçons d'analyse fonctionnelle,* Akadémiai Kiadó, Budapest, 1952. *MR* **14**—286.

[8] Rivlin, T. J., *An introduction to the approximation of functions,* Blaisdell Publishing Co., Waltham, Mass.—Toronto, Ont.—London, 1969. *MR* **40** #3126.

[9] Saad, Y., Étude de la convergence des éléments propres calculés par les méthodes de Lánczos et Lánczos par blocs, Université Scientifique et Médicale de Grenoble, N⁰ 284, 1977.

[10] Szegő, G., *Orthogonal polynomials,* American Mathematical Society Colloquium Publications Vol. 23, Revised ed., American Mathematical Society, Providence, R. I., 1959. *MR* **21** #5029.

(Received March 24, 1980)

INSTYTUT MATEMATYCZNY PAN
UL. ŚNIADECKICH 8
P. O. BOX 137
PL—00-950 WARSZAWA
POLAND

ALTERNATING APPROXIMATION ON SHORTER INTERVALS

by

CHARLES B. DUNHAM

Abstract

Chebyshev approximation on intervals by a family with the alternating property for error curves of best approximations is studied. Whether uniform convergence occurs if a slightly shorter interval is used is examined.

For $g\in C[\alpha,\beta]$ define

$$\|g\|_\gamma = \sup\{|g(x)|:\ \alpha \leqq x \leqq \gamma\} \qquad \alpha < \gamma \leqq \beta.$$

Consider approximation of $f\in C(\alpha,\beta]$ with respect to this norm on $C[\alpha,\gamma]$. Let the family of approximations (F, P) be an alternating family ([26], 15 ff; [9], [13]) on $[\alpha,\beta]$: for every $A\in P$, there exists $\varrho(A)$ (the *degree* of F at A) such that $F(A,.)$ is best to $g\in C[\alpha,\beta]$ if and only if $g-F(A,.)$ alternates at least $\varrho(A)$ times on $[\alpha,\beta]$. It is known that a best approximation on $[\alpha,\beta]$ is unique (if it exists): denote it by $S(\gamma)$. In this note we consider the dependence of $S(\gamma)$ on γ as $\gamma\to\beta$.

This problem has practical interest. First, β might be a number not representable on the arithmetic at hand: for example 1/10 is not representable exactly in binary. Secondly, RALSTON suggests that trouble in ordinary rational approximations will occur if an approximation with degeneracy of one is best on a slightly longer interval ([25], 159, 170). CODY expresses the same idea less explicitly in [6] (407-case (i), 408-top diagram in Figure 1). We exhibit some of that trouble in this paper.

The major mathematical interest of the paper is the exhibition of simple and natural sequences of subsets for which uniform convergence of discretization fails. It is noted after the theorem that these sequences need not be intervals.

Approximation on $[\alpha,\gamma]$ is a special case of approximation on subsets, treated by the author in [13], [15], [21]. In [13] is shown

THEOREM. *Let $S(\beta)$ be non-degenerate then $\|S(\beta)-S(\gamma)\|_\beta\to 0$ as $\gamma\to\beta$.*

We observe also that if $f-S(\beta)=f-F(A,.)$ alternates $\varrho(A)$ times on $[\alpha,\gamma_0]$ for some $\gamma_0<\beta$, then $f-S(\beta)$ alternates $\varrho(A)$ times on $[\alpha,\gamma]$, $\gamma_0\leqq\gamma\leqq\beta$, implying $S(\gamma)=S(\beta)$. We are therefore left only with the case

HYPOTHESIS D. (a) $S(\beta)$ is degenerate;

(b) $f-S(\beta)=f-F(A,.)$ does not alternate $\varrho(A)$ times on $[\alpha,\gamma]$ for any $\gamma<\beta$.

1980 *Mathematics Subject Classification.* Primary 41A50.
Key words and phrases. Minimax, alternating, subinterval, non-uniform convergence.

We first examine this case for some simple but important alternating families of variable degree and then go on to more general families.

For several simple approximating families, in particular

(i) the ordinary rational family $R_m^0[\alpha, \beta]$ of ratios of constants to m^{th} degree polynomials $q, q>0$ on $[\alpha, \beta]$;

(ii) several families of the form $F(A, x)=a_1\varphi(a_2x)$ [12], of which the case $F(A, x)=a_1 \exp(a_2x)$ is of special importance, the only approximation not of maximum degree is the zero function, which has degree 1. Let Hypothesis D hold. For $\gamma<\beta, f-0$ does not alternate once, hence $S(\gamma)\neq 0$ and $S(\gamma)$ is of degree >1.There is no sequence $\gamma_k\to\beta$ such that $\|S(\gamma_k)-S(\beta)\|_\beta\to 0$, for $f-S_{\gamma(k)}$ alternates more than once on $[\alpha, \gamma_k]$ and $f-0$ alternates exactly once on $[\alpha, \beta]$.

Perhaps nearly all more complicated alternating families of practical interest are covered by

THEOREM. *Let $F(A, .)$ be of less than maximum degree. Let $G=\{F(B,.): \varrho(B)<$ $<$maximum degree$\}$ be alternating on $[\alpha, \beta]$ and $F(A,.)$ be of degree $\leqq\varrho(A)-1$ in G. Let $f-F(A, .)$ alternate $\varrho(A)$ times on $[\alpha, \beta]$ but not alternate $\varrho(A)$ times on $[\alpha, \gamma]$ for any $\gamma<\beta$. Then for any γ in*

$$[\sup\{x: |f(x)-F(A, x)| = \|f-F(A, .)\|_\beta, x<\beta\}, \beta),$$

$S(\gamma)$ (if it exists) is of degree $>\varrho(A)$.

PROOF. Suppose a γ exists for which the assertion fails. $f-F(A,.)$ does not alternate $\varrho(A)$ times on $[\alpha, \gamma]$, hence it is not best on $[\alpha, \gamma]$, hence

$$(*) \qquad \|f-F(B, .)\|_\gamma \equiv \|f-S(\gamma)\|_\gamma < \|f-F(A, .)\|_\gamma = \|f-F(A, .)\|_\beta.$$

By failure of the assertion $\varrho(B)\leqq\varrho(A)$ and $F(B, .)\in G$. Let $\{x_0, ..., x_{n-1}, x_n\}$ be any alternant of $f-F(A,.)$ on $[\alpha, \beta]$. By choice of $\gamma, x_{n-1}\leqq\gamma$. Hence $f-F(A, .)$ alternates $\varrho(A)-1$ times on $[\alpha, \gamma]$. But this implies that $F(A, .)$ is uniquely best in G on $[\alpha, \beta]$, giving a contradiction to $(*)$.

Use of previous reasoning for the case of 0 degenerate gives

COROLLARY. *There is no sequence $\{\gamma_k\}\to\beta$ with $\|S(\gamma_k)-S(\beta)\|_\beta\to 0$.*

An examination of the proof shows that the result generalizes to approximation on subsets X_k each containing $\{x_0, ..., x_{n-1}\}$ from an alternant $\{x_0, ..., x_{n-1}, x_n\}$ of $f-F(A, .)$.

This approach gives a much easier construction of failure of uniform convergence in approximation on subsets than that of the author in [11]: on the other hand, that latter construction made no assumption of location of error extrema except that they be nowhere dense and non-existence was not a problem.

We now consider applying the theorem. One application is to ordinary rational functions $R_m^n[\alpha, \beta]$ for $n\geqq 1$, perhaps with a multiplicative weight. It should be noted that rational families sometimes depend on the domain (e.g. [8]), but this possibility does not matter in this paper. Let $F(A, .)$ be an element of $R_m^n[\alpha, \beta]$ of less than maximum degree $(n+m+1)$. By the theory of ACHIESER ([28], p. 55), the following happens. If $F(A, .)\not\equiv 0$, it is of degree $n+m+1-d(A)$ in $R_m^n[\alpha, \beta]$, where $d(A)$ is the defect of $F(A,.)$ in $R^n_m[\alpha, \beta]$ and is >0. $F(A,.)\not\equiv 0$ is of degree $n+m-d(A)$ in $G=R^{n-1}_{m-1}[\alpha, \beta]$, an alternating family. If $F(A,.)=0$, it is of degree $n+1$ in $R^n_m[\alpha, \beta]$ and degree n in $R^{n-1}_{m-1}[\alpha, \beta]$. Thus the degree of $F(A,.)$ in G is precisely $\varrho(A)-1$ and the theorem applies.

Another application is to alternating generalized γ-polynomials [5], [20] $V_{n,m}(\Phi)$ of the form

$$F(A, x) = \sum_{k=1}^{n} a_k \phi(a_{n+k}x) + \sum_{k=1}^{m} a_{2n+k} x^{k-1}.$$

Normally, the degree is $2n+m-d(A)$, $d(A)$ the number of zeros in $\{a_1, \ldots, a_n\}$. A special case of these are alternating γ-polynomials [3]: the best-known of these are simple sums of exponentials ([24], 176ff; [27]). Other families have been studied in [14], [17], [19], [29]. In all known cases with families of maximum degree $\geqq 3$, the set of elements of less than maximum degree, namely those with at least one of $a_1, \ldots, a_n$ equal to zero, form the family $V_{n-1,m}(\varphi)$, an alternating family and elements of degree $\varrho(A)<2n+m$ (maximum) in $V_{n+m}(\varphi)$ are of degree $\varrho(A)-1$ in $V_{n-1,m}(\varphi)$. Thus the theorem also applies.

We now give an example to show that problems do not necessarily occur for $F(A,.)$ merely degenerate

EXAMPLE. Let $[\alpha, \beta]=[-1, 1]$. Approximate by

$$F(A, x) = \begin{cases} a_1 \exp(a_2 x) & a_1 > 0 \\ a_1 & a_1 \leqq 0. \end{cases}$$

Use of theory for the family $F(A, x)=a_1 \exp(a_2 x)$ and the family of constants shows that the above is an alternating family on $[-1, 1]$ with $\varrho(A)=2$ if $a_1>0$ and $\varrho(A)==1$ if $a_1\leqq 0$. The zero approximation is degenerate by the definition of [13], p. 99. Now let $f(x)=x$. As $f-0$ alternates once on $[-1, 1]$, 0 is best on $[-1, 1]$. For any $\gamma<1$, there is a negative constant $S(\gamma)$ such that $f-S(\gamma)$ alternates exactly once on $[-1, \gamma]$. Theory for approximation by constants guarantees that $\|S(\gamma)-S(B)\|_\beta\to 0$ as $\gamma\to\beta$.

It is not difficult to see that the theorem can be extended to generalizations of the Chebyshev problem, including simultaneous approximation [7], use of a weight function [9], and restricted range approximation [16]. Alternating families can be replaced by alternating families with a fixed point [18], partly alternating families [10], and perhaps the families varisolvent in the sense of GILLOTTE and MCLAUGHLIN [22].

A non-standard alternating theory for approximation on $[0, \infty]$ by reciprocals of polynomials (and more generally, Williams' type decaying rationals) is due to BRINK and TAYLOR [4]. The exceptional case in their theory is when the polynomial in the best approximation is of less than maximum (polynomial) degree and one less alternation than standard on $[0, \infty]$ occurs. As standard alternation occurs on all finite subintervals $[0, \gamma]$, we cannot in this exceptional case have uniform convergence on $[0, \infty]$ of $S(\gamma)$ to $S(\infty)$.

What happens in the problem of BLATT [1] would be of interest.

The case where (F, P) is an alternating family on $[\alpha, \delta]$, where $\delta > \beta$, and $\gamma \to \beta$ from above is also of interest. Uniform convergence of $S(\gamma)$ to $S(\beta)$ when $S(\beta)$ is non-degenerate is proven in [21]. The situation for $S(\beta)$ degenerate is open. It appears that much more powerful techniques than those of this paper are needed.

REFERENCES

[1] BLATT, H. P., Rationale Tschebyscheff-Approximation über unbeschränkten Intervallen, *Numer. Math.* **27** (1976/77), 179—190. *MR* **57** #13303.

[2] BRAESS, D., Chebyshev approximation by exponentials on finite subsets, *Math. Comp.* **27** (1973), 327—331. *MR* **48** #9191.

[3] BRAESS, D., Chebyshev approximation by γ-polynomials, *J. Approximation Theory* **9** (1973), 20—43. *MR* **51** #13533.

[4] BRINK, D. and TAYLOR, G. D., Chebyshev approximation by reciprocals of polynomials on $[0, \infty)$, *J. Approximation Theory* **16** (1976), 142—149. *MR* **53** #1112.

[5] BRINK—SPALINK, J., Tschebyscheff-Approximation durch γ-Polynome mit teilweise fixierten Frequenzen, *J. Approximation Theory* **15** (1975), 60—77. *MR* **52** #6269.

[6] CODY, W. J., A survey of practical rational and polynomial approximation of functions, *SIAM Rev.* **12** (1970), 400—423. *MR* **42** #2627.

[7] DUNHAM, C., Simultaneous Chebyshev approximation of functions on an interval, *Proc. Amer. Math. Soc.* **18** (1967), 472—477. *MR* **35** #3334.

[8] DUNHAM, C., Rational Chebyshev approximation on subsets, *J. Approximation Theory* **1** (1968), 484—487. *MR* **38** #6279.

[9] DUNHAM, C., Chebyshev approximation with respect to a weight function, *J. Approximation Theory* **2** (1969), 223—232. *MR* **40** #6137.

[10] DUNHAM, C., Partly alternating families, *J. Approximation Theory* **6** (1972), 378—386. *MR* **51** # 13546.

[11] DUNHAM, C., Approximation by alternating families on subsets, *Computing (Arch. Elektron. Rechnen)* **9** (1972), 261—265. *MR* **47** #1242.

[12] DUNHAM, C., Chebyshev approximation by $A^*\varphi(Bx)$, *Z. Angew. Math. Mech.* **53** (1973), 353—354. *MR* **52** #8741.

[13] DUNHAM, C., Alternating Chebyshev approximation, *Trans. Amer. Math. Soc.* **178** (1973), 95—109. *MR* **47** #7282.

[14] DUNHAM, C., Chebyshev approximation by logarithmic families, *Z. Angew. Math. Mech.* **53** (1973), 352—353. *MR* **52** #8740.

[15] DUNHAM, C., Varisolvent Chebyshev approximation on subsets, *Approximation Theory* (Proc. Internat. Sympos., Univ. Texas, Austin, Tex., 1973), ed. by G. G. Lorentz, Academic Press, New York, 1973, 337—340. *MR* **48** #11859.

[16] DUNHAM, C., Alternating minimax approximation with unequal restraints, *J. Approximation Theory* **10** (1974), 199—205. *MR* **50** #5301.

[17] DUNHAM, C., Chebyshev approximation by $A+B^*\varphi(Cx)$, *Computing (Arch. Elektron. Rechnen)* **13** (1974), 205—213. *MR* **52** #14773.

[18] DUNHAM, C., Alternation with a null point, *J. Approximation Theory* **15** (1975), 157—160. *MR* **52** #11430.

[19] DUNHAM, C., Chebyshev approximation by polynomial-rational sums, *Z. Angew. Math. Mech.* **55** (1975), 758—759. *MR* **53** #6175.

[20] DUNHAM, C., Behavior of the Chebyshev operator of best approximation from a curve of functions, *J. Approximation Theory* **21** (1977), 209—213. *MR* **57** #975.

[21] DUNHAM, C., Nearby alternating Chebyshev approximation, *Studia Sci. Math. Hungar.* **10** (1975), 381—387. *MR* **80m**: 41015.

[22] GILLOTTE, M. and MCLAUGHLIN, H., On nonlinear uniform approximation, *J. Approximation Theory* **21** (1977), 174—187. *MR* **56** #16222.

[23] LOEB, H. L. and WOLFE, J. M., Discrete nonlinear approximation, *J. Approximation Theory* **7** (1973), 365—385. *MR* **50** #843.

[24] MEINARDUS, G., *Approximation of functions: Theory and numerical methods,* Springer Tracts in Natural Philosophy, Vol. 13, Springer-Verlag New York, Inc., New York, 1967. *MR* **36** # 571.

[25] RALSTON, A., Some aspects of degeneracy in rational approximation, *J. Inst. Math. Appl.* **11** (1973), 157—170. *MR* **49** # 7672.

[26] RICE, J., *The approximation of functions, Vol. 2: Nonlinear and multivariate theory,* Addison-Wesley Publishing Co., Reading, Mass.—London—Don Mills, Ont., 1969. *MR* **39** # 5989.

[27] WERNER, H. K. E., Tschebyscheff-approximation with sums of exponentials, *Approximation Theory* (Proc. Sympos., Lancaster, 1969), ed. by A. Talbot, Academic Press, London, 1970, 109—136. *MR* **42** # 3479.

[28] ACHIESER, N. I., *Theory of approximation,* Frederick Ungar Publishing Co., New York, 1956. *MR* **20** # 1872.

[29] DUNHAM, C. B., Chebyshev approximation by polynomials, plus $B\varphi(CX)$, *J. Approx. Theory* **24** (1978), 344—349. *MR* **80h**: 41014.

(Received March 28, 1980)

COMPUTER SCIENCE DEPARTMENT
UNIVERSITY OF WESTERN ONTARIO
LONDON, ONTARIO
N6A 5B9
CANADA

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 415—418

(n, T)-TOWERS FOR APERIODIC TRANSFORMATIONS T

by

CHRISTOPH KOPF

Abstract

(n, T)-towers are introduced for nonsingular invertible transformations T and it is shown that the existence of (n, T)-towers for every $n \geqq 2$ is necessary and sufficient for the aperiodicity of T. As a consequence the uniform approximation theorem is obtained for arbitrary probability spaces.

Let $(\Omega, \mathfrak{A}, \mu)$ denote a probability space and let T denote an invertible nonsingular transformation from Ω into itself, i.e. $T: \Omega \to \Omega$ is bijective, T and T^{-1} are measurable and $\mu(T^{-1}A)=0=\mu(TA)$ for all $A \in \mathfrak{A}$ with $\mu(A)=0$. An invertible transformation T is called *aperiodic* iff for every natural number $n \geqq 1$ and every measurable set A with $\mu(A)>0$ there exists a measurable subset B of A such that $\mu(B \triangle T^{-n}B)>0$. We write $A=B$ for $\mu(A \triangle B)=0$.

DEFINITION. Let $n \geqq 2$ be fixed. A family $(U_{i,j})_{1 \leqq i \leqq n, 1 \leqq j < n+i}$ of measurable sets $U_{i,j}$ is called *(n, T)-tower* iff

$$U_{i,j} \cap U_{k,l} = \emptyset \quad \text{for} \quad (i,j) \neq (k,l), \quad 1 \leqq i, k \leqq n, \quad 1 \leqq j < n+i, \quad 1 \leqq l < n+k;$$

$$T^{-1} U_{i,j} = U_{i,j-1} \quad \text{for} \quad 1 \leqq i \leqq n, \quad 2 \leqq j < n+i;$$

$$T^{-1}\left(\bigcup_{i=1}^{n} U_{i,1}\right) = \bigcup_{i=1}^{n} U_{i,n+i-1}$$

and

$$\Omega = \bigcup_{\substack{1 \leqq i \leqq n \\ 1 \leqq j < n+i}} U_{i,j}.$$

The set $\bigcup_{i=1}^{n} U_{i,1}$ is called the *basis* and the set $\bigcup_{i=1}^{n} U_{i,n+i-1}$ is called the *top* of the (n, T)-tower $(U_{i,j})$.

THEOREM. *Let T denote an invertible nonsingular transformation on Ω. The following conditions are equivalent:*

(i) *T is aperiodic;*

(ii) *For every $n \geqq 2$ there exists a measurable set G with $\bigcup_{i=-n+1}^{n-1} T^i G = \Omega$ such that $G, T^{-1}G, \ldots, T^{-n+1}G$ are mutually disjoint;*

AMS (MOS) subject classifications (1970). Primary 28A65; Secondary 54H20.
Key words and phrases. Tower of a transformation, nonsingular transformation, aperiodic transformation, periodic transformation, uniform approximation theorem.

(iii) *For every* $n\geqq 2$ *there exists an* (n, T)*-tower with a basis (or with a top) of* μ*-measure less or equal than* $\frac{\mu(\Omega)}{n}$.

We note that there is a maximal invariant measurable subset of Ω on which T is aperiodic (see e.g. [3]).

PROOF.

(i)$\Rightarrow$(ii).

Let $n\geqq 2$ be fixed. Since T is aperiodic there exists an $(n-1, \varepsilon)$-Rohlin set for every $\varepsilon>0$, i.e., a measurable set D with $\mu\left(\bigcup_{i=0}^{n-1} T^{-i}D\right)>\mu(\Omega)-\varepsilon$ such that $D, T^{-1}D, \dots, T^{-n+1}D$ are mutually disjoint (see [2]). For $\varepsilon_1=\sup\,\{\mu(A)|A\in\mathfrak{A};\ A, T^{-1}A, \dots, T^{-n+1}A$ are mutually disjoint$\}$ we therefore conclude $\varepsilon_1>0$. Assume a measurable set A such that $\mu(A)>2^{-1}\varepsilon_1$ and $A, T^{-1}A, \dots, T^{-n+1}A$ are mutually disjoint. Let

$$\varepsilon_2=\sup\Big\{\mu(B)|B\in\mathfrak{A};\ B\subset\Omega\setminus\bigcup_{i=-n+1}^{n-1} T^iA;\ B, T^{-1}B, \dots, T^{-n+1}B \text{ are mutually disjoint}\Big\}.$$

If $\mu\left(\Omega\setminus\bigcup_{i=-n+1}^{n-1} T^iA\right)>0$ it follows from Rohlin's theorem that $\varepsilon_2>0$ and we choose a measurable subset B of $\Omega\setminus\bigcup_{i=-n+1}^{n-1} T^iA$ such that $\mu(B)>2^{-1}\varepsilon_2$ and $B, T^{-1}B, \dots, T^{-n+1}B$ are mutually disjoint. We repeat this argument on $\Omega\setminus\bigcup_{i=-n+1}^{n-1} T^i(A\cup B)$ and so on. We obtain a measurable set G with $\bigcup_{i=-n+1}^{n-1} T^iG=\Omega$ such that $G, T^{-1}G, \dots, T^{-n+1}G$ are mutually disjoint.

(ii)$\Rightarrow$(iii).

Let $n\geqq 2$ be fixed and let G be a measurable set with $\bigcup_{i=-n+1}^{n-1} T^iG=\Omega$ such that $G, T^{-1}G, \dots, T^{-n+1}G$ are mutually disjoint. From $\bigcup_{i=1}^{2n-1} T^{-i}G=\Omega=\bigcup_{j=1}^{2n-1} T^jG$ we conclude $G\subset\sum_{i=n}^{2n-1} T^{-i}G$ because the n sets $T^{-n+1}G, \dots, T^{-1}G, G$ are mutually disjoint and in the same way we conclude $G\subset\sum_{j=n}^{2n-1} T^jG$ because the n sets $G, TG, \dots, T^{n-1}G$ are mutually disjoint. We define $H_i=T^{n-1+i}G\cap G$ for $1\leqq i\leqq n$ and obtain $\sum_{i=1}^{n} H_i=G=\sum_{j=1}^{n} T^{-n+1-j}H_j$.

For an arbitrary integer k_0 we set $U_{i,j}=T^{k_0-n-i+j}H_i$ for $1\leqq i\leqq n$, $1\leqq j<n+i$ and obtain thereby the desired (n, T)-tower. To show

$$U_{i,j}\cap U_{k,l}=T^{k_0+j-1}G\cap T^{k_0-n-i+j}G\cap T^{k_0+l-1}G\cap T^{k_0-n-k+l}G=\emptyset$$

we first assume $j<l$. If $0<l-j<n$ then

$$T^{-k_0-j+1}(U_{i,j}\cap U_{k,l})\subset G\cap T^{l-j}G=\emptyset.$$

If $n \leqq l-j < n+k-1$ then

$$T^{-k_0-j+1}(U_{i,j} \cap U_{k,l}) \subset G \cap T^{-n-k+l-j+1} G = \emptyset$$

because $-n+1 \leqq -n-k+l-j+1 < 0$.

If $j=l$ and $i \neq k$ then

$$T^{-k_0+n-j}(U_{i,j} \cap U_{k,j}) \subset T^{-i} G \cap T^{-k} G = \emptyset.$$

The relations $T^{-1}U_{i,j} = U_{i,j-1}$ for $1 \leqq i \leqq n, 2 \leqq j < n+i$ are evident.

Furthermore we have

$$T^{-1}\left(\bigcup_{i=1}^{n} U_{i,1}\right) = T^{-1}\left(\bigcup_{i=1}^{n} T^{k_0-n-i+1} H_i\right) = T^{k_0-1} G = \bigcup_{i=1}^{n} T^{k_0-1} H_i = \bigcup_{i=1}^{n} U_{i,n+i-1}.$$

Finally we show

$$\Omega = \bigcup_{\substack{1 \leqq i \leqq n \\ 1 \leqq j < n+i}} U_{i,j}.$$

Let $\omega \in \sum_{i=-n+1}^{n-1} T^{k_0+i} G = \Omega$ be fixed. If $\omega \in \sum_{j=0}^{n-i} T^{k_0+j} G$, for instance $\omega \in T^{k_0+j_0} G$ with $0 \leqq j_0 \leqq n-1$, then, since $G = \sum_{i=1}^{n} T^{-n+1-i} H_i$ for some i_0 with $1 \leqq i_0 \leqq n$ we have

$$\omega \in T^{k_0+j_0-n+1-i_0} H_{i_0} = U_{i_0, j_0+1}.$$

If $\omega \in \sum_{j=-n+1}^{-1} T^{k_0+j} G$, for instance $\omega \in T^{k_0-n+j_0} G$ with $1 \leqq j_0 \leqq n-1$, then, since $G = \sum_{i=1}^{n} H_i$, for some i_0 with $1 \leqq i_0 \leqq n$ we have

$$\omega \in T^{k_0-n+j_0} H_{i_0} = U_{i_0, j_0+i_0}.$$

Since $T^{k_0} G$ is the basis and $T^{k_0-1} G$ is the top of the (n, T)-tower $(U_{i,j})$ by a suitable choice of the integer k_0 we get $\mu(T^{k_0} G)$ or $\mu(T^{k_0-1} G)$ less or equal than $\dfrac{\mu(\Omega)}{n}$.

(iii)$\Rightarrow$(i).

Assume a measurable set A with $\mu(A) > 0$ and assume a natural number $n \geqq 1$. Let $(U_{i,j})$ denote an $(n+1, T)$-tower. Since $\Omega = \bigcup_{\substack{1 \leqq i \leqq n+1 \\ 1 \leqq j < n+i+1}} U_{i,j}$ there exist natural numbers i_0 and j_0 with $1 \leqq i_0 \leqq n+1$ and $1 \leqq j_0 < n+i_0+1$ such that $\mu(A \cap U_{i_0, j_0}) > 0$. We obtain

$$\mu\big((A \cap U_{i_0, j_0}) \triangle T^{-n}(A \cap U_{i_0, j_0})\big) > 0$$

because $U_{i_0, j_0} \cap T^{-n} U_{i_0, j_0} = \emptyset$.

As a consequence we get the uniform approximation theorem for arbitrary probability spaces, which is proved in [1], §7 for the unit interval by using a more special definition of periodicity. We call a nonsingular invertible transformation T periodic on Ω iff there is a natural number $n \geqq 1$ such that $A = T^{-n} A$ for every measurable subset A of Ω.

COROLLARY. *For every* $\varepsilon>0$ *and for every nonsingular invertible transformation* T *on* Ω *there exists a nonsingular invertible periodic transformation* $\bar{T}$ *on* Ω *such that* $\mu(\{\omega\in\Omega|T\omega\neq\bar{T}\omega\})<\varepsilon$, *i.e. the periodic transformations are dense in the set of nonsingular invertible transformations with respect to the uniform topology.*

PROOF. The essential step is to show that for every $\varepsilon>0$ and for every nonsingular invertible aperiodic transformation T on Ω there exists a nonsingular invertible periodic transformation $\bar{T}$ on Ω such that $\mu(\{\omega\in\Omega|T\omega\neq\bar{T}\omega\})<\varepsilon$. This follows from the existence of an (n, T)-tower with $\frac{\mu(\Omega)}{n}<\varepsilon$ and with a top of μ-measure less or equal than $\frac{\mu(\Omega)}{n}$ by setting

$$\bar{T}\omega=\begin{cases}T\omega & \text{for } \omega\in U_{i,j},\ 1\leqq i\leqq n,\ 1\leqq j<n+i-2,\\ T^{-n-i+2}\omega & \text{for } \omega\in U_{i,j},\ 1\leqq i\leqq n,\ j=n+i-1.\end{cases}$$

By decomposing Ω into the aperiodic part and into the periodic parts ([3], Theorem 4) the assertion now follows easily.

REFERENCES

[1] FRIEDMAN, N. A., *Introduction to ergodic theory,* Van Nostrand Reinhold Mathematical Studies. No. 29, Van Nostrand Reinhold Co., New York—Toronto, Ont.—London, 1970, *MR* **55** #8310.

[2] KOPF, CH., On a theorem of V. A. Rohlin, *Math. Nachr.* **92** (1979), 161—162.

[3] KOPF, CH., Negative nonsingular transformations, *Ann. Inst. Henri Poincaré Sect. B* **18** (1982), 81—102.

(Received March 28, 1980)

INSTITUT FÜR MATHEMATIK DER
UNIVERSITÄT INNSBRUCK
INNRAIN 52
A—6020 INNSBRUCK
AUSTRIA

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 419—425

ON CONNECTEDNESS OF A RANDOM GRAPH WITH A SMALL NUMBER OF EDGES

by

Z. FÜREDI

Abstract

Consider the $n \times n$ lattice graph $G(n)$. Let $G(n, p)$ denote the random subgraph of $G(n)$ defined by choosing the edges of $G(n)$ with probability p mutually independently.

We prove: $p < 1/3$ is fixed then $G(n, p)$ is highly nonconnected, i. e. the largest component of $G(n, p) < C(p) \log n$, if $p > 2/3$ then $G(n, p)$ is nearly connected, i.e. there is a giant component. But $G(n, p)$ will be connected if p is very close to 1, more precisely if $p = 1 - c/\sqrt{n}$ then

$$\lim_{\infty \to n} \text{Prob}\,(G(n, p) \text{ is connected}) = e^{-c^4}.$$

It was P. Erdős and A. Rényi who first posed the problem of investigating the properties of random graphs (see [2], [3], [4], [5]). One of their most known result is as follows [2].

Let us consider the $\binom{n}{2}$ edges of the complete graph on n points as (completely) independent random variables. Let us denote by e the random variable which corresponds to the edge e. It takes the value 1 or 0 according to whether this edge belongs to our random graph or not. Further we suppose that $p = \text{Prob}\,(e=1) = \frac{\log n}{n} + \frac{c}{n}$ and $\text{Prob}\,(e=0) = 1-p$. This random graph (or random vector variable) will be denoted by $G_{n,p}$. Now as n tends to infinity and c is fixed,

$$\lim \text{Prob}\,(G_{n,p} \text{ is connected}) = e^{-e^{-c}} =$$

$$= \lim \text{Prob}\,(G_{n,p} \text{ has no isolated point}).$$

This theorem was generalized in many ways, e.g. [8] if

$$p = \frac{\log n}{n} + \frac{w(n)}{n}$$

then

$$\lim_{n\to\infty} \text{Prob}\,(G_{n,p} \text{ contains a Hamiltonian circuit}) = 1 \text{ or } 0$$

according to $\lim_{n\to\infty} w(n) = \infty$ resp. $-\infty$.

(The threshold function is not yet known.)

1980 *Mathematics Subject Classification.* Primary 05C40; Secondary 60C05.
Key words and phrases. Random graph, threshold function.

Another direction of generalization which is considered in this paper, is that one regards not a complete graph but some other one as the underlying graph. In this case only the edges of this underlying graph are chosen randomly. There are several theorems similar to the above cited one in this topic, e.g. for the complete bipartite graph or the n-cube [1], [6], [7]. However, the investigated underlying graphs usually have $O\left(\binom{|V(G)|}{2}\right)$ edges [9]. It is surprising that the case when G has only few edges has not been investigated. Questions of this type are of similar character and have important applications in physics.

In this paper we consider the following special case. Let $G(n)$ be the graph formed by an $n\times n$ square lattice. $|V(G)|=(n+1)^2$ $|E(G)|=2n(n+1)$. Choose the edges of $G(n)$ independently with probability p, and denote this random graph by $G(n,p)$.

THEOREM 1. *If* $p=1-\frac{c}{\sqrt{n}}$ *where* c *is fixed, then*

$$\lim_{n\to\infty} \operatorname{Prob}\big(G(n,p) \text{ is connected}\big) = e^{-c^4} =$$

$$= \lim_{n\to\infty} \operatorname{Prob}\big(G(n,p) \text{ has no isolated point}\big).$$

THEOREM 2. *Let* p *be fixed not depending on* n. *Then*
a) *If* $p>2/3$ *then* $G(n,p)$ *contains a giant component and the*

$$\text{2-nd largest component of } G(n,p) < \left(\frac{\log n}{\log(1/3(1-p))}+w(n)\right)^2.$$

b) *If* $p<1/3$ *then*

$$\text{the largest component of } G(n,p) < C(p)\log n.$$

(The inequalities are meant in the sense that they hold true with probability tending to 1 as $n\to\infty$.)

The Theorems mean that $G(n,p)$ is highly nonconnected if p is small, but for large p $G(n,p)$ is nearly connected and it will be really connected if p is very close to 1. It seems to be difficult to determine at what value of p the threshold value lies, but it is very likely 1/2 (cf. [6]). The method of the proof of Theorem 2 yields that the constant 2/3 can be improved to 0.658.

PROOF of Theorem 1.

A cutset of $G(n)$ is said to be *connected* if its edges can be linearly ordered in such a way that consecutive edges are neighbouring. The edges are *neighbouring* if they belong to the same small square in $G(n)$ (see Fig. 3). It is easy to see that if $G(n,p)$ is not connected then there exists a cutset of $\overline{G(n,p)}$ so it has a connected cutset by its planarity.

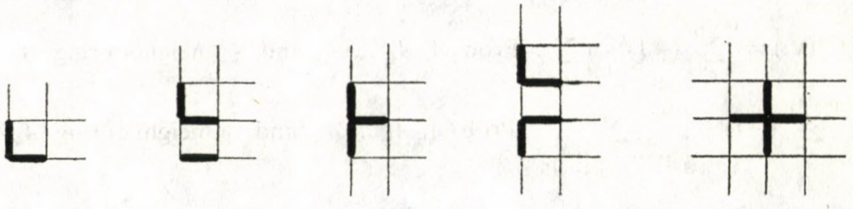
We shall see from the proof of Theorem 2 that

Prob (there exists a connected cut with k or more edges) $< 2n(n+1)3^k(1-p)^k/$ $/(1-3(1-p))$.

Put $k=5$, then

$$\text{Prob (there exists a connected cutset with} \geqq 5 \text{ edges)} <$$
$$< 2n(n+1)\big(3c/\sqrt{n}\big)^5/\big(1-3c/\sqrt{n}\big) = O\big(c^5/\sqrt{n}\big).$$

On the other hand any connected cutset with 2, 3 or 4 edges must be one of the following five types (Figure 1).



(i) (ii) (iii) (iv) (v)

Fig. 1

(i) *A connected cutset with 2 edges is in the corner.* There are 4 such cuts.

$$\text{Prob}\,((\text{i})) < 4(1-p)^2 = 4c^2/n.$$

(ii) *A connected cutset with 3 edges is in the corner.* There are 8 such cutsets.

$$\text{Prob}\,((\text{ii})) < 8(1-p)^3 = 8c^3/n\sqrt{n}.$$

(iii) *Isolated point on the boundary of $G(n)$.* There are $4(n-1)$ such cutsets.

$$\text{Prob}\,((\text{iii})) < 4(n-1)(1-p)^3 < 8c^3\sqrt{n}.$$

(iv) *A connected cutset with 4 edges on the boundary of $G(n)$.* There are $4(n-2)$ such cutsets.

$$\text{Prob}\,((\text{iv})) < 4(n-2)(1-p)^4 < 4c^4/n.$$

(v) *Isolated point inside $G(n)$.* There are $(n-2)^2$ such cutsets.

From these we get.

Prob ($G(n,$ p) has no isolated point)=
$=O(1)+$Prob (There are no isolated points and no connected cutsets with $\geqq 5$ edges)=
$=O(1)+$Prob (There are no cutsets with $\geqq 2$ edges)$+O(1)=$
$=O(1)+$Prob ($G(n,$ p) is connected).

We are finished with the proof of the equation

$$\lim_{n\to\infty} \text{Prob}\,(G(n,p) \text{ is connected}) = \lim_{n\to\infty} \text{Prob}\,(G(n,p) \text{ has no isolated point}).$$

Further

$$\text{Prob}\,(G(n,p) \text{ has no isolated point}) =$$
$$= \text{Prob (There is no isolated point inside the lattice } G(n,p)) - o(1) =$$
$$= \text{Prob}\,(\neg(\text{v})) - o(1).$$

By an application of the sieve method we determine Prob $(\neg(\mathrm{v}))$. Write A_i for the event that the i-th inner point is isolated in $G(n,p)$ $(1 \leq i \leq (n-1)^2)$. A_u and A_v are *neighbouring* if the u-th and v-th points are joined by an edge in $G(n)$.

$$\operatorname{Prob}(\neg(\mathrm{v})) = \operatorname{Prob}(\neg(\mathrm{v}) \text{ and there are no neighbouring } A_u A_v \text{ at all}).$$

(Clearly, the non-existence of isolated points implies the non-existence of pairs of neighbouring isolated points.)

$$\operatorname{Prob}(\neg(\mathrm{v})) = \sum_{k=0}^{(n-1)^2} (-1)^k \sum_{i_1, i_2, \ldots, i_k} \operatorname{Prob}(A_{i_1} A_{i_2} \ldots A_{i_k} \text{ and } \nexists^1 \text{ neighbouring } A_u A_v) =$$

$$= \sum_{k=0}^{(n-1)^2} (-1)^k \sum_{\substack{i_1, i_2, \ldots, i_k \text{ and} \\ \nexists A_{i_\alpha} A_{i_\beta} \text{ neighbouring}}} \operatorname{Prob}(A_{i_1} A_{i_2} \ldots A_{i_k} \text{ and } \nexists \text{ neighbouring } A_u A_v).$$

In the last sum it suffices to consider only those k with $0 \leq k < \log_2 n \cdot \max(6c^4, 3)$, because if k is more than this upper bound then

$$0 < \sum_{\substack{i_1, i_2, \ldots, i_k \text{ and} \\ \nexists \text{ neighbouring } A_{i_\alpha} A_{i_\beta}}} \operatorname{Prob}(A_{i_1} \ldots A_{i_k} \text{ and } \nexists \text{ neighbouring } A_u A_v) \leq$$

$$\leq \sum_{\substack{i_1, i_2, \ldots, i_k \text{ and} \\ \nexists \text{ neighbouring } A_{i_\alpha} A_{i_\beta}}} \operatorname{Prob}(A_{i_1} \ldots A_{i_k}) = \sum_{\substack{i_1, i_2, \ldots, i_k \text{ and} \\ \nexists \text{ neighbouring } A_{i_\alpha} A_{i_\beta}}} (1-p)^{4k} <$$

$$< \binom{(n-1)^2}{k} (1-p)^{4k} < \frac{n^{2k}}{k!} \frac{c^{4k}}{n^{2k}} < \left(\frac{c^4 e}{k}\right)^k < \frac{1}{n^3}.$$

So it suffices to sum over $k < (6c^4+3) \log_2 n$.

$$\operatorname{Prob}(\neg(\mathrm{v})) =$$

$$= O\left(\frac{1}{n}\right) + \sum_{k=0}^{c \log_2 n} (-1)^k \sum_{\substack{i_1, i_2, \ldots, i_k \\ \nexists A_{i_\alpha} A_{i_\beta} \text{ neighbouring}}} \operatorname{Prob}(A_{i_1} \ldots A_{i_k} \text{ and } \nexists \text{ neighbouring } A_u A_v) =$$

$$= o(1) + \sum_{k=1}^{c \log_2 n} (-1)^k \sum_{\substack{i_1, i_2, \ldots, i_k \text{ and} \\ \nexists A_{i_\alpha} A_{i_\beta} \text{ neighbouring}}} \operatorname{Prob}(A_{i_1} \ldots A_{i_k}) -$$

$$- \sum_{k=0}^{c \log_2 n} (-1)^k \sum_{\substack{i_1, i_2, \ldots, i_k \text{ and} \\ \nexists A_{i_\alpha} A_{i_\beta} \text{ neighbouring}}} \operatorname{Prob}(A_{i_1} \ldots A_{i_k} \text{ and } \exists \text{ neighbouring } A_u A_v) =$$

$$= o(1) + S_1 - S_2,$$

where S_1 and S_2 are the sums.

Now we show that S_1 is equal to $e^{-c^4} + o(1)$. To do so we have to count how many times one can choose k out of the A_i' s so that among the chosen k ones there

---

[1] The symbol $\nexists$ stands for $\neg\exists$.

are no neighbours. The number of such choices is at most $\binom{(n-1)^2}{k}$ and at least $\binom{(n-1)^2}{k}-4(k-1)\binom{(n-1)^2}{k}$. From this we get that

$$S_1 = 1-\frac{c^4}{1!}+\frac{c^8}{2!}-\ldots+O\left(\frac{\log n}{n^2}\right) = e^{-c^4}+o(1).$$

We divide every term of S_2 into two parts.

$$\sum_{\substack{i_1, i_2, \ldots, i_k \text{ and} \\ \nexists A_{i_\alpha} A_{i_\beta} \text{ neighbouring}}} \operatorname{Prob}(A_{i_1} \ldots A_{i_k} \text{ and } \exists \text{ neighbouring } A_u A_v) =$$

$$= \sum_{\substack{A_{i_1}, \ldots, A_{i_k} \\ \text{contain no neighbouring pair,} \\ \text{but they imply a neighbouring pair} \\ A_u A_{i_\alpha} \text{ (see Fig. 2)}}} \operatorname{Prob}(A_{i_1} \ldots A_{i_k} \text{ and } \exists \text{ neighbouring } A_u A_v) +$$

$$+ \sum_{\substack{A_{i_1}, \ldots, A_{i_k} \\ \text{contain no eighbouring pair, and} \\ \text{they do not imply a neighbouring pair.}}} \operatorname{Prob}(A_{i_1} \ldots A_{i_k} \text{ and } \exists \text{ neighbouring } A_u A_v).$$

If the position of $A_{i_1}; \ldots, A_{i_k}$ implies the existence of a neighbouring pair (see Fig. 2) (but A_{i_α} and A_{i_β} are not neighbouring) then

$$\operatorname{Prob}(A_{i_1} \ldots A_{i_k} \text{ and } \exists \text{ neighbouring } A_u A_v) =$$
$$= \operatorname{Prob}(A_{i_1} \ldots A_{i_k}) = (1-p)^{4k} = c^{4k}/n^{2k}.$$

But the number of $A_{i_1}, \ldots, A_{i_k}$ in such a position is at most $(n-1)^2\binom{(n-1)^2}{k-4}$.

If this is not the case then

$$\operatorname{Prob}(A_{i_1} \ldots A_{i_k} \text{ and } \exists \text{ neighbouring } A_u A_v) =$$
$$= \operatorname{Prob}(A_{i_1} \ldots A_{i_k} \text{ and } \exists \text{ neighbouring } A_u A_v \text{ and}$$
$$[\text{one of } A_{i_\alpha} \text{ and } A_u \text{ or } A_v \text{ are neighbouring}]) +$$
$$+ \operatorname{Prob}(A_{i_1} \ldots A_{i_k} \text{ and } \exists \text{ neighbouring } A_u A_v \text{ and } \neg$$
$$[\text{one of } A_{i_\alpha} \text{ and } A_u \text{ or } A_v \text{ are neighbouring}]) <$$
$$< (1-p)^{4k}(4k(1-p)+2n^2(1-p)^k).$$

Using these facts we get that S_2 is $o(1)$, consequently

$$\operatorname{Prob}(\neg(\mathrm{v})) = e^{-c^4}+o(1). \quad \text{Q. E. D.}$$

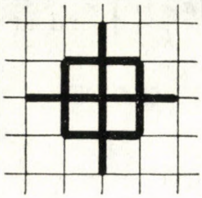


Fig. 2

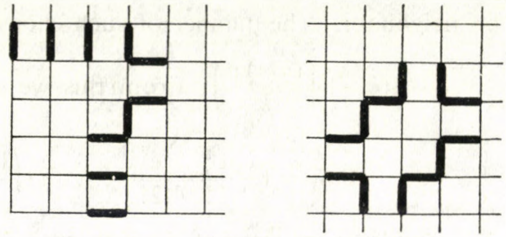


Fig. 3

PROOF of Theorem 2.

a) To begin with we show that the number of connected cuts of $G(n)$ with k edges is at most $n(n+1)3^{k-1}\cdot 4$ (see Fig. 3). This follows from the fact that on every connected cutset we can choose a starting edge and find (at most) two orderings of the edges of this cutset. Thus we can encode the form of the cut by 4 signs. We get a sequence of k signs with no consecutive identical signs. The number of such sequences is at most $4\cdot 3^{k-1}$. Hence

$$\operatorname{Prob}\big(G(n,p) \text{ has a connected cut with } \geqq k \text{ edges}\big) \leqq$$

$$\leqq \sum_{i=k}^{\infty} \operatorname{Prob}\big(G(n,p) \text{ has a connected cut with } i \text{ edges}\big) <$$

$$< \sum_{i=k}^{\infty} n(n+1)4\cdot 3^{k-1}(1-p)^i = n(n+1)4\cdot 3^{k-1}(1-p)^k \frac{1}{1-3(1-p)}.$$

Thus if $k > 2\log n/\log\big(1/3(1-p)\big)+w(n)$ then

$$\operatorname{Prob}\big(G(n,p) \text{ has a connected cut with } \geqq k \text{ edges}\big) = o(1).$$

In view of the fact that a cutset with k edges surrounds a set of at most $\frac{1}{4}k^2$ points:

$$\operatorname{Prob}\left(G(n,p) \text{ has a component with } \geqq \frac{k^2}{4} \text{ points}\right) = o(1).$$

b) Starting from an arbitrary point A let us go on the edges of $G(n)$, in every point P we can choose from at most three edges. Those new points connected with P are called the successors of P. As we can list the edges of $G(n)$ arbitrarily, so there is an appropriate branching process for the building of the component of $G(n,p)$ containing A (see [0]). Since

$$\text{E (number of successors of } P) \leqq 3p < 1,$$

this process is subcritical (Galton—Watson), i.e. it extincts exponentially, more precisely, there are positive constants c, ε such that

Prob (the cardinality of the component of $G(n,p)$ containing A is more than $k) < ce^{-\varepsilon}K$.

Thus

$$\text{Prob}\,(G(n,p) \text{ contains a component greater then } C(p)\log n) \leqq$$

$$\leqq \sum_A \text{Prob}\,(G(n,p) \text{ contains a component containing } A \text{ and greater than } C(p)\log n) <$$

$$< n^2 c e^{-c(p)\log n} = o(1).$$

REMARK. The methods presented here can be generalized to connected graphs for which the maximal degree is very small compared with the number of vertices, e.g. the lattice points of d dimensional cubes of size n ($n \to \infty$, d is fixed).

REFERENCES

[0] ATHREYA, K. B. and NEY, P. E., *Branching processes,* Die Grundlehren der mathematischen Wissenschaften, Band 196, Springer-Verlag, New York—Heidelberg, 1972. *MR* **51** #9242.

[1] BURTIN, YU. D., The probability of connectedness of a random subgraph of an n-dimensional cube, *Problemy Peredači Informacii* **13** (1977), 90—95 (in Russian). *MR* **58** #21833.

[2] ERDŐS, P. and RÉNYI, A., On the evolution cf random graphs, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5** (1960), 17—61. *MR* **23** #A2338.

[3] ERDŐS, P. and RÉNYI, A., Asymmetric graphs, *Acta Math. Acad. Sci. Hungar.* **14** (1963), 295—315. *MR* **27** #6258.

[4] ERDŐS, P. and RÉNYI, A., On the existence of a factor of degree one of a connected random graph, *Acta Math. Acad. Sci. Hungar.* **17** (1966), 359—368. *MR* **34** #85.

[5] ERDŐS, P. and SPENCER, J., *Probabilistic methods in combinatorics,* Probability and Mathematical Statistics, Vol. 17, Akadémiai Kiadó, Budapest, 1974; Academic Press, New York—London, 1974. *MR* **52** #2895.

[6] ERDŐS, P. and SPENCER, J. H., Evolution of the n-cube, *Comput. Math. Appl.* **5** (1979), 33—39. *MR* **80g**: 05054

[7] HART, S., A note on the edges of the n-cube, *Discrete Math.* **14** (1976), 157—163. *MR* **53** #161.

[8] KOMLÓS, J. and SZEMERÉDI, E., Limit distribution for the existence of Hamilton cycles in a random graph, *Discrete Math.* (to appear).

[9] MARGULIS, G. A., Probabilistic characteristics of graphs with large connectivity, *Problemy Peredači Informacii* **10** (1974), 101—108. *MR* **57** #12300.

(Received April 22, 1980)

MTA MATEMATIKAI KUTATÓ INTÉZETE
REÁLTANODA U. 13—15.
H—1053 BUDAPEST
HUNGARY

Thus

[illegible]

[illegible] greater than [illegible]

[illegible]

Remark. [illegible]

REFERENCES

[illegible]

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 427—430

ELEMENTS OF THE SOCLE OF A SEMI-SIMPLE BANACH ALGEBRA

by

ABDULLAH H. AL-MOAJIL

The purpose of this paper is to give different characterizations of the elements of $\operatorname{soc} A$ for a semi-simple Banach algebra A. In [1] J. C. ALEXANDER proved that $\operatorname{soc} A$ exists and $x \in \operatorname{soc} A$ if and only if the operator $a \to xax$ has finite rank. That is xAx is finite dimensional. On the other hand A. W. TULLO proved in [6] that the condition "$A = \operatorname{soc} A$" is equivalent to the finite dimensionality of A. Other conditions which are equivalent to finite dimensionality of A (and hence to the condition "$A = \operatorname{soc} A$") can be found in [2], [3] and [4]. These conditions include: (i) Every element of A is algebraic (i.e. satisfies a polynomial identity) ([2] and [3]), (ii) The spectrum of every element in A is finite ([2] and [4]) and (iii) Every closed right ideal in A contains an idempotent and A contains only a finite set of orthogonal idempotents ([2]).

Following the general pattern of proofs in [2] and [4] we will attempt to characterize elements of $\operatorname{soc} A$. We will show that imposing conditions similar to (i), (ii) and (iii) above on the right ideal xA gives equivalences to the condition that $\operatorname{soc} A$ exists and $x \in \operatorname{soc} A$. Our results also give an alternative proof to Alexander's result.

Let A be an algebra over the field of complex numbers $\mathbf{C}$. We mean by an idempotent in A a non-zero element $e \in A$ such that $e^2 = e$. Two idempotents e and f are orthogonal if $ef = fe = 0$. The idempotent e is said to be minimal if eAe is a division algebra. We mention that if A is normed then in this case, the Gelfand—Mazur theorem implies that eAe consists of scalar multiples of e. Moreover, in a semi-simple Banach algebra A, an idempotent e is minimal if and only if eA (Ae) is a minimal right (left) ideal in A. It is also well-known that if J is a minimal right ideal in A such that $J^2 \neq 0$ then there exists a minimal idempotent e such that $J = eA$ [5; 2.15]. If A contains minimal right (left) ideals then their sum is called the right (left) socle of A. If the right and left socles exist and are equal, then the resulting two-sided ideal is called the socle of A and is denoted by $\operatorname{soc} A$. The elements of $\operatorname{soc} A$ are precisely the finite sums of elements from minimal right (left) ideals. By previous remarks one can easily see that if (0) is the only ideal in A whose square is (0) then $\operatorname{soc} A$ exists if and only if A contains one-sided minimal ideals (right or left), [5, 2, 1.12].

Our terminology is consistent with that of [5], and algebras considered are over the field of complex numbers $\mathbf{C}$. We recall that if A is an algebra and $x \in A$ then the spectrum of x in A (denoted by $\sigma(x)$) is the set $\{\lambda \in \mathbf{C} : \lambda - x \text{ is not invertible}\}$ if A has an identity. If A does not possess an identity then $\sigma(x) = \{0\} \cup \{\lambda \in \mathbf{C} : \lambda \neq 0$ and $\frac{1}{\lambda}x$ is not quasiregular$\}$.

AMS (MOS) subject classifications (1970). Primary 46H05, 46H10; Secondary 16A32.
Key words and phrases. Minimal idempotent, minimal ideal, socle, spectrum, closed ideal.

Let A be a semi-simple normed algebra and suppose that e and f are minimal idempotents in A. If $x \in A$ and $exf \neq 0$ then $exfA \neq (0)$. Therefore, by minimality of eA we get $exfA = eA$ and hence $exfAf = eAf$. But $fAf = \mathbf{C}f$ by the Gelfand—Mazur theorem and therefore $eAf = \mathbf{C}exf$. Therefore eAf is either (0) or 1-dimensional. This observation appears as a lemma in [2] and other parts of the literature. It is used in the proofs of our theorems.

THEOREM 1. *If A is a semi-simple Banach algebra and $x \in A$ then the following conditions are equivalent:*

a) $\operatorname{soc} A$ *exists and* $x \in \operatorname{soc} A$.

b) *The subalgebra xAx is finite dimensional.*

c) *The right ideal xA is algebraic of bounded degree.*

d) *For every closed right ideal J of A either $J \cap xA = (0)$ or $J \cap xA$ contains a nonzero idempotent. Moreover, xA contains at most a finite number of pairwise orthogonal idempotents.*

e) *There exists orthogonal minimal idempotents $e_1, \dots, e_d$ in xA such that* $x = \left(\sum_{i=1}^{d} e_i\right) x$.

PROOF. If (a) holds, then since A is semi-simple, there exist minimal idempotents $e_1, \dots, e_n; f_1, \dots, f_m$ and elements $x_1, \dots, x_n; y_1, \dots, y_m$ such that $x = \Sigma e_i x_i = \Sigma y_j f_j$. Therefore $xAx \subset \sum_{i,j=1,1}^{n,m} e_i A f_j$. But $e_i A f_j$ is either (0) or 1-dimensional. Therefore, xAx is finite dimensional.

Now suppose that (b) holds. Let $d = \text{dimension}\,(xAx) + 2$. If $y \in A$, then $xyx, (xy)^2 x, \dots, (xy)^{d-1} x$ are linearly dependent, since xAx has dimension $d-2$. Therefore there exist scalars $\alpha_1, \dots, \alpha_{d-1}$ not all zeros such that $\sum_{i=1}^{d-1} \alpha_i (xy)^i x = 0$. Hence multiplying on the right by y we get $\sum_{i=2}^{d} \alpha_{i-1} (xy)^i = 0$. Thus xy satisfies a polynomial identity of degree d. That is xA is algebraic of the bounded degree d.

Next assume that (c) holds. Let d be a positive integer such that every element of xA satisfies a polynomial identity of degree d. Then, by the spectral mapping theorem [5; 1.6.10], every element of xA has at most d elements in its spectrum. (If $p(y) = 0$ then $p(\sigma(y)) = \sigma(p(y)) = \sigma(0) = \{0\}$. Therefore every $\lambda \in \sigma(y)$ is a root of p. If p is of degree d then p has at most d distinct roots.) Now, suppose that $J \cap xA \neq (0)$. Then, by semi-simplicity there is $y \in J \cap xA$ such that $\sigma(y) \neq 0$. Let B be the subalgebra generated by y. Then $B \subset J \cap xA$. Since y is algebraic, then B is finite dimensional, therefore, closed. That is B is a commutative Banach algebra, and there is a 1—1 correspondence between the set of maximal modular ideals in B and the set of non-zero elements in $\sigma(y)$ which is finite. Thus B has a finite number of maximal modular ideals. It follows that if R is the Jacobson radical of B, then B/R is the direct sum of finite number of copies of the field of complex numbers. Therefore, B/R contains idempotents which can be lifted to B [5; 2.3.9]. Thus B contains idempotents. But $B \subset J \cap xA$. Hence, $J \cap xA$ contains idempotents.

This proves the first assertion of (d). On the other hand if xA contains orthogonal idempotents $e_1, \dots, e_n$ with $n > d$, then setting $y = \sum_{k=1}^{n} \frac{1}{k} e_k$ we get $y \in xA$ and

$ye_k = \frac{1}{k} e_k$. That is $\left(\frac{1}{k} - y\right) e_k = 0$. This implies that $\frac{1}{k} \in \sigma(y)$. Therefore, $\sigma(y)$ contains at least n elements, namely $1, \frac{1}{2}, \ldots, \frac{1}{n}$. This contradicts the fact that $\sigma(y)$ contains at most d elements. Therefore, xA contains at most d orthogonal idempotents, which concludes (d).

Finally, let (d) hold. Let e_1 be an idempotent in xA. If e_1 is not minimal, then $e_1 A \subset xA$ is not a minimal closed ideal [5,2.1.10] and must contain an idempotent f by (d). Let $e_2 = fe_1 \in e_1 A e_1$. If e_2 is not minimal this process can continue to produce a sequence $\{e_i\}$ with $e_i \in e_j A e_j$ for $i \geqq j$. But if $f_i = e_{i+1} - e_i$ then the f_i's are orthogonal and hence $\{f_i\}$ is finite. Therefore for some n we must have e_n minimal. Therefore xA contains minimal idempotents. Choose $E = \{e_1, \ldots, e_d\}$ to be a maximal family of orthogonal minimal idempotent.

Now, let $f = e_1 + \ldots + e_d$. Then f is an idempotent and hence $(1-f)A$ is a closed ideal in A (note that no identity is needed to define $(1-f)A$). We have $(1-f)A \cap xA = (1-f)xA$ since $f \in xA$. If $(1-f)xA \neq (0)$, then by the assumption of (d) $(1-f)xA$ contains an idempotent $g \neq 0$. We have $g = (1-f)g$. Let $h = g(1-f)$. Then h is an idempotent. Moreover, h is orthogonal to e_i for $i = 1, \ldots, d$. But by the argument in the last paragraph hAh contains a minimal idempotent which is necessarily orthogonal to e_i for $i = 1, \ldots, d$. It follows by the maximality of E, that $h = 0$, which implies that $g = g^2 = hg = 0$ which is a contradiction. Hence, our assumption that $(1-f)xA \neq (0)$ is false and we have $(1-f)xA = (0)$. Therefore, by semi-simplicity, $(1-f)x = 0$, i.e. $x = fx = \left(\sum_{i=1}^{d} e_i\right) x$. This proves (e). Since it is trivial that (e) implies (a), this concludes the proof of the theorem.

Remark 1. Let H be a separable Hilbert space and $B(H)$ the algebra of bounded operators on H. Then $\operatorname{soc} B(H)$ consists of the operators on H with finite rank. We observe that condition (e) of Theorem 1 in this case says that if $T \in \operatorname{soc} B(H)$ then T can be represented as an infinite matrix with only a finite number of non-zero rows.

Remark 2. In part (b) of Theorem 1 the condition about xAx can be relaxed by replacing it with the condition that there is a positive integer n such that $(xA)^n x$ is finite dimensional. This is evident in the proof of "(b) implies (c)" as we can take d in this case to be dimension $(xA)^n x + 2$. Then for $y \in A$ we have $(xy)^n x, (xy)^{n+1} x, \ldots, (xy)^{n+d-2}$ are linearly dependent. The rest of the proof now proceeds as before. Nevertheless, Theorem 1 does not hold if the ideal xA was replaced by the subalgebra xAx in the statement of the theorem. For example if H is a separable Hilbert space with orthonormal basis $\{x_1, x_2, \ldots\}$ and $A = B(H)$. Let $T \in A$ be defined by $Tx_{2i-1} = x_{2i}$ and $Tx_{2i} = 0$. Then $T^2 = 0$, so TAT satisfies the identity $\lambda^2 = 0$. Therefore, TAT satisfies the condition stated about xA in Theorem 1 (c). Nevertheless, as T is not of finite rank it does not belong to $\operatorname{soc} A$.

We now use condition (c) of the theorem to show that $x \in \operatorname{soc} A$ implies that xA is closed. This helps to relax some of the conditions in Theorem 1.

Theorem 2. *If A is a semi-simple Banach algebra and $x \in A$ then the following conditions are equivalent:*

a) $\operatorname{soc} A$ *exists and* $x \in \operatorname{soc} A$.

b) *The right ideal xA is closed and algebraic.*
c) *The right ideal xA is closed and $\sigma(y)$ is finite for every $y \in xA$.*

PROOF. Let (a) hold. Then, by condition (c) of Theorem 1, there exist orthogonal minimal idempotents $e_1, \ldots, e_n$ in xA such that $x = \left(\sum_{i=1}^{n} e_i\right) x$. Now let (xy_n) be a sequence in xA converging to y. Then, $xy_n = (\Sigma\, e_i) xy_n$. Therefore, $(\Sigma\, e_i) y = (\Sigma\, e_i) \lim_n xy_n = \lim_n (\Sigma\, e_i) xy_n = \lim_n xy_n = y$. But $e_i \in xA$ and therefore $y = (\Sigma\, e_i) y \in xA$. This says that xA is closed. Condition (c) of Theorem 1 implies that xA is algebraic, and hence (b) follows.

If (b) holds, then as in the proof of Theorem 1, the spectral mapping theorem implies that every element of xA has a finite spectrum. Therefore (c) holds.

Next, let (c) hold. We will show that (d) of Theorem 1 holds which is equivalent to (a). Let J be a closed right ideal. If $J \cap xA \neq (0)$, then $J \cap xA$ is a non-zero closed right ideal, and by semi-simplicity it contains an element y with non-zero spectrum. Moreover, by condition (c), $\sigma(y)$ is finite. If B is the closed subalgebra generated by y, then $B \subset J \cap xA$ since $J \cap xA$ is closed. Now, the same argument as in the proof of "(c) implies (d)" in Theorem 1 applies and we conclude that $J \cap xA$ contains an idempotent. Moreover, if $\{e_1, e_2, \ldots\}$ is an infinite family of pairwise orthogonal idempotents in xA, we can choose $\{\lambda_i\}_{i=1}^{\infty}$ distinct numbers such that $|\lambda_i|\, \|e_i\| < 2^{-i}$. Then, since xA is closed, $\Sigma\, \lambda_i e_i$ converges to an element $y \in A$. Since $ye_i = \lambda_i e_i$ for each i we have $\lambda_i \in \sigma(y)$ for $i = 1, 2, \ldots$. This contradicts the assumption that $\sigma(y)$ is finite. Hence xA contains only a finite number of orthogonal idempotents. Therefore, (d) of Theorem 1 holds which concludes the proof.

We mention that in the example of Remark 2 following the proof of Theorem 1, TAT is closed (the reader can verify that). Therefore TAT satisfies condition (b) and (c) which are stated about xA in Theorem 2. Nevertheless, as seen before, $T \notin \operatorname{soc} B(H)$.

REFERENCES

[1] ALEXANDER, J. C., Compact Banach algebras, *Proc. London Math. Soc.* (3) **18** (1968), 1—18. *MR* **37** #4618.

[2] AL-MOAJIL, A. H., Characterization of finite dimensionality for semi-simple Banach algebras, *Manuscripta Math.* **33** (1981), 315—325.

[3] DIXON, P. G., Locally finite Banach algebras, *J. London Math. Soc.* (2) **8** (1974), 325—328. *MR* **50** #996.

[4] KAPLANSKY, I., Ring isomorphisms of Banach algebras, *Canadian J. Math.* **6** (1954), 374—381. *MR* **16**—49.

[5] RICKART, C. E., *General theory of Banach algebras,* The University Series in Higher Mathematics, D. van Nostrand Co., Inc., Princeton, N. J.—Toronto—London—New York, 1960. *MR* **22** # 5903.

[6] TULLO, A. W., Conditions on Banach algebras which imply finite dimensionality, *Proc. Edinburgh Math. Soc.* (2) **20** (1976), 1—5. *MR* **54** #3407.

(Received May 7, 1980)

DEPARTMENT OF MATHEMATICAL SCIENCES
UNIVERSITY OF PETROLEUM AND MINERALS
DHAHRAN
SAUDI ARABIA

Present address:
GULF ORGANIZATION FOR INDUSTRIAL CONSULTING
P. O. BOX 5114
DOHA
QATAR

GENERALIZATIONS OF PRÖSSDORF'S THEOREMS

by

L. LEINDLER

1. Let $f(x)$ be a continuous and 2π-periodic function and let

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \tag{1}$$

be its Fourier series. Denote $s_n(x) = s_n(f; x)$ the n-th partial sum of (1). Let $\lambda = \{\lambda_n\}$ be a monotone nondecreasing sequence of integers such that $\lambda_1 = 1$ and $\lambda_{n+1} - \lambda_n \leqq 1$. The mean

$$V_n(\lambda; x) = \frac{1}{\lambda_n} \sum_{\nu=n-\lambda_n}^{n-1} s_\nu(x) \qquad (n \geqq 1)$$

defines the n-th generalized de la Vallée Poussin mean of the sequence $\{s_n(x)\}$ generated by the sequence λ.

The usual supremum norm will be denoted by $\|\cdot\|_C$.

Let $\omega(\delta)$ be a nondecreasing continuous function on the interval $[0, 2\pi]$ having the properties:

$$\omega(0) = 0, \quad \omega(\delta_1 + \delta_2) \leqq \omega(\delta_1) + \omega(\delta_2).$$

Such functions will be called moduli of continuity. If

$$\omega(f; \delta) := \sup_{|x-y| \leqq \delta} |f(x) - f(y)|$$

denotes the modulus of continuity of $f \in C_{2\pi}$, then the class of function $f \in C_{2\pi}$ for which

$$\omega(f; \delta) \leqq A\omega(\delta)$$

will be denoted by H^ω, and

$$\|f\|_\omega := \|f\|_C + \sup_{x,y} |\Delta^\omega f(x, y)|,$$

where

$$\Delta^\omega f(x, y) = \frac{|f(x) - f(y)|}{\omega(|x-y|)}, \quad x \neq y,$$

defines a norm in H^ω. In the case $\omega(\delta) = \delta^\alpha$ $(0 < \alpha \leqq 1)$ we write, as usual, H^α, $\Delta^\alpha f(x, y)$ and $\|f\|_\alpha$ instead of H^{δ^α}, $\Delta^{\delta^\alpha} f(x, y)$ and $\|f\|_{\delta^\alpha}$, respectively.

1980 *Mathematics Subject Classification.* Primary 41A25.
Key words and phrases. Fourier series, class H^ω, modulus of continuity.

Recently Z. Stypiński [3], generalizing a theorem of S. Prössdorf [2], proved the following

Theorem A. *Let* $f \in H^\alpha (0<\alpha\leqq 1)$ *and* $0\leqq\beta<\alpha$. *Then*

$$\|V_n - f\|_\beta = \begin{cases} O(\lambda_n^{\beta-\alpha}), & 0<\alpha<1, \\ O(\lambda_n^{\beta-1}(1+\log\lambda_n)), & \alpha=1. \end{cases} \tag{2}$$

Theorem A, in the special case $\lambda_n = n$, reduces to a theorem of S. Prössdorf ([2], Theorem 2).

In the present paper first we extend the validity of Theorem A to the class H^ω and improve its second statement.

Let us define $\alpha=\alpha(\omega)$ as the infimum of those α' for which there exists a natural number $\mu=\mu(\alpha')$ such that

$$2^{\mu\alpha'}\omega(2^{-n-\mu}) > 2\omega(2^{-n}) \tag{3}$$

holds for all n. It is easy to see that for any modulus of continuity $\omega(\delta)$ the definition of α has sense and $0\leqq\alpha\leqq 1$, namely if $\alpha'>1$ then (3) holds obviously by the monotonicity of $2^n\omega(2^{-n})$, furthermore (3) with $\alpha'\leqq 0$ will never be satisfied. (We mention that a definition being equivalent to α was given by V. Totik [4], p. 152. He denoted this number by ω_0.)

It is also clear that if (3) holds for a certain α' and $0<\varepsilon<\dfrac{1}{2\mu}$ then

$$2^{2\mu(\alpha'-\varepsilon)}\omega(2^{-n-2\mu}) \geqq \frac{1}{2}2^{2\mu\alpha'}\omega(2^{-n-2\mu}) = \frac{1}{2}2^{\mu\alpha'}\big(2^{\mu\alpha'}\omega(2^{-n-\mu-\mu})\big) \geqq$$

$$\geqq \frac{1}{2}2^{\mu\alpha'}\big(2\omega(2^{-n-\mu})\big) \geqq \frac{1}{2}2^2\omega(2^{-n}) = 2\omega(2^{-n})$$

is also true, and this shows that (3) holds if and only if $\alpha'>\alpha$, i.e. (3) is not valid for $\alpha'=\alpha$.

Thus (3) determines for any modulus of continuity $\omega(\delta)$ a number α uniquely, and this enables us to use the notation $\omega_\alpha(\delta)$ for any modulus of continuity satisfying (3) for any $\alpha'>\alpha$.

Since for $\omega(\delta)=\delta^\alpha$ $(0<\alpha\leqq 1)$ (3) holds for any $\alpha'>\alpha$ we can write that $\delta^\alpha=\omega_\alpha^*(\delta)$. We remark that a modulus of continuity $\omega_\alpha(\delta)$ does not necessarily coincide with δ^α.

Now we define a certain subset of the moduli of continuity $\omega_\alpha(\delta)$ to be considered later on.

Let Ω_α denote the set of the moduli of continuity $\omega_\alpha(\delta)$ having the following additional property beside (3): For any natural number μ there exists a natural number $N(\mu)$ such that if $n>N(\mu)$ then

$$2^{\mu\alpha}\omega_\alpha(2^{-n-\mu}) \leqq 2\omega_\alpha(2^{-n}) \tag{4}$$

holds.

Our result reads:

THEOREM 1. *If* $0 \leqq \beta < \alpha \leqq 1$, $\omega_\beta(\delta) \in \Omega_\beta$, $\omega_\alpha(\delta) \in \Omega_\alpha$ *and* $f \in H^{\omega_\alpha}$ *then*

$$\|V_n - f\|_{\omega_\beta} = \begin{cases} O\left(\dfrac{\omega_\alpha(1/\lambda_n)}{\omega_\beta(1/\lambda_n)}\right), & \text{if } \alpha < 1 \text{ or } \beta > 0, \\ O\left(\dfrac{\omega_1(1/\lambda_n)}{\omega_0(1/\lambda_n)}(1+\log\lambda_n)\right), & \text{if } \beta = 0 \text{ and } \alpha = 1 \end{cases} \tag{5}$$

holds.

It is clear that Theorem 1 includes Theorem A and if $\beta > 0$, $\alpha = 1$ then it gives a better approximation order than (2) does.

In [2] Prössdorf also proved the following

THEOREM B. *If* $f \in H^\alpha$ $(0 < \alpha \leqq 1)$ *and* $0 \leqq \beta < \alpha$ *then*

$$\|s_n - f\|_\beta = O(n^{\beta-\alpha}\log n). \tag{6}$$

If $\lambda_n \equiv 1$ then $V_{n+1}(\lambda; x) \equiv s_n(x)$, but (6) gives a better result than (2) (or (5)) does in this special case. Thus in order to generalize (6) we have to refine our proof. Next we prove a generalization of Theorem B.

THEOREM 2. *If* $0 \leqq \beta < \alpha \leqq 1$, $\omega_\beta(\delta)$ *and* $\omega_\alpha(\delta)$ *belong to the class* Ω_β *and* Ω_α, *respectively, and* $f \in H^{\omega_\alpha}$ *then*

$$\|s_n - f\|_{\omega_\beta} = O\left(\frac{\omega_\alpha(1/n)}{\omega_\beta(1/n)}\log n\right) \tag{7}$$

holds.

2. To prove our theorems we require the following lemmas.

LEMMA 1. *If* $0 \leqq \beta < \alpha \leqq 1$, $\omega_\beta(\delta) \in \Omega_\beta$ *and* $\omega_\alpha(\delta) \in \Omega_\alpha$ *then for any* n

$$\sum_{k=1}^{n} \frac{\omega_\beta(2^{-k})}{\omega_\alpha(2^{-k})} \leqq K_{\beta,\alpha} \frac{\omega_\beta(2^{-n})}{\omega_\alpha(2^{-n})} \tag{2.1}$$

and

$$\sum_{k=n}^{\infty} \frac{\omega_\alpha(2^{-k})}{\omega_\beta(2^{-k})} \leqq K_{\beta,\alpha} \frac{\omega_\alpha(2^{-n})}{\omega_\beta(2^{-n})} \tag{2.2}$$

hold, where $K_{\beta,\alpha}$ *is a positive constant independent of* n.

PROOF. If $\beta' > \beta$ then there exists a natural number $\nu = \nu(\beta')$ such that

$$2^{\nu\beta'}\omega_\beta(2^{-n-\nu}) > 2\omega_\beta(2^{-n})$$

holds for any n. Hence it follows obviously that

$$2^{k\nu\beta'}\omega_\beta(2^{-n-\nu k}) > 2^k\omega_\beta(2^{-n}). \tag{2.3}$$

Since $\omega_\alpha(\delta) \in \Omega_\alpha$ thus for this ν there exists a natural number $N(\nu)$ such that if $m > N(\nu)$ then

$$2^{\nu\alpha}\omega_\alpha(2^{-m-\nu}) \leqq 2\omega_\alpha(2^{-m})$$

holds, and consequently for any k

$$2^{kv\alpha}\omega_\alpha(2^{-m-vk}) \leqq 2^k\omega_\alpha(2^{-m}). \tag{2.4}$$

If $\beta<\beta'<\alpha$ and $m>N(v(\beta'))$ then (2.3) and (2.4) imply that

$$\frac{\omega_\beta(2^{-n-vk})}{\omega_\alpha(2^{-m-vk})} > 2^{kv(\alpha-\beta')}\frac{\omega_\beta(2^{-n})}{\omega_\alpha(2^{-m})}, \tag{2.5}$$

whence

$$\frac{\omega_\beta(t)}{\omega_\alpha(t)} \to \infty \quad \text{as} \quad t\to 0 \tag{2.6}$$

follows.

Now we can prove (2.1). Let $\beta'=\dfrac{\alpha+\beta}{2}$, and let us choose $j=j(N(v))$ and $l=l(n)$ such that

$$v(j-1)\leqq N(v)<vj \quad \text{and} \quad v(l-1)\leqq n<vl$$

hold. Then, by (2.5), we can make the following estimations:

$$\sum_{k=1}^{n}\frac{\omega_\beta(2^{-k})}{\omega_\alpha(2^{-k})} \leqq \left(\sum_{k=1}^{vj}+\sum_{k=vj+1}^{vl}\right)\frac{\omega_\beta(2^{-k})}{\omega_\alpha(2^{-k})} = \left(\sum_{k=1}^{vj}+\sum_{i=j}^{l-1}\sum_{k=iv+1}^{(i+1)v}\right)\frac{\omega_\beta(2^{-k})}{\omega_\alpha(2^{-k})} \leqq$$

$$\leqq K(v)+\sum_{i=j}^{l-1} v\frac{\omega_\beta(2^{-iv})}{\omega_\alpha(2^{-(i+1)v})} \leqq K(v)+vK(\alpha-\beta)\frac{\omega_\beta(2^{-(l-1)v})}{\omega_\alpha(2^{-lv})} \leqq \tag{2.7}$$

$$\leqq K(v)+v2^{2v}K(\alpha-\beta)\frac{\omega_\beta(2^{-n})}{\omega_\alpha(2^{-n})}$$

where $K(v)$ and $K(\alpha-\beta)$ denote positive constants depending on the parameters denoted in the brackets.

Since $v=v(\beta')$ depends also just on $\dfrac{\alpha+\beta}{2}(=\beta')$; thus in respect to (2.6) (2.1) is proved by (2.7).

The estimations (2.2), by (2.5), can be proved similarly, we omit the details.

LEMMA 2. *If* $0\leqq\beta<\alpha\leqq 1$ *and* $f\in H^{\omega_\alpha}$, $\omega_\beta(\delta)\in\Omega_\beta$; $\omega_\alpha(\delta)\in\Omega_\alpha$, *furthermore*

$$\varphi_x(t)=f(x+t)+f(x-t)-2f(x),$$

then

$$|\varphi_x(t)-\varphi_y(t)| \leqq K\omega_\beta(|x-y|)\frac{\omega_\alpha(t)}{\omega_\beta(t)} \tag{2.8}$$

holds[1] *for any* x, y *and positive* t.

PROOF. It is clear that $f\in H^{\omega_\alpha}$ implies the following estimations:

$$|\varphi_x(t)-\varphi_y(t)| \leqq K_1\,\omega_\alpha(t) \qquad (t\geqq 0) \tag{2.9}$$

[1] $K, K_1, K_2, \ldots$ denote positive constants not necessarily the same at each occurrence.

and

$$|\varphi_x(t)-\varphi_y(t)| \leqq K_2\omega_\alpha(|x-y|). \tag{2.10}$$

The proof of (2.8) for $0<t\leqq|x-y|$ is trivial, namely then by (2.9)

$$|\varphi_x(t)-\varphi_y(t)| \leqq K_1\omega_\beta(t)\frac{\omega_\alpha(t)}{\omega_\beta(t)} \leqq K_1\omega_\beta(|x-y|)\frac{\omega_\alpha(t)}{\omega_\beta(t)}.$$

If $t>|x-y|$ then, using the result (2.1) of Lemma 1, we obtain, by a simple consideration on account of the monotonicity of ω_α and ω_β, that

$$\frac{\omega_\beta(t)}{\omega_\alpha(t)} \leqq K_3\frac{\omega_\beta(|x-y|)}{\omega_\alpha(|x-y|)},$$

whence and from (2.10) our statement (2.8) follows obviously.

LEMMA 3. *For any nonnegative sequence* $\{a_n\}$ *the inequality*

$$\sum_{n=1}^{m} a_n \leqq Ka_m \qquad (m=1,2,\ldots;\ K>0) \tag{2.11}$$

holds if and only if there exist a positive number c *and a natural number* μ *such that for any* n

$$a_{n+1} > ca_n \tag{2.12}$$

and

$$a_{n+\mu} \geqq 2a_n \tag{2.13}$$

are valid.

PROOF. An elementary calculation gives the proof. Namely, (2.12) follows from (2.11) if $c=1/K$. Furthermore, by (2.11), we have for any μ that

$$a_{n+\mu} \geqq \frac{1}{K}\sum_{i=n+1}^{n+\mu} a_i \geqq \frac{1}{K}\sum_{i=n+1}^{n+\mu}\frac{1}{K}a_n = \frac{\mu}{K^2}a_n$$

holds, and if $\mu\geqq 2K^2$ then this proves (2.13).

Conversely, if (2.12) and (2.13) hold, and $(n-1)\mu<m\leqq n\mu$, then

$$\sum_{i=1}^{m} a_i = \sum_{k=1}^{n-1}\sum_{i=\mu(k-1)+1}^{\mu k} a_i + \sum_{i=(n-1)\mu+1}^{m} a_i \leqq$$

$$\leqq 2\sum_{i=(n-2)\mu+1}^{(n-1)\mu} a_i + \sum_{i=(n-1)\mu+1}^{m} a_i \leqq 2\sum_{i=1}^{2\mu}\left(\frac{1}{c}\right)^i a_m = Ka_m$$

which proves (2.11).

3. PROOF of THEOREM 1. The proof runs on the same line as that of our first theorem in connection with the convergence of the means $V_n(\lambda; x)$ ([1], Theorem 2, see also [3]).

A standard computation gives that

$$\varrho_n(x) = V_n(\lambda; x) - f(x) = \frac{1}{\pi\lambda_n} \int_0^{\pi/2} \varphi_x(2t) K_n(t)\, dt,$$

where

$$K_n(t) = \frac{\sin \lambda_n t \sin (2n-\lambda_n)t}{\sin^2 t}.$$

Hence

$$|\varrho_n(x) - \varrho_n(y)| \leqq \frac{1}{\pi\lambda_n} \int_0^{\pi/2} |\varphi_x(2t) - \varphi_y(2t)| K_n(t)\, dt \equiv I.$$

Let us split the integral I into three parts:

$$I = \int_0^{\pi/2} = \int_0^{1/n} + \int_{1/n}^{1/\lambda_n} + \int_{1/\lambda_n}^{\pi/2} \equiv I_1 + I_2 + I_3.$$

These integrals, by (2.1), (2.2) and (2.8), can be estimated by elementary methods:

$$I_1 \leqq Kn \int_0^{1/n} |\varphi_x(2t) - \varphi_y(2t)|\, dt \leqq K_1 n \omega_\beta(|x-y|) \int_0^{1/n} \frac{\omega_\alpha(t)}{\omega_\beta(t)}\, dt \leqq$$

$$\leqq K_2 \omega_\beta(|x-y|) \frac{\omega_\alpha(1/n)}{\omega_\beta(1/n)}.$$

$$I_2 \leqq K_3 \int_{1/n}^{1/\lambda_n} \frac{|\varphi_x(2t) - \varphi_y(2t)|}{t}\, dt \leqq K_4 \omega_\beta(|x-y|) \int_{1/n}^{1/\lambda_n} \frac{\omega_\alpha(t)}{t\omega_\beta(t)}\, dt \leqq$$

$$\leqq K_4 \omega_\beta(|x-y|) \sum_{k=\lambda_n}^{n-1} \int_{1/k+1}^{1/k} \frac{\omega_\alpha(t)}{t\omega_\beta(t)}\, dt \leqq K_5 \omega_\beta(|x-y|) \sum_{k=\lambda_n}^{n} \frac{1}{k} \frac{\omega_\alpha(1/k)}{\omega_\beta(1/k)} \leqq$$

$$\leqq K_6 \omega_\beta(|x-y|) \sum_{m=\log \lambda_n}^{\log n} \frac{\omega_\alpha(2^{-m})}{\omega_\beta(2^{-m})} \leqq^2 K_7 \omega_\beta(|x-y|) \frac{\omega_\alpha(1/\lambda_n)}{\omega_\beta(1/\lambda_n)},$$

and finally

$$I_3 \leqq K_8 \frac{1}{\lambda_n} \int_{1/\lambda_n}^{\pi/2} \frac{|\varphi_x(2t) - \varphi_y(2t)|}{t^2}\, dt \leqq K_9 \frac{1}{\lambda_n} \int_{1/\lambda_n}^{\pi/2} \omega_\beta(|x-y|) \frac{\omega_\alpha(t)}{t^2 \omega_\beta(t)}\, dt \leqq$$

$$\leqq K_{10} \frac{1}{\lambda_n} \omega_\beta(|x-y|) \sum_{k=1}^{\lambda_n} \frac{\omega_\alpha(1/k)}{\omega_\beta(1/k)} \leqq K_{11} \frac{1}{\lambda_n} \omega_\beta(|x-y|) \sum_{m=0}^{\log \lambda_n} 2^m \frac{\omega_\alpha(2^{-m})}{\omega_\beta(2^{-m})}.$$

Here the last sum can be estimated easily if $\alpha < 1$; namely then

$$\sum_{m=0}^{\log \lambda_n} 2^m \frac{\omega_\alpha(2^{-m})}{\omega_\beta(2^{-m})} \leqq \frac{K_1}{\omega_\beta(1/\lambda_n)} \sum_{m=0}^{\log \lambda_n} 2^m \omega_\alpha(2^{-m}) \leqq K_2 \lambda_n \frac{\omega_\alpha(1/\lambda_n)}{\omega_\beta(1/\lambda_n)}$$

(see Lemma 3 and (3) with $\alpha' = 1$).

---

[2] $\sum_{m=a}^{b}$, where a and b are not integers, means a sum over all integers between a and b. The logarithm is used with basis 2.

If $\alpha=1$ and $\beta>0$ we obtain the same upper estimation for this sum but then (3) holds if and only if $\alpha'>1(=\alpha)$. Using the monotonicity of the sequence $2^{m(1+\beta/2)}\omega_\alpha(2^{-m})$ we get that

$$\sum_{m=0}^{\log\lambda_n} 2^m \frac{\omega_\alpha(2^{-m})}{\omega_\beta(2^{-m})} \leqq K_3\lambda_n^{1+\beta/2}\omega_\alpha(1/\lambda_n)\sum_{m=0}^{\log\lambda_n}\frac{1}{2^{m\beta/2}\omega_\beta(2^{-m})}$$

and if we show that

$$\sum_{m=0}^{\log\lambda_n} 2^{-m\beta/2}(\omega_\beta(2^{-m}))^{-1} \leqq K_4\lambda_n^{-\beta/2}(\omega_\beta(1/\lambda_n))^{-1} \tag{3.1}$$

holds, then our statement is verified.

To prove (3.1) first we show that there exists a natural number μ such that

$$2^{-(m+\mu)\beta/2}(\omega_\beta(2^{-m-\mu}))^{-1} > 2\cdot 2^{-m\beta/2}(\omega_\beta(2^{-m}))^{-1}. \tag{3.2}$$

Since $\omega_\beta(\delta)\in\Omega_\beta$ thus for any μ there exists an index $N(\mu)$ such that if $m>N(\mu)$ then

$$2^{\mu\beta}\,\omega_\beta(2^{-m-\mu}) \leqq 2\omega_\beta(2^{-m});$$

and if $\mu>4/\beta$ then hence we get that

$$2^{\mu\beta}\,\omega_\beta(2^{-m-\mu}) < \frac{1}{2}\,2^{\mu\beta/2}\,\omega_\beta(2^{-m}),$$

which implies (3.2).

A standard calculation similar to the proof of Lemma 3 shows that (3.2) implies (3.1).

In the case $\alpha=1$ and $\beta=0$ the sum investigated before does not exceed

$$K\lambda_n(1+\log\lambda_n)\frac{\omega_1(1/\lambda_n)}{\omega_0(1/\lambda_n)},$$

namely $\{2^m\omega_1(2^{-m})/\omega_0(2^{-m})\}$ is a nondecreasing sequence.

Consequently, collecting the partial results, we have that

$$I = |\varrho_n(x)-\varrho_n(y)| \leqq \begin{cases} K\omega_\beta(|x-y|)\dfrac{\omega_\alpha(1/\lambda_n)}{\omega_\beta(1/\lambda_n)} & \text{if}\quad \alpha<1 \quad\text{or}\quad \beta>0,\\[2ex] K\omega_\beta(|x-y|)\dfrac{\omega_\alpha(1/\lambda_n)}{\omega_\beta(1/\lambda_n)}(1+\log\lambda_n) & \text{if}\quad \alpha=1 \quad\text{and}\quad \beta=0,\end{cases}$$

whence (5) obviously follows.

Theorem 1 is proved.

PROOF of Theorem 2. In the proof we follow Prössdorf's method. Denote

$$s_n^*(x) = s_n(x)-\frac{1}{2}(a_n\cos nx+b_n\sin nx).$$

An elementary calculation shows that if $f\in H^{\omega_\alpha}$ then $a_n=O\left(\omega_\alpha\left(\frac{1}{n}\right)\right)$ and $b_n=O\left(\omega_\alpha\left(\frac{1}{n}\right)\right)$. Hence we get that

$$\|s_n-s_n^*\|_{\omega_\beta}=O\left(\omega_\alpha\left(\frac{1}{n}\right)\Big/\omega_\beta\left(\frac{1}{n}\right)\right), \tag{3.3}$$

namely if $|x-y|>\frac{1}{n}$ then

$$F_n^\beta(x,y):=\frac{|s_n(x)-s_n^*(x)-s_n(y)+s_n^*(y)|}{\omega_\beta(|x-y|)}\leqq\frac{K\omega_\alpha(1/n)}{\omega_\beta(|x-y|)}\leqq K\frac{\omega_\alpha(1/n)}{\omega_\beta(1/n)};$$

and if $|x-y|\leqq\frac{1}{n}$ then by Lagrange's theorem

$$F_n^\beta(x,y)\leqq K\frac{\omega_\alpha(1/n)}{\omega_\beta(|x-y|)}n|x-y|\leqq K_1\frac{\omega_\alpha(1/n)}{\omega_\beta(1/n)},$$

where we have used that

$$\frac{\omega_\beta(t)}{t}\uparrow \quad\text{as}\quad t\to 0.$$

In respect to (3.3) it is enough to consider the functions $R_n(x)=s_n^*(x)-f(x)$ and to estimate the difference $|R_n(x)-R_n(y)|$. Using a known estimation (see [5], p. 107) we obtain that

$$|R_n(x)-R_n(y)|\leqq J_1+J_2+J_3+J_4,$$

where with $\eta=\frac{\pi}{n}$

$$J_1=\frac{1}{\pi}\int_\eta^\pi t^{-1}|\varphi_x(t)-\varphi_y(t)+\varphi_y(t+\eta)-\varphi_x(t+\eta)|\,dt,$$

$$J_2=\eta\int_\eta^\pi t^{-2}|\varphi_x(t)-\varphi_y(t)|\,dt,$$

$$J_3=2\eta^{-1}\int_0^{2\eta}|\varphi_x(t)-\varphi_y(t)|\,dt$$

and

$$J_4=\int_{\pi-\eta}^{\pi}|\varphi_x(t)-\varphi_y(t)|\,dt.$$

Using (2.6) and the obvious inequality $\omega_\beta(1/n)\leqq Kn\omega_\alpha(1/n)$ J_4 can be estimated as follows:

$$J_4\leqq K_1\omega_\alpha(|x-y|)\frac{1}{n}\leqq K_2\omega_\beta(|x-y|)\frac{\omega_\alpha(1/n)}{\omega_\beta(1/n)}.$$

Following the same considerations as in the estimation of I_1 at the proof of Theorem 1 we obtain that

$$J_3\leqq K_3\omega_\beta(|x-y|)\frac{\omega_\alpha(1/n)}{\omega_\beta(1/n)}.$$

If we put $\lambda_n = n$ into the estimation obtained for I_3 we immediately get that

$$J_2 \leqq K_4 \omega_\beta(|x-y|) \frac{\omega_\alpha(1/n)}{\omega_\beta(1/n)} \log n.$$

Finally, an elementary consideration shows that

$$J_1 \leqq K \min\left(\omega_\alpha(|x-y|),\ \omega_\alpha\left(\frac{1}{n}\right)\right) \log n,$$

whence

$$J_1 \leqq K\omega_\beta(|x-y|) \frac{\omega_\alpha(1/n)}{\omega_\beta(1/n)} \log n$$

follows. Namely if $1/n \leqq |x-y|$ then

$$\omega_\alpha\left(\frac{1}{n}\right) = \omega_\beta(1/n) \frac{\omega_\alpha(1/n)}{\omega_\beta(1/n)} \leqq \omega_\beta(|x-y|) \frac{\omega_\alpha(1/n)}{\omega_\beta(1/n)},$$

and if $0 < |x-y| < 1/n$ then by (2.6)

$$\omega_\alpha(|x-y|) = \omega_\beta(|x-y|) \frac{\omega_\alpha(|x-y|)}{\omega_\beta(|x-y|)} \leqq K\omega_\beta(|x-y|) \frac{\omega_\alpha(1/n)}{\omega_\beta(1/n)}.$$

Now collecting our partial results we obtain that (7) holds and this completes the proof.

REFERENCES

[1] LEINDLER, L., On summability of Fourier series, *Acta Sci. Math. (Szeged)* **29** (1968), 147—162. *MR* **39** #707.

[2] PRÖSSDORF, S., Zur Konvergenz der Fourierreihen hölderstetiger Funktionen, *Math. Nachr.* **69** (1975), 7—14. *MR* **52** #14812.

[3] STYPIŃSKI, Z., On a generalization of the theorem of Prössdorf, *Funct. Approx. Comment. Math.* **7** (1979), 101—104. *MR* **80k**: 42007.

[4] TOTIK, V., On the strong approximation of Fourier series, *Acta Math. Acad. Sci. Hungar.* **35** (1980), 151—172.

[5] ZYGMUND, A., *Trigonometrical series,* I, Izdat. Mir, Moscow, 1965 (in Russian). *MR* **31** #2554. [*Trigonometrical series,* Dover Publications, New York, 1955. *MR* **17**—361].

(Received May 12, 1980)

JÓZSEF ATTILA TUDOMÁNYEGYETEM
BOLYAI INTÉZETE
ARADI VÉRTANUK TERE 1.
H—6720 SZEGED
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 441—452

r-ZUGÄNGLICHE UNTERDECKUNGEN DER EBENE DURCH KONGRUENTE KREISE

von

G. BLIND UND R. BLIND

1. Eine Packung von offenen Einheitskreisen $\{K_i\}_{i\in N}$ im E^2 heißt *r-zugänglich,* wenn an jeden Kreis K_i aus beliebig großer Entfernung ein Kreis vom Radius r so bis zur Berührung herangeschoben werden kann, daß er in keiner Lage einen inneren Punkt mit einem Kreis K_i gemeinsam hat.

Das Problem der dichtesten r-zugänglichen Kreispackung wurde in [5] aufgeworfen und untersucht. In [1] wurde gezeigt, daß für die Unterdeckungs-Dichte $\bar{D}$ einer r-zugänglichen Kreispackung gilt

$$\bar{D} \leqq \frac{\pi}{\sqrt{4(1+r)^2-1}+\sqrt{3}} \quad \text{für} \quad r>0,$$

wobei ein Beispiel zeigt, daß diese Schranke nicht verbessert werden kann.

In der vorliegenden Arbeit wird ein erheblich einfacherer Beweis dieser Schranke angegeben. Benützt werden Methoden aus [2].

2. Es ist zweckmäßig, die Unterdeckungs-Dichte von allgemeineren als den r-zugänglichen Kreispackungen nach oben abzuschätzen: Im E^2 sei $\{K_i\}_{i\in N}$ eine Packung von Einheitskreisen (Systemkreisen) mit der folgenden Eigenschaft: Die Indexmenge **N** kann in zwei disjunkte Teilmengen I_1 und I_2 zerlegt werden mit

(1) An jeden Kreis K_i mit $i\in I_1$ kann aus beliebig großer Entfernung ein *Fahrkreis* vom Radius r so bis zur Berührung herangeschoben werden, daß er in jeder Lage disjunkt ist zu allen Systemkreisen $K_i (i\in \mathbf{N})$.

(2) Für jeden Kreis K_i mit $i\in I_2$ gilt, daß im Innern des zu ihm konzentrischen Kreises vom Radius $2+2r$ kein weiterer Systemkreismittelpunkt liegt.

Die r-zugänglichen Kreispackungen erhält man speziell für $I_2=\emptyset$. Weil die Unterdeckungs-Dichte der Packung $\{K_i\}_{i\in N}$ nach oben abgeschätzt werden soll, kann o.B.d.A. angenommen werden, daß es zu jedem Punkt $P\in E^2$ einen Systemkreismittelpunkt im Innern des Kreises um P mit Radius $2+2r$ gibt; denn sonst kann man der Packung einen Systemkreis mit Mittelpunkt P hinzufügen, ohne die Eigenschaften (1) und (2) zu verletzen.

3. Der Mittelpunkt eines Systemkreises K_i heiße O_i. Nach der Methode der Stützkreise (siehe z.B. [3] oder [4, S. 228], jeder Stützkreis hat nach **2** einen Radius

1980 *Mathematics Subject Classifications.* Primary 52A45.

Key words and phrases. Convex sets, packing and covering, packing of circles in the euclidean plane.

$<2+2r$) gibt es im E^2 ein Polygonnetz M' aus konvexen Polygonen mit den Ecken O_i, das E^2 schlicht und lückenlos überdeckt. Der Mittelpunkt eines Stützkreises ist der Mittellotschnittpunkt des zugehörigen M'-Polygons. Wenn also P_1 und P_2 zwei M'-Polygone mit gemeinsamer Seite $P_1 \cap P_2$ und Mittellotschnittpunkten Q_1 bzw. Q_2 sind, und wenn das Mittellot auf $P_1 \cap P_2$ in Richtung von P_1 nach P_2 orientiert ist, so liegt Q_1 vor Q_2 auf dem Mittellot.

Aus der Definition von M' folgt außerdem, daß M' dual ist zum durch $\{K_i\}_{i\in N}$ definierten Netz der Dirichletschen Zellen: Die Ecken der Dirichletschen Zellen sind genau die Mittelpunkte der Stützkreise, d.h. genau die Mittellotschnittpunkte der M'-Polygone; die Kanten der Dirichletschen Zellen verbinden genau die Mittellotschnittpunkte von zwei M'-Polygonen mit gemeinsamer Seite, und sie liegen auf den Mittelloten dieser Seiten.

Durch geeignete Zerlegung der Polygone von M' erhält man ein Dreiecksnetz M mit den Ecken O_i, das E^2 schlicht und lückenlos überdeckt. Der Mittellotschnittpunkt eines Dreiecks von M ist identisch mit dem Mittellotschnittpunkt des entsprechenden Polygons von M'.

Weil der Radius jedes Stützkreises $<2+2r$ ist, ist $2(2+2r)$ eine obere Schranke für die Länge der Kanten von M und $\pi(2+2r)^2$ eine obere Schranke für den Flächeninhalt eines M-Dreiecks.

4. Es sei $\mathcal{Q}$ die Menge aller Punkte Q mit $\overline{QO}_i \geqq 1+r$ für alle $i \in \mathrm{N}$. Die Teilmenge von $\mathcal{Q}$ aus allen Punkten, die als Mittelpunkte für einen Fahrkreis in Frage kommen, besteht aus Wegzusammenhangs-Komponenten $\mathcal{Q}_j$ $(j \in J)$, die alle nicht beschränkt sind.

Γ' sei der folgende Graph: Die Knoten von Γ' seien die in $\mathcal{Q}$ liegenden Mittellotschnittpunkte von M-Dreiecken; zwei Knoten von Γ' werden genau dann durch eine Kante verbunden, wenn die ihnen entsprechenden M-Dreiecke eine gemeinsame Seite haben und wenn die Verbindungsstrecke der beiden Knoten ganz in $\mathcal{Q}$ liegt. Die Knoten von Γ' sind also höchstens 3-wertig.

Zur genaueren Untersuchung von Γ' betrachten wir zunächst die Systemkreismittelpunkte O_i $(i \in I_2)$. Alle Punkte auf dem Rand des Kreises $K(O_i, 1+r)$ um O_i mit Radius $1+r$ liegen in $\mathcal{Q}$, wegen der Voraussetzung (2) über $\{K_i\}$. $K(O_i, 1+r)$ liegt deshalb ganz in der K_i enthaltenden Dirichletschen Zelle, so daß jeder ihrer Randpunkte in $\mathcal{Q}$ liegt. Daraus folgt: Die M-Dreiecke mit der Ecke O_i haben als Mittellotschnittpunkt einen Punkt aus $\mathcal{Q}$, und die Verbindungsstrecke der Mittellotschnittpunkte von zwei nebeneinanderliegenden solchen Dreiecken liegt ganz in $\mathcal{Q}$. Zu jedem O_i $(i \in I_2)$ enthält also Γ' einen Kreis, und Γ_i sei die Komponente von Γ', die diesen Kreis enthält.

Zu gegebenem $j \in J$ seien nun $Q_1, Q_2 \in \mathcal{Q}_j$ Mittellotschnittpunkte von M-Dreiecken. Wegen des Wegzusammenhangs von $\mathcal{Q}_j$ gibt es dann einen ganz in $\mathcal{Q}_j$ verlaufenden Weg W, der Q_1 und Q_2 verbindet. Wenn W nicht nur aus Kanten von Dirichletschen Zellen besteht, gibt es eine einen Kreis K_i enthaltende Dirichletsche Zelle D_i $(i \in \mathrm{N})$ so, daß Punkte von W im Innern von D_i liegen. Verändert man den Weg W, indem man $W \cap D_i$ ersetzt durch die Projektion von $W \cap D_i$ von O_i aus auf den Rand von D_i, so erhält man wieder einen Q_1 und Q_2 verbindenden Weg, der ganz in $\mathcal{Q}_j$ verläuft. Daher kann o.B.d.A. angenommen werden, daß W nur aus Kanten von Dirichletschen Zellen besteht. Der Untergraph Γ_j von Γ', der durch die in $\mathcal{Q}_j$ liegenden Knoten von Γ' erzeugt wird, ist also zusammenhängend. Dabei ist Γ_j Komponente von Γ'. Weil mit einem Punkt einer Dirichletschen Zelle auch eine Ecke der Dirichletschen Zelle zu $\mathcal{Q}_j$ gehört, und weil $\mathcal{Q}_j$ nicht beschränkt ist, ist Γ_j unendlich.

Nun betrachten wir die Systemkreismittelpunkte O_i $(i\in I_1)$. Weil an jeden Kreis K_i $(i\in I_1)$ ein Fahrkreis bis zur Berührung herangeschoben werden kann, gibt es auf dem Rand des Kreises $K(O_i, 1+r)$ um O_i mit Radius $1+r$ einen Punkt $Q\in\mathcal{Q}_j$ für ein geeignetes $j\in J$. Die K_i enthaltende Dirichletsche Zelle enthält Q, so daß eine ihrer Ecken in $\mathcal{Q}_j$ liegt. Deshalb gibt es ein M-Dreieck mit der Ecke O_i so, daß sein Mittellotschnittpunkt in $\mathcal{Q}_j$ liegt, also Knoten von Γ_j ist.

Es sei schließlich Γ der Graph, der durch Vereinigung aller Graphen Γ_j $(j\in J)$ und aller Graphen Γ_i $(i\in I_2)$ entsteht. Nach Konstruktion hat dann Γ die folgenden Eigenschaften:

(3) Die Knoten von Γ sind mindestens 1-wertig und höchstens 3-wertig.

(4) Jede Komponente von Γ ist auch Komponente von Γ'.

(5) Jede endliche Komponente von Γ enthält einen Kreis.

(6) Jeder Systemkreismittelpunkt O_i $(i\in\mathrm{N})$ ist Ecke eines solchen M-Dreiecks, dessen Mittellotschnittpunkt Knoten einer unendlichen Komponente von Γ ist, oder die Mittellotschnittpunkte aller O_i enthaltenden M-Dreiecke sind Knoten von Γ.

5. Es sei O ein beliebig, aber fest gewählter Ursprungspunkt der Ebene, und $K(O, R)$ sei der Kreis mit Mittelpunkt O und Radius R. Es sei $n=n(R)$ die Zahl der Systemkreise, die ganz innerhalb $K(O, R)$ liegen. Dann ist die Unterdeckungs-Dichte $\bar{D}$ der Kreispackung $\{K_i\}_{i\in N}$ gegeben durch

$$\bar{D}=\varlimsup_{R\to\infty}\frac{n(R)}{R^2}.$$

Weil eine obere Schranke für $\bar{D}$ angegeben werden soll, kann o.B.d.A. im folgenden R jeweils so groß angenommen werden, daß die durchgeführten Konstruktionen sinnvoll sind.

Alle M-Dreiecke, deren Ecken Mittelpunkte von ganz in $K(O, R)$ liegenden Systemkreisen sind, sollen den Bereich G_0 bilden. G_0 besitzt dann n Eckpunkte, und es sei k die Anzahl jener M-Kanten, die Randkanten von G_0 sind. Nach dem Eulerschen Polyedersatz besteht dann G_0 aus $2n-k-2$ M-Dreiecken.

Diejenigen Systemkreise, die die k Randecken von G_0 als Mittelpunkte haben, liegen ganz außerhalb des Kreises $K\big(O, R-(2(2+2r)+2)\big)$, weil nach **3** $2(2+2r)$ eine obere Schranke für die Kantenlänge von M-Dreiecken ist. Diese k Systemkreise liegen also ganz innerhalb des Kreisrings $K(O, R)\setminus K\big(O, R-(6+4r)\big)$ mit einem Flächeninhalt kleiner als $2\pi R(6+4r)$, so daß gilt

$$k<2R(6+4r). \tag{7}$$

Weil die Länge der Randkanten von G_0 unabhängig von R durch $2+2r$ beschränkt ist, liegen für genügend großes R auch die Randkanten von G_0 außerhalb des Kreises $K\big(O, R-(6+4r)\big)$, so daß $G_0\supset K\big(O, R-(6+4r)\big)$. Weil die Fläche eines

M-Dreiecks nach **2** kleiner als $\pi(2+2r)^2$ ist, gilt für die Anzahl der M-Dreiecke von G_0

$$2n-k-2 \geqq \frac{\pi(R-(6+4r))^2}{\pi(2+2r)^2}.$$

Wegen $n > n - \dfrac{k+2}{2}$ folgt daraus zusammen mit (7)

$$\frac{k}{n} < \frac{2R(6+4r)}{\frac{1}{2}\left(\frac{R-(6+4r)}{2+2r}\right)^2}. \tag{8}$$

6. Zu jedem M-Dreieck, das eine Randkante von G_0 als Seite hat und nicht in G_0 liegt, betrachten wir nun das gleichschenklige Dreieck mit derselben Randkante als Basis, mit demselben Mittellotschnittpunkt, und auf derselben Seite der Randkante gelegen. Alle diese k gleichschenkligen Dreiecke bilden zusammen mit G_0 einen Bereich G; dabei können Dreiecke von G gemeinsame innere Punkte haben.

Die Länge jeder Kante von G ist wie die von M-Dreiecken $\geqq 2$.

Es ist $G_0 \subset K(O, R)$, und die k Dreiecke von $G \setminus G_0$ haben alle einen Flächeninhalt $< \pi(2+2r)^2$. Deshalb gilt für den Flächeninhalt F aller G-Dreiecke

$$F < \pi R^2 + k\pi(2+2r)^2. \tag{9}$$

Wir teilen nun die Dreiecke von G in zwei Klassen ein: Ein Dreieck von G heiße *Dreieck 1. Art,* wenn sein Mittellotschnittpunkt Knoten von Γ ist, sonst heiße es *Dreieck 2. Art.* Zwei Dreiecke 1. Art heißen im folgenden *benachbart,* wenn die ihnen entsprechenden Knoten von Γ benachbart sind.

7. Der Flächeninhalt eines G-Dreiecks Δ läßt sich gemäß den Figuren 1 und 2 aus den Flächeninhalten von 3 rechtwinkligen Dreiecken berechnen, die alle den Mittellotschnittpunkt von Δ als Ecke haben. Der Flächeninhalt eines rechtwinkligen Dreiecks ist

$$F(\varepsilon, u) = \frac{1}{2} u^2 \operatorname{ctg} \varepsilon,$$

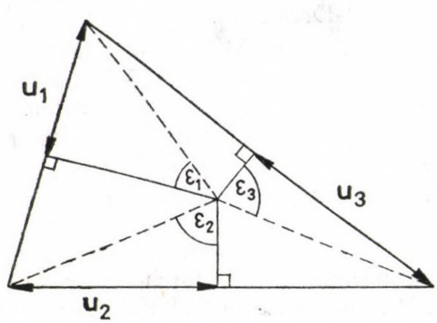


Fig. 1

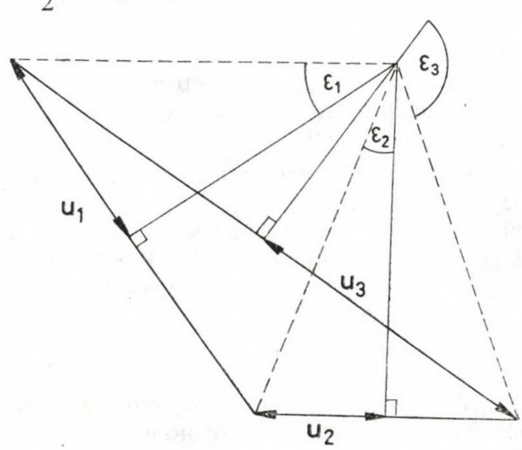


Fig. 2

wenn u eine Kathete des Dreiecks ist und ε der u gegenüberliegende Winkel. Weil nun

$$-F(\varepsilon, u) = F(\pi-\varepsilon, u), \tag{10}$$

gilt für den Flächeninhalt $F(\Delta)$ von Δ

$$F(\Delta) = 2\sum_{m=1}^{3} F(\varepsilon_m, u_m), \tag{11}$$

wobei man die Definitionen von ε_m und u_m den Figuren 1 und 2 entnimmt.

Ein so am Mittellotschnittpunkt von Δ auftretender Winkel ε_m heiße γ-Winkel, wenn Δ ein Dreieck 2. Art ist; wenn Δ ein Dreieck 1. Art ist, so heiße ε_m α-Winkel, wenn der eine Schenkel von ε_m Mittellot einer solchen Seite von Δ ist, die gemeinsame Seite von zwei benachbarten G-Dreiecken 1. Art ist, sonst heiße ε_m β-Winkel.

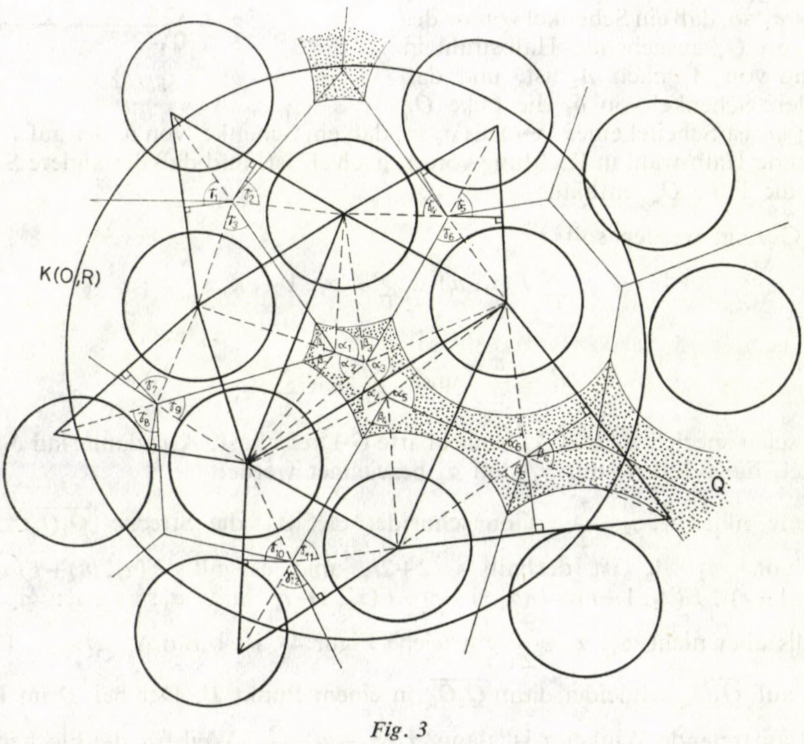


Fig. 3

Damit gilt für den Flächeninhalt F aller G-Dreiecke

$$\frac{1}{2}F = \sum_l F(\alpha_l, u_l) + \sum_p F(\beta_p, u_p) + \sum_q F(\gamma_q, u_q), \tag{12}$$

wobei die Summationen sich über alle Teildreiecke von G-Dreiecken mit α- bzw. β- bzw. γ-Winkeln erstrecken (siehe Fig. 3), und wobei

$$\sum_l \alpha_l + \sum_p \beta_p + \sum_q \gamma_q = (2n-k-2+k)\pi = (2n-2)\pi. \tag{13}$$

Im folgenden wird aus (12) und (13) eine Abschätzung für F abgeleitet. Zur Vorbereitung dient **8**.

8. Es seien Δ_1 und Δ_2 zwei G-Dreiecke mit einer gemeinsamen Seite $\Delta_1 \cap \Delta_2$. Q_1 bzw. Q_2 seien ihre Mittellotschnittpunkte, die auf dem Mittellot L zu $\Delta_1 \cap \Delta_2$ liegen. Q_1 und Q_2 liegen nach **3** auf L in der durch Δ_1 und Δ_2 gegebenen Reihenfolge. O_{i_0} sei eine gemeinsame Ecke von Δ_1 und Δ_2. Dann ist Q_1 Scheitel eines in **7** definierten Winkels δ_1 so, daß ein Schenkel von δ_1 der auf L von Q_1 ausgehende Halbstrahl in Richtung von Δ_1 nach Δ_2 ist, und daß der andere Schenkel von δ_1 die Ecke O_{i_0} enthält; Q_2 ist Scheitel eines Winkels δ_2 so, daß ein Schenkel von δ_2 der auf L von Q_2 ausgehende Halbstrahl in Richtung von Δ_2 nach Δ_1 ist, und daß der andere Schenkel von δ_2 die Ecke O_{i_0} enthält.

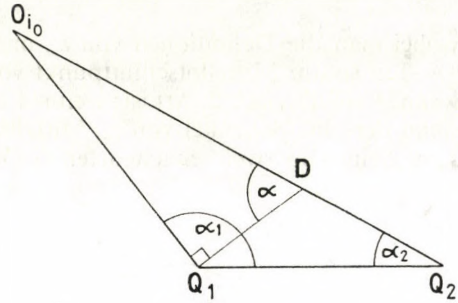


Fig. 4

9. Gezeigt werden soll

$$\sum_l F(\alpha_l, u_l) \geqq \sum_l F(\alpha_l', 1+r), \tag{14}$$

wobei

$$\alpha_l' \leqq \frac{\pi}{2} \quad \text{und} \quad \sum_l \alpha_l' = \sum_l \alpha_l.$$

Es seien speziell Δ_1 und Δ_2 benachbarte G-Dreiecke 1. Art; dann sind δ_1 und δ_2 α-Winkel, die o.B.d.A. mit α_1 und α_2 bezeichnet werden.

Wenn nun $\alpha_1, \alpha_2 \leqq \frac{\pi}{2}$, dann schneidet $\Delta_1 \cap \Delta_2$ die Strecke $\overline{Q_1Q_2} \subset \mathcal{Q}$. Die Länge von $\Delta_1 \cap \Delta_2$ ist deshalb $\geqq 2+2r$, und es gilt $F(\alpha_1, u_1)+F(\alpha_2, u_2) \geqq$ $\geqq F(\alpha_1, 1+r)+F(\alpha_2, 1+r) = F(\alpha_1', 1+r)+F(\alpha_2', 1+r)$ mit $\alpha_1' := \alpha_1$, $\alpha_2' := \alpha_2$.

Falls aber nicht $\alpha_1, \alpha_2 \leqq \frac{\pi}{2}$ gilt (siehe Figur 4), ist o.B.d.A. $\alpha_1 > \frac{\pi}{2}$. Das Lot von Q_1 auf $\overline{Q_1O_{i_0}}$ schneidet dann $\overline{Q_2O_{i_0}}$ in einem Punkt D. Der bei D im Dreieck $DQ_1O_{i_0}$ auftretende Winkel α ist dann $\alpha = \alpha_1+\alpha_2-\frac{\pi}{2}$. Weil für die Flächeninhalte der Dreiecke $DQ_1O_{i_0}$ und $Q_2Q_1O_{i_0}$ gilt $F(DQ_1O_{i_0}) \leqq F(Q_2Q_1O_{i_0})$ und $F(DQ_1O_{i_0}) =$ $= F(\alpha, \overline{Q_1O_{i_0}})$ und

$$F(Q_2Q_1O_{i_0}) = F(\alpha_1, u_1)+F(\alpha_2, u_2),$$

ist

$$F(\alpha_1, u_1)+F(\alpha_2, u_2) \geqq F(\alpha, \overline{Q_1O_{i_0}}) \geqq F(\alpha, 1+r)$$

wegen $Q_1 \in \mathcal{Q}$. Mit $\alpha_1' := \alpha_1 + \alpha_2 - \frac{\pi}{2}$, $\alpha_2' := \frac{\pi}{2}$ ist also auch in diesem Fall

$$F(\alpha_1, u_1) + F(\alpha_2, u_2) \geqq F(\alpha_1', 1+r) + F(\alpha_2', 1+r)$$

mit $\alpha_1', \alpha_2' \leqq \frac{\pi}{2}$ und $\alpha_1' + \alpha_2' = \alpha_1 + \alpha_2$.

Weil man durch Betrachten aller Paare von benachbarten G-Dreiecken 1. Art jeden α-Winkel genau einmal erhält, ist (14) gezeigt.

10. Gezeigt werden soll

$$\sum_p F(\beta_p, u_p) + \sum_q F(\gamma_q, u_q) \geqq \sum_p F(\beta_p, 1) + \sum_q F(\gamma_q, 1). \tag{15}$$

Es seien speziell Δ_1 und Δ_2 nicht benachbarte G-Dreiecke, dann sind δ_1 und δ_2 beide keine α-Winkel, sondern β- oder γ-Winkel. Die Länge von $\Delta_1 \cap \Delta_2$ ist $\geqq 2$. Wenn nun $\delta_1, \delta_2 \leqq \frac{\pi}{2}$, so gilt

$$F(\delta_1, u_1) + F(\delta_2, u_2) \geqq F(\delta_1, 1) + F(\delta_2, 1).$$

Falls aber nicht $\delta_1, \delta_2 \leqq \frac{\pi}{2}$ gilt, ist o.B.d.A. $\delta_1 > \frac{\pi}{2}$, und wegen $\delta_1 + \delta_2 \leqq \pi$ ist $\delta_2 \leqq \pi - \delta_1 \leqq \frac{\pi}{2}$. Die 1. Ableitung nach u von $F(\delta_1, u) + F(\delta_2, u)$ ist für $\delta_2 \leqq \leqq \pi - \delta_1 \leqq \frac{\pi}{2}$ nichtnegativ, d.h. $F(\delta_1, u) + F(\delta_2, u)$ ist monoton wachsend in u. Somit ist auch in diesem Fall

$$F(\delta_1, u_1) + F(\delta_2, u_2) \geqq F(\delta_1, 1) + F(\delta_2, 1),$$

wegen $u_1 = u_2$.

Durch Betrachten aller Paare von nicht benachbarten G-Dreiecken mit einer gemeinsamen Seite erhält man alle solchen β- und γ-Winkel genau einmal, bei denen der eine Schenkel nicht Mittellot einer Randkante von G ist. Nun sind die Randkanten von G die von den Basen verschiedenen Seiten der gleichschenkligen Dreiecke von $G \setminus G_0$; diejenigen β- und γ-Winkel, bei denen der eine Schenkel Mittellot einer Randkante von G ist, sind also $\leqq \frac{\pi}{2}$. Die Länge jeder Randkante von G ist nach **6** außerdem $\geqq 2$. Damit ist (15) gezeigt.

11. Für alle Winkel β_p gilt

$$\beta_p \leqq \frac{\pi}{2}. \tag{16}$$

Nach **10** gilt (16) für alle diejenigen β-Winkel, bei denen der eine Schenkel Mittellot einer Randkante von G ist. Wenn also (16) falsch ist, so gibt es zwei nicht benachbarte G-Dreiecke Δ_1 und Δ_2 mit gemeinsamer Seite $\Delta_1 \cap \Delta_2$ so, daß $\delta_1 > \frac{\pi}{2}$ und ein β-Winkel ist. Dann ist Δ_1 ein Dreieck 1. Art, also $Q_1 \in \mathcal{Q}$ und $\overline{Q_1 O_{i_0}} \geqq 1 + r$. Wegen

$\delta_1 > \frac{\pi}{2}$ ist dann für alle $Q \in \overline{Q_1 Q_2}$ auch $\overline{QO_{i_0}} \geqq 1+r$. Nun ist der Kreis $K(Q_1, \overline{Q_1 O_{i_0}})$ um Q_1 mit Radius $\overline{Q_1 O_{i_0}}$ Stützkreis, ebenso der Kreis $K(Q_2, \overline{Q_2 O_{i_0}})$, so daß im Innern dieser Kreise kein Systemkreismittelpunkt liegt. Für jedes $Q \in \overline{Q_1 Q_2}$ enthält deshalb auch der Kreis $K(Q, \overline{QO_{i_0}})$ keinen Systemkreismittelpunkt in seinem Inneren. Für alle $i \in \mathbf{N}$ ist somit $\overline{QO_i} \geqq \overline{QO_{i_0}} \geqq 1+r$ und folglich $\overline{Q_1 Q_2} \subset \mathcal{Q}$.

Nach (4) ist also auch Q_2 ein Knoten von Γ, der zu Q_1 benachbart ist, im Widerspruch dazu, daß Δ_1 und Δ_2 nicht benachbart waren.

12. Es folgen Abschätzungen für die Anzahl der G-Dreiecke 1. und 2. Art und der α- und β-Winkel.

Dazu betrachten wir die Einschränkung von Γ auf die Knoten, die Mittellotschnittpunkte von G-Dreiecken sind; dieser Graph werde mit $\Gamma|G$ bezeichnet.

G_0 hat $n-k$ innere Eckpunkte. Genau s von ihnen seien Ecken von solchen M-Dreiecken, deren Mittellotschnittpunkte Knoten von unendlichen Komponenten von Γ sind. Jede Komponente von $\Gamma|G$, die Teil einer unendlichen Komponente von Γ ist, besitzt mindestens einen Knoten, der Mittellotschnittpunkt eines Dreiecks von $G \setminus G_0$ ist; deshalb gibt es höchstens k solche Komponenten. Wenn eine Komponente von $\Gamma|G$ e_1 Knoten hat, und die durch die Knoten gegebenen Dreiecke insgesamt e_2 Ecken haben, so gilt $e_1 \geqq e_2 - 2$. Damit gibt es mindestens $s-2k$ Knoten von $\Gamma|G$, die auf unendlichen Komponenten von Γ liegen.

Für die restlichen $n-k-s$ inneren Eckpunkte von G_0 gilt nach (6), daß alle sie enthaltenden M-Dreiecke von 1. Art sind; daher besitzt $\Gamma|G$ noch mindestens $n-k-s$ weitere Knoten.

Insgesamt hat $\Gamma|G$ also mindestens $(s-2k)+(n-k-s)=n-3k$ Knoten. Damit gibt es $n-3k+t$ $(0 \leqq t \leqq n+3k-2)$ Dreiecke 1. Art und $n+3k-2-t$ Dreiecke 2. Art.

Am Mittellotschnittpunkt eines Dreiecks 1. Art treten so viele α-Winkel auf, wie die Wertigkeit des dem Mittellotschnittpunkt entsprechenden Knotens von $\Gamma|G$ ist. Nach (3) sind höchstens solche Knoten von $\Gamma|G$ 0-wertig, die Dreiecken von $G \setminus G_0$ entsprechen; $\Gamma|G$ hat also höchstens k 0-wertige Knoten.

Eine Komponente von $\Gamma|G$, deren Knoten nur G_0-Dreiecke entsprechen, ist eine endliche Komponente von Γ, die nach (5) einen Kreis enthält, so daß die Anzahl ihrer 1-wertigen Knoten kleiner oder gleich der Anzahl ihrer 3-wertigen Knoten ist.

Alle anderen Komponenten von $\Gamma|G$ besitzen mindestens einen Knoten, der Mittellotschnittpunkt eines Dreiecks von $G \setminus G_0$ ist, so daß es höchstens k solche Komponenten gibt; für jede solche Komponente ist die Anzahl der 1-wertigen Knoten kleiner oder gleich der um 2 vermehrten Anzahl der 3-wertigen Knoten.

Wenn also v_1 bzw. v_3 die Anzahl der 1- bzw. 3-wertigen Knoten von $\Gamma|G$ ist, so ist $v_1 \leqq v_3 + 2k$. Weil es nun $n-3k+t$ Dreiecke 1. Art gibt, gibt es also mindestens $2(n-3k+t-k)-2k=2(n-5k+t)$ α-Winkel und höchstens $n+k+t$ β-Winkel.

Gemäß (8) sei R so groß, daß $n-5k \geqq 0$ ist.

13. (12) bis (16) ergeben zusammen

$$\frac{1}{2} F \geqq \sum_{l} F(\alpha_l', 1+r) + \sum_{p} F(\beta_p, 1) + \sum_{q} F(\gamma_q, 1) \tag{17}$$

mit

$$\alpha_l' \leqq \frac{\pi}{2},\ \beta_p \leqq \frac{\pi}{2} \quad \text{und} \quad \sum_l \alpha_l' + \sum_p \beta_p + \sum_q \gamma_q = 2(n-2)\pi.$$

Im folgenden wird $\sum_q F(\gamma_q, 1)$ weiter abgeschätzt: Wir betrachten ein Dreieck 2. Art; an seinem Mittellotschnittpunkt treten drei γ-Winkel auf, o.B.d.A. die Winkel $\gamma_1, \gamma_2, \gamma_3$.

Wenn $\gamma_1, \gamma_2, \gamma_3 \leqq \frac{\pi}{2}$, so gilt mit Hilfe der Jensenschen Ungleichung

$$F(\gamma_1, 1) + F(\gamma_2, 1) + F(\gamma_3, 1) \geqq 3F\left(\frac{\gamma_1+\gamma_2+\gamma_3}{3}, 1\right) = 3F\left(\frac{\pi}{3}, 1\right).$$

Falls aber nicht $\gamma_1, \gamma_2, \gamma_3 \leqq \frac{\pi}{2}$, ist o.B.d.A. $\gamma_1 > \frac{\pi}{2}$. Aus $\gamma_1+\gamma_2+\gamma_3=\pi$ folgt dann $\gamma_3 \leqq \frac{\pi}{2}$ und $\gamma_2 \leqq \pi-\gamma_1 \leqq \frac{\pi}{2}$. Dann ist

$$F(\gamma_2, 1) + F(\gamma_1, 1) = F(\gamma_2, 1) - F(\pi-\gamma_1, 1) \qquad \text{(nach (10))}$$

$$\geqq F\left(\frac{\pi}{2} - ((\pi-\gamma_1)-\gamma_2), 1\right) - F\left(\frac{\pi}{2}, 1\right)$$

$\left(\text{weil } F(\varepsilon, 1) \text{ für } 0 < \varepsilon \leqq \frac{\pi}{2} \text{ monoton fallend und konvex ist}\right)$

$$= F\left(\gamma_1+\gamma_2-\frac{\pi}{2}, 1\right) + F\left(\frac{\pi}{2}, 1\right)$$

$$\geqq 2F\left(\frac{\gamma_1+\gamma_2}{2}, 1\right) \quad \text{(nach der Jensenschen Ungleichung).}$$

Damit ist mit Hilfe der Jensenschen Ungleichung auch in diesem Fall

$$F(\gamma_1, 1) + F(\gamma_2, 1) + F(\gamma_3, 1) \geqq 3F\left(\frac{\gamma_1+\gamma_2+\gamma_3}{3}, 1\right) = 3F\left(\frac{\pi}{3}, 1\right).$$

Weil es nun nach **12** genau $n+3k-2-t$ Dreiecke 2. Art gibt, ist also

$$\text{(18)} \qquad \sum_q F(\gamma_q, 1) \geqq (n+3k-t-2)3F\left(\frac{\pi}{3}, 1\right).$$

14. Es folgt eine Abschätzung für

$$\sum_l F(\alpha_l', 1+r) + \sum_p F(\beta_p, 1),$$

wobei

$$\alpha_l' \leqq \frac{\pi}{2}, \quad \beta_p \leqq \frac{\pi}{2}$$

und

$$\sum_l \alpha_l' + \sum_p \beta_p = (n-3k+t)\pi.$$

Nach **12** gibt es mindestens $2(n-5k+t)$ α-Winkel und höchstens $n+k+t$ β-Winkel. Deshalb gilt

$$\sum_l F(\alpha_l', 1+r)+\sum_p F(\beta_p, 1) = \sum_{l=1}^{2(n-5k+t)} F(\alpha_l', 1+r)+\sum_{l=2(n-5k+t)+1}^{\dots} F(\alpha_l', 1+r)+$$

$$+\sum_p F(\beta_p, 1) \geqq \sum_{l=1}^{2(n-5k+t)} F(\alpha_l', 1+r)+\sum_{p=1}^{n+k+t} F(\beta_p', 1),$$

wegen der Monotonie von $F(\varepsilon, u)$ in u und nach Umbezeichnung gewisser α-Winkel in β-Winkel. Wegen der Jensenschen Ungleichung gilt weiter

$$\sum_{l=1}^{2(n-5k+t)} F(\alpha_l', 1+r)+\sum_{p=1}^{n+k+t} F(\beta_p', 1)\geqq 2(n-5k+t)F(\alpha, 1+r)+(n+k+t)F(\beta, 1) \tag{19}$$

mit

$$\alpha = \left(\sum_{l=1}^{2(n-5k+t)} \alpha_l'\right)\frac{1}{2(n-5k+t)} \quad \text{und} \quad \beta = \left(\sum_{p=1}^{n+k+t} \beta_p'\right)\frac{1}{n+k+t}.$$

Daher ist

$$2(n-5k+t)\alpha+(n+k+t)\beta = \Sigma\, \alpha_l'+\Sigma\, \beta_p = (n-3k+t)\pi,$$

oder

$$\beta = \frac{(n-3k+t)\pi-2(n-5k+t)\alpha}{n+k+t}. \tag{20}$$

Nun ist die rechte Seite von (19) monoton fallend in α für $\alpha\leqq\beta$, und es ist β monoton fallend in α. Deshalb gilt (19) auch unter der Annahme $\alpha\geqq\beta$, was nach (20) äquivalent ist zu $\alpha\geqq\frac{\pi}{3}$. Damit ergibt sich schließlich

$$\sum_l F(\alpha_l', 1+r)+\sum_p F(\beta_p, 1) \geqq 2(n-5k+t)F(\alpha, 1+r)+(n+k+t)F(\beta, 1) \tag{21}$$

für ein α mit $\frac{\pi}{3}\leqq\alpha\leqq\frac{\pi}{2}$ und für

$$\beta = \frac{(n-3k+t)\pi-2(n-5k+t)\alpha}{n+k+t}.$$

15. (9), (17), (18) und (21) ergeben zusammen

$$\begin{gathered}\frac{1}{2}\left(\pi R^2+k\pi(2+2r)^2\right) \geqq \\ \geqq 2(n-5k+t)F(\alpha, 1+r)+(n+k+t)F(\beta, 1)+(n+3k-t-2)3F\left(\frac{\pi}{3}, 1\right),\end{gathered} \tag{22}$$

für ein α mit $\frac{\pi}{3}\leqq\alpha\leqq\frac{\pi}{2}$ und für

$$\beta = \frac{(n-3k+t)\pi-2(n-5k+t)\alpha}{n+k+t}.$$

Nun ist die Funktion auf der rechten Seite von (22) für alle α mit $\frac{\pi}{3} \leqq \alpha \leqq \frac{\pi}{2}$ monoton wachsend in t, so daß gilt

$$\begin{aligned} &\frac{1}{2}\left(\pi R^2 + k\pi(2+2r)^2\right) \geqq \\ (23)\qquad &\geqq 2(n-5k)F(\alpha, 1+r) + (n+k)F(\beta, 1) + (n+3k-2)3F\left(\frac{\pi}{3}, 1\right), \end{aligned}$$

für ein α mit $\frac{\pi}{3} \leqq \alpha \leqq \frac{\pi}{2}$ und für

$$\beta = \frac{(n-3k)\pi - 2(n-5k)\alpha}{n+k}.$$

Die rechte Seite von (23) ist eine Funktion von α, ihre erste Ableitung nach α ist für $\alpha = \frac{\pi}{3}$ kleiner als 0 und für $\alpha = \frac{\pi}{2}$ genau dann größer als 0, wenn $\frac{1}{1+r} > \sin\left(\frac{2k\pi}{n+k}\right)$, was nach (8) für genügend großes R erfüllt ist. Deshalb nimmt die rechte Seite von (23) in $\frac{\pi}{3} \leqq \alpha \leqq \frac{\pi}{2}$ ihr Minimum an für $\bar{\alpha}$ und $\bar{\beta} := \beta(\bar{\alpha})$ mit

$$(24)\qquad \sin\bar{\alpha} = (1+r)\sin\bar{\beta} = (1+r)\sin\frac{\left(1-3\frac{k}{n}\right)\pi - 2\left(1-5\frac{k}{n}\right)\bar{\alpha}}{1+\frac{k}{n}}.$$

16. Damit folgt aus (23) für die Dichte $\bar{D}$

$$(25)\qquad \bar{D} = \varlimsup_{R\to\infty} \frac{n}{R^2} \leqq \varlimsup_{R\to\infty} \frac{1}{R^2} \frac{\frac{1}{2}\left(\pi R^2 + k\pi(2+2r)^2\right) + 10kF(\bar{\alpha}, 1+r)}{2F(\bar{\alpha}, 1+r) + F(\bar{\beta}, 1) + 3F\left(\frac{\pi}{3}, 1\right)}.$$

Wegen $\frac{\pi}{3} \leqq \bar{\alpha} \leqq \frac{\pi}{2}$ ist

$$(26)\qquad F(\bar{\alpha}, 1+r) \leqq F\left(\frac{\pi}{3}, 1+r\right).$$

Nach (24) ist weiter $\bar{\alpha}$ eine implizite Funktion von $\frac{k}{n}$, die für $\frac{k}{n} = 0$ stetig ist. Wegen (8) ist $\lim_{R\to\infty} \frac{k}{n} = 0$, so daß also $\lim_{R\to\infty} \bar{\alpha}$ existiert, und nach (24) ist

$$(27)\qquad \bar{\bar{\alpha}} := \lim_{R\to\infty} \bar{\alpha} = \arccos\frac{1}{2(1+r)}, \quad \lim_{R\to\infty} \bar{\beta} = \pi - 2\bar{\bar{\alpha}}.$$

(25) zusammen mit (7), (26) und (27) ergibt nun

$$\bar{D} \leqq \frac{1}{2}\pi \frac{1}{2F(\bar{\bar{\alpha}}, 1+r)+F(\pi-2\bar{\bar{\alpha}}, 1)+3F\left(\frac{\pi}{3}, 1\right)},$$

oder wegen (27)

$$\bar{D} \leqq \frac{\pi}{\sqrt{4(1+r)^2-1}+\sqrt{3}}.$$

LITERATURVERZEICHNIS

[1] BLIND, G., r-zugängliche Unterdeckungen der Ebene durch kongruente Kreise, I und II, *J. Reine Angew. Math.* **288** (1976), 1—23 und **289** (1977), 1—29. *MR* **56** #1205a, 1205b.
[2] BLIND, G., ϱ-zugängliche Unterdeckungen der Sphäre durch kongruente Kreise, *Resultate der Math.* **4** (1981), 141—154.
[3] DELONE, B. N. and RYŠKOV, S. S., Extremal problems of the theory of positive quadratic forms, *Trudy Mat. Inst. Steklov.* **112** (1971), 203—223, 387 (in Russian). *MR* **49** #4939.
[4] FEJES TÓTH, L., *Reguläre Figuren,* Akadémiai Kiadó, Budapest, 1965. *MR* **30** #3408.
[5] HEPPES, A., On the densest packing of circles not blocking each other, *Studia Sci. Math. Hungar.* **2** (1967), 257—263. *MR* **35** #6037.

(Eingegangen am 15. Mai 1980)

UNIVERSITÄT STUTTGART
MATHEMATISCHES INSTITUT B
POSTFACH 80 11 40
PFAFFENWALDRING 57
D—7000 STUTTGART 80
WEST GERMANY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 453—454

STANDARD ELEMENTS OF ORDER TWO OF A LATTICE

by

IQBALUNNISA and S. AKILANDAM

G. GRÄTZER and E. T. SCHMIDT in [1] defined an element s of L to be standard element of order two if it is a standard element of a standard ideal S of L. They posed the problem "Do the standard elements of order two form a sublattice of L?" (cf. p. 74, Problem 18 in [1]). In this paper we give an answer to this problem in the negative. We show that the intersection of a finite number of standard elements of order two is a standard element of order two, while the sum of two standard elements of order two is not necessarily a standard element of order two of L. In other words, the set of standard elements of order two of L is closed with respect to intersection but it is not with respect to union.

We begin with

LEMMA 1. *If s_1 and t_1 are standard elements of order two of L, then $s_1 \wedge t_1$ is a standard element of order two of L.*

PROOF. Let s_1, t_1 be standard in S, T, where S and T are standard ideals of L. Let $S_1=(s_1]$ and $T_1=(t_1]$. Now S_1 is standard in S and $S \wedge T \subseteqq S$. Thus by Lemma 9 (page 38 of [1]) $S_1 \wedge T$ is standard in $S \wedge T$. Similarly $T_1 \wedge S$ is standard in $S \wedge T$. Therefore $(S_1 \wedge T) \wedge (T_1 \wedge S)$ is standard in $S \wedge T$, as the standard elements form a sublattice $S \wedge T$(by Theorem 3, p. 33 in [1]).

But $(S_1 \wedge T) \wedge (T_1 \wedge S) = S_1 \wedge T_1$, therefore $S_1 \wedge T_1$ is standard in $S \wedge T$. Further $S_1 \wedge T_1 = (s_1 \wedge t_1]$. Thus $s_1 \wedge t_1$ is a standard element of order two in L.

LEMMA 2. *If s_1 and t_1 are standard elements of order two in L, then $s_1 \vee t_1$ need not be a standard element of order two in L.*

PROOF. By a counterexample, consider the lattice L of Fig. 1 and the elements s, t, s_1 and t_1 as marked there in. Now s is standard in L because the congruence generated by $(s]$ is a standard congruence whose congruence classes are $(s]$ and $[b)$.

Now consider the congruence generated by $(t]$. Clearly, the congruence generated by $(t]$ is a standard congruence as its congruence classes are $(t]$, intervals (s, f), (e, s), (b, g) and $\{1\}$. Now consider the congruence generated by $(s_1]$ on $(s]$. This is a standard congruence on $(s]$ as its congruence classes are $(s_1]$, intervals (t_1, d), (t, f) $\{e\}$ and $\{s\}$. Clearly, the congruence generated by $(t_1]$ on $(t]$ is a standard congruence on $(t]$ as $(t]$ is a distributive lattice.

1980 *Mathematics Subject Classification.* Primary 06B10.
Key words and phrases. Lattice, standard ideal, standard element, standard congruence.

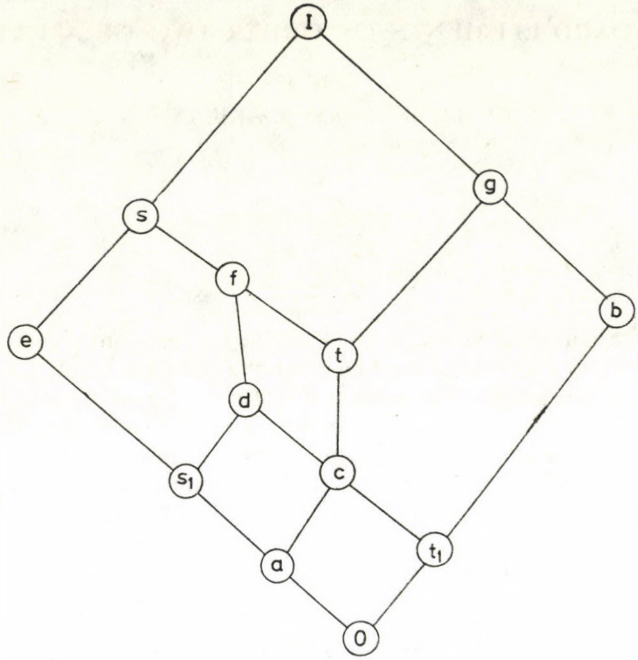


Fig. 1

Consider $d=s_1 \vee t_1$. We claim that d is not a standard element of order two in L. Firstly, we observe that d is not standard in L and $(s]$ for

$$f \wedge (d \vee e) = f \wedge s = f,$$

$$(f \wedge d) \vee (f \wedge e) = d \vee s_1 = d.$$

Therefore $f \wedge (d \vee e) \neq (f \wedge d) \vee (f \wedge e)$. The only possibility left is that d is standard in $(f]$ and f is standard in L. This fails as f is not standard in L. For

$$s \wedge (f \vee g) = s \wedge I = s,$$

$$(s \wedge f) \vee (s \wedge g) = f \vee t = f.$$

Therefore $s \wedge (f \vee g) \neq (s \wedge f) \vee (s \wedge g)$.

Thus we conclude the

THEOREM. *The set of standard elements of order two of L do not form a sublattice of L, in general.*

REFERENCE

GRÄTZER, G. and SCHMIDT, E. T., Standard ideals in lattices, *Acta Math. Acad. Sci. Hungar.* **12** (1961), 17—86. *MR* **24** # A3099.

(Received May 29, 1980)

RAMANUJAN INSTITUTE FOR
ADVANCED STUDY IN MATHEMATICS
UNIVERSITY OF MADRAS
IND—600 005 MADRAS
INDIA

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 455—459

A REMARK ON THE CONCENTRATION FUNCTION OF COMBINATORIAL NUMBER THEORY

by

JÓZSEF BECK

§ 1.

Throughout this paper $c_1, c_2, \ldots$ will denote large enough positive absolute constants. $[\alpha]$ denotes the integer part of α. The notation $|\ |$ will be used in two meanings, for a real number α let $|\alpha|$ denote the absolute value of α and for $T \subset \mathbf{R}$ let $|T|$ denote the Lebesgue measure of T, but it will not cause any misunderstanding.

Let $\mathscr{A}=\{a_1, a_2, \ldots, a_n\}$ be a sequence of positive integers and for a fixed integer x, let $Q(\mathscr{A}, x)$ denote the number of the solutions of

$$\sum_{k=1}^{n} \varepsilon_k a_k = x$$

where ε_k may have the value -1 or $+1$.

Let

$$Q_r = Q_r(\mathscr{A}) = \max_{-\infty<x<+\infty} \left\{ \sum_{j=0}^{r-1} Q(\mathscr{A}, x+j) \right\},$$

that is, Q_r denote the maximal number of sums which can fall into an open interval having length r. Q_r is called the concentration function of the sequence $\mathscr{A}$, its variable is the length r. For simplicity let $Q=Q(\mathscr{A})=Q_1(\mathscr{A})$. Clearly, $Q_r \leqq rQ$ and in what follows we shall investigate Q only.

For the basic results and methods on the concentration function see G. HALÁSZ [3].

In the following we shall assume that the elements of the sequence $\mathscr{A}$ are pairwise different, i.e.

(1) $$0 < a_1 < a_2 < \ldots < a_n.$$

In this case P. ERDŐS and L. MOSER proved that

$$Q < c_1 2^n n^{-3/2} (\log n)^{3/2},$$

and they conjectured

(2) $$Q < c_2 2^n n^{-3/2}.$$

A. SÁRKÖZY and E. SZEMERÉDI [1] proved that for $\varepsilon > 0$, $n > n_0(\varepsilon)$, (2) holds with $c_2 = (1+\varepsilon)\dfrac{8}{\sqrt{\pi}}$. J. L. NICOLAS [2] recently improved on the value of the constant

1980 *Mathematics Subject Classification.* Primary 10K99; Secondary 60G50.
Key words and phrases. Concentration function, characteristic function, combinatorics.

c_2, but the best possible value of c_2 remained to be unknown. Halász [3] investigated a multidimensional generalization of the problem above; his results involve (2) as a very simple corollary.

(2) is sharp as far as the order of magnitude is concerned, e.g. let $a_k=k$. A natural way of excluding this extremal case is to assume

$$a_{k+1}-a_k \neq a_{l+1}-a_l \quad \text{if} \quad k \neq l. \tag{3}$$

If we require that (3) holds, then we obtain an improvement on (2).

THEOREM 1. *If in addition to* (1) *also* (3) *holds, then*

$$Q < c_3 2^n n^{-5/2}.$$

This estimate is also sharp, the order of magnitude can be attained, e.g. if $a_k=k^2$.

We state also the case of a more scattered distribution.

THEOREM 2. *Let* $\mathcal{A}=\{a_1, a_2, \dots, a_n\}$ *be a sequence of positive integers such that*

$$\Delta a_k = a_{k+1}-a_k \geqq 1, \quad 1 \leqq k \leqq n-1;$$

$$\Delta^2 a_k = \Delta(\Delta a_{k+1}-\Delta a_k) = a_{k+2}-2a_{k+1}+a_k \geqq 1, \quad 1 \leqq k \leqq n-2;$$

$$\vdots \qquad\qquad \vdots$$

$$\Delta^l a_k = \Delta(\Delta^{l-1} a_{k+1}-\Delta^{l-1} a_k) = \sum_{j=0}^{l} (-1)^j \binom{l}{j} a_{k+l-j} \geqq 1, \quad 1 \leqq k \leqq n-l;$$

then

$$Q < c_4(l) 2^n n^{-l-\frac{1}{2}}.$$

Finally, we mention that Theorem 2 for $l=2$ is weaker than Theorem 1.

§ 2

Our proofs are based on the following inequality

$$Q(\mathcal{A}) \leqq \frac{2^{n-1}}{\pi} \int_{-\pi}^{\pi} \prod_{k=1}^{n} |\cos a_k t| \, dt.$$

Indeed, using the orthogonality of the functions e^{int}

$$Q(\mathcal{A}, x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \prod_{k=1}^{n} (e^{ia_k t}+e^{-ia_k t}) \right\} e^{-ixt} \, dt =$$

$$= \frac{2^{n-1}}{\pi} \int_{-\pi}^{\pi} \left\{ \prod_{k=1}^{n} \cos a_k t \right\} e^{-ixt} \, dt \leqq \frac{2^{n-1}}{\pi} \int_{-\pi}^{\pi} \prod_{k=1}^{n} |\cos a_k t| \, dt.$$

Applying the elementary inequality $\alpha \leqq e^{-(1-\alpha)}$ if $\alpha \geqq 0$ we obtain

$$|\cos a_k t| = \sqrt{\cos^2 a_k t} \leqq \exp\left\{-\frac{1}{2}(1-\cos^2 a_k t)\right\} = \exp\left\{-\frac{1}{4}(1-\cos 2a_k t)\right\},$$

multiplying for $k=1, 2, \ldots, n$,

$$\prod_{k=1}^{n} |\cos a_k t| \leqq \exp\left\{-\frac{1}{4} f(t)\right\},$$

where

$$f(t) = \sum_{k=1}^{n} (1-\cos 2a_k t),$$

therefore

$$Q \leqq \frac{2^{n-1}}{\pi} \int_{-\pi}^{\pi} \exp\left\{-\frac{1}{4} f(t)\right\} dt. \tag{4}$$

We shall estimate the integral of $\exp\left\{-\frac{1}{4} f(t)\right\}$ in Lebesgue's sense as we have learned in Halász's paper [3], that is,

$$\int_{-\pi}^{\pi} \exp\left\{-\frac{1}{4} f(t)\right\} dt \leqq \sum_{j=0}^{+\infty} |\{t\in(-\pi, \pi]\colon j \leqq f(t) \leqq j+1\}| e^{-j/4} \leqq$$
$$\leqq \sum_{j=0}^{+\infty} |\{t\in(-\pi, \pi]\colon f(t) \leqq j+1\}| e^{-j/4}. \tag{5}$$

In order to bound the size of sets $\{t\in(-\pi, \pi]\colon f(t)\leqq h\}$ we need the following lemma of Halász.

LEMMA (Halász). *Let $b_1, b_2, \ldots, b_m$ be a sequence of distinct positive integers, then*

$$\left|\left\{t\in(-\pi, \pi]\colon \sum_{k=1}^{m} (1-\cos b_k t) \leqq h\right\}\right| \leqq c_5 \frac{\sqrt{h}}{m^{3/2}}.$$

PROOF (Halász, see [3] and [4]). The following measure theoretic lemma forms the core of the proof:

Let S be a closed set on the real line, periodic with 2π, symmetric with respect to the origin and S_k be the set of points representable as $u = \sum_{i=1}^{k} u_i$ with $u_i\in S$. Then either S_k is the whole line or $|S_k|\geqq k|S|$, where $|S_k|$ and $|S|$ denote their Lebesgue measure mod 2π.

The proof is given, e.g. in [4].

For notational convenience let

$$T(h) = \left\{t\in(-\pi, \pi]\colon \sum_{k=1}^{m} (1-\cos b_k t) \leqq h\right\}.$$

From the elementary inequality

$$1-\cos \sum_{j=1}^{k} \alpha_j = 2\left(\sin \sum_{j=1}^{k} \alpha_j/2\right)^2 \leqq 2k \sum_{j=1}^{k} (\sin \alpha_j/2)^2 = k \sum_{j=1}^{k} (1-\cos \alpha_j) \tag{6}$$

immediately follows that $t \in T(k^2h)$ if it is representable as the sum of at most k numbers from $T(h)$, therefore by the measure theoretic lemma

$$(7) \qquad |T(k^2h)| \geqq \min(2\pi, k|T(h)|).$$

On the other hand, by Parseval's equality

$$|T(m/2)|\left(\frac{m}{2}\right)^2 \leqq \int_{-\pi}^{\pi}\left(\sum_{k=1}^{m}\cos b_k t\right)^2 dt = \pi m.$$

We thus have

$$|T(m/2)| \leqq \frac{4\pi}{m}.$$

Returning now to (7) we obtain

$$|T(h)| \leqq \left[\frac{2h}{m}\right]^{1/2}|T(m/2)| \leqq \left[\frac{2h}{m}\right]^{1/2}\frac{4\pi}{m} \qquad (m \geqq 2),$$

which was to be proved.

§ 3

PROOF of Theorem 1. Let $\Delta a_k = a_{k+1} - a_k$ and rearrange $\{\Delta a_k\}_{k=1}^{n-1}$ in increasing order

$$\Delta a_{k_1} < \Delta a_{k_2} < \ldots < \Delta a_{k_{n-1}}.$$

The proof is based on the fact that the numbers

$$a_{k_j} - \Delta a_{k_i}, \quad \frac{n}{2} < j \leqq n, \quad 1 \leqq i \leqq \frac{n}{2},$$

are pairwise different. Indeed,

$$a_{k_j} > a_{k_j} - \Delta a_{k_i} \geqq a_{k_j} - (a_{k_j} - a_{k_j-1}) = a_{k_j-1}.$$

Applying Halász's lemma to the sequence

$$\left\{2(a_{k_j} - \Delta a_{k_i}) \colon \frac{n}{2} < j \leqq n,\ 1 \leqq i \leqq \frac{n}{2}\right\}$$

we obtain

$$\left|\left\{t \in (-\pi, \pi] \colon \sum_{j=[n/2]+1}^{n}\sum_{i=1}^{[n/2]}(1 - \cos 2(a_{k_j} - \Delta a_{k_i})t) \leqq h\right\}\right| \leqq \frac{c_5\sqrt{h}}{[n/3]^3}.$$

By (6) used for $k=3$

$$\sum_{j=[n/2]+1}^{n}\sum_{i=1}^{[n/2]}(1 - \cos 2(a_{k_j} - \Delta a_{k_i})t) \leqq$$

$$\leqq 3\sum_{j=[n/2]+1}^{n}\sum_{i=1}^{[n/2]}\{(1 - \cos 2a_{k_j}t) + (1 - \cos 2a_{k_i+1}t) + (1 - \cos 2a_{k_i}t)\} \leqq$$

$$\leqq 9\left[\frac{n}{2}\right]\sum_{k=1}^{n}(1 - \cos 2a_k t).$$

From this follows

$$\left|\left\{t\in(-\pi,\pi]:\ \sum_{k=1}^{n}(1-\cos 2a_k t)\leqq\frac{h}{9[n/2]}\right\}\right|\leqq c_5\frac{\sqrt{h}}{[n/2]^3},$$

that is,

$$\left|\left\{t\in(-\pi,\pi]:\ \sum_{k=1}^{n}(1-\cos 2a_k t)\leqq y\right\}\right|\leqq\frac{3c_5\sqrt{y}}{[n/2]^{5/2}}.$$

Let us now return to (4) and (5)

$$Q\leqq\frac{2^{n-1}}{\pi}\int_{-\pi}^{\pi}\exp\left\{-\frac{1}{4}f(t)\right\}dt\leqq$$

$$\leqq\frac{2^{n-1}}{\pi}\sum_{y=0}^{+\infty}\left|\left\{t\in(-\pi,\pi]:\ \sum_{k=1}^{n}(1-\cos 2a_k t)\leqq y+1\right\}\right|e^{-y/4}\leqq$$

$$\leqq\frac{3c_5 2^{n-1}}{\pi[n/2]^{5/2}}\sum_{y=0}^{+\infty}(1+y)^{1/2}e^{-y/4}<c_2\frac{2^n}{n^{5/2}},$$

and Theorem 1 is proved.

PROOF of Theorem 2. Our basic observation is that the numbers

$$a_{k_0}-\Delta a_{k_1}-\Delta^2 a_{k_2}-\ldots-\Delta^l a_{k_l}$$

where $n\geqq k_0$ and $k_0-2\geqq k_1>k_2>\ldots>k_l\geqq 1$, are pairwise different. Indeed, in this case

$$a_{k_0}>a_{k_0}-\Delta a_{k_1}-\Delta^2 a_{k_2}-\ldots-\Delta^l a_{k_l}\geqq a_{k_0-1}.$$

Applying again Halász's lemma and (6) we can complete the proof similarly as above.

ACKNOWLEDGEMENT. I am especially grateful to Mr. G. Halász, who has simplified my proof.

REFERENCES

[1] SÁRKÖZY, A. and SZEMERÉDI, E., Über ein Problem von Erdős und Moser, *Acta Arith.* **11** (1965), 205—208. *MR* **32** #102.

[2] NICOLAS, J. L., Sur un problème de Erdős et Moser, *C. R. Acad. Sci. Paris Sér. A—B* **282** (1976), A9—A12. *MR* **53** #7934.

[3] HALÁSZ, G., Estimates for the concentration function of combinatorial number theory and probability, *Period. Math. Hungar.* **8** (1977), 197—211. *MR* **58** #13338.

[4] HALÁSZ, G., On the distribution of additive arithmetic functions, *Acta Arith.* **27** (1975), 143—152. *MR* **51** #5527.

(Received June 13, 1980)

MTA MATEMATIKAI KUTATÓ INTÉZETE
REÁLTANODA U. 13—15.
H—1053 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 14 (1979), 461—469

ON A GENERALIZATION OF THE GAME GO-MOKU, II

by

L. CSIRMAZ and ZS. NAGY

Abstract

Two players, I and II play the following game. They pick alternately the points of a set A until either all elements of A have been chosen or α moves have been made. The first and every limit move (if any) is I's turn. I wins if he picks all elements of some set of the winning family $\mathscr{F}\subset P(A)$, otherwise the winner is II. If the elements of $\mathscr{F}$ are finite and I has a winning strategy, then I has a winning strategy in finitely many moves. The cases when the elements $\mathscr{F}$ are countable are discussed in details. Various consistency results are given for undetermined and determined games. Several interesting problems are stated.

We study here another possible generalization of this well-known Oriental game. This part of our paper can be read independently from the previous one [1], but we do not repeat here the motivation behind our concepts.

1. Definitions

Our set theoretical notation will be standard. Ordinal numbers are denoted by α, β, etc. cardinal numbers by $\varkappa, \lambda$. If A and B are sets, then ${}^{A}B$ denotes the family of functions from A to B, and, by definition, $[A]^{\varkappa}=\{X\subset A: |X|=\varkappa\}$, $[A]^{<\varkappa}=\{X\subset A: |X|<\varkappa\}$. In this paper the inclusion $A\subset B$ allows the sets A and B to be equal.

The game we are going to deal with is denoted by $(A, \mathscr{F})^{\alpha}$, and consists of the board A, the family of winning sets $\mathscr{F}\subset P(A)$, and the ordinal α. The game is played by the players I and II as follows. The players pick elements of A alternately, every element can be picked at most once. I starts and every limit step (if any) is I's turn. The game ends if either all elements of A have been chosen or if α moves have been made.

The winner is I if he picked all elements of some $X\in\mathscr{F}$, otherwise the winner is II.

The game $(A, \mathscr{F})$ denotes the game $(A, \mathscr{F})^{|A|^{+}}$. In this paper under the word "game" we always mean game of this type.

The notion of strategy, that of a play according to a strategy can be found in [6]. A game is undetermined if neither I nor II has a winning strategy, abbreviated as WS.

1980 *Mathematics Subject Classification.* Primary 04A20; Secondary 54A35.
Key words and phrases. Infinite games, Martin's axiom.

2. Basic results

We start with some trivial observations.

PROPOSITION 2.1. *If I has a WS in* $(A, \mathcal{F})^\alpha$ *and* $A_1 \supset A, \mathcal{F}_1 \supset \mathcal{F}, \alpha_1 \geqq \alpha$ *then I has a WS in* $(A_1, \mathcal{F}_1)^{\alpha_1}$. *Moreover, if II has a WS in* $(A_1, \mathcal{F}_1)^{\alpha_1}$ *then he has a WS in* $(A, \mathcal{F})^\alpha$.

Now suppose that I has a WS in $G=(A, \mathcal{F})$. Then there exists a least $\alpha \in \mathrm{On}$ such that I still has a WS in $(A, \mathcal{F})^\alpha$, and a least cardinal $\varkappa$, such that I still has a WS in $(A, \mathcal{F}')$ for some $\mathcal{F}' \in [\mathcal{F}]^\varkappa$. Let us say that $\alpha = \mathrm{ord}(G)$, and $\varkappa = \mathrm{card}(G)$.

PROPOSITION 2.2. *For every ordinal* α *there is a game* G_1 *and for every cardinal* $\varkappa$ *there is a game* G_2 *such that* $\mathrm{ord}(G_1) = \alpha$ *and* $\mathrm{card}(G_2) = \varkappa$.

PROOF. We construct a game $(A, \mathcal{F})^\alpha$ such that $|\mathcal{F}| = |\alpha|$, I has a WS in it but II has a WS in

(i) $(A, \mathcal{F})^\beta$ if $\beta < \alpha$;

(ii) $(A, \mathcal{F}')^\alpha$ if $\mathcal{F}' \subset \mathcal{F}, \quad \mathcal{F}' \neq \mathcal{F}$.

Let $A = \{P_\beta, Q_\beta : \beta < \alpha\}$ and let the elements of $\mathcal{F}$ be $\{P_\gamma : \gamma \leqq \beta\} \cup \{Q_\beta\}$ for each $\beta < \alpha$ and the set $\{P_\gamma : \gamma < \alpha\}$. I can win only by picking the points P_γ in succession, so the game $(A, \mathcal{F})$ has all the properties required.

PROPOSITION 2.3. $\mathrm{card}(G)^+ \geqq \mathrm{ord}(G)$.

PROOF. For every $\beta < \mathrm{ord}(G)$ there is a play of length β which is not a win for I. Obviously, we may assume that II kills at least one winning set by his every move, i.e. $\mathrm{card}(G) \geqq |\beta|$. From this the proposition follows immediately.

REMARK. The proof shows that if $\mathrm{ord}(G)$ is not a cardinal number then $\mathrm{card}(G)^+ > \mathrm{ord}(G)$, and the game defined in the proof of 2.2 gives examples where $\mathrm{card}(G) < \mathrm{ord}(G)$.

PROBLEM. Is $\mathrm{card}(G)^+ > \mathrm{ord}(G)$ always true?

PROPOSITION 2.4. $2^{|\mathrm{ord}(G)|} \geqq \mathrm{card}(G)$.

PROOF. Let $\alpha = \mathrm{ord}(G)$, and $G^* = (A, \mathcal{F})^\alpha$. I has a WS in G^* so we may assume $\mathcal{F} \subset [A]^{\leqq |\alpha|}$. This strategy is function F from ${}^{\underset{\smile}{\alpha}}A = \bigcup \{{}^\beta A : \beta < \alpha\}$ to A, which gives the response of I to the series of moves $\langle b_\gamma : \gamma < \beta \rangle$ of II for every $\beta < \alpha$. Now there is a subset $B \subset A$, $|B| \leqq 2^{|\alpha|}$ which is closed under F, i.e. whenever $b_\gamma \in B$ for $\gamma < \beta$, then $F(\langle b_\gamma : \gamma < \beta \rangle) \in B$. This means that $F \restriction B$ is a I-WS in $(B, \mathcal{F} \cap P(B))^\alpha$. Therefore, by 2.1, I has a WS in $(A, \mathcal{F} \cap P(B))^\alpha$, i.e.

$$\mathrm{card}(G) \leqq |\mathcal{F} \cap P(B)| \leqq |{}^\alpha B| \leqq 2^{|\alpha|}.$$

REMARK. Let A be the points of a normal binary tree of height α, $\mathcal{F}$ be the family of the (maximal) branches. If $G = (A, \mathcal{F})$ then $\mathrm{ord}(G) = \alpha$ and $\mathrm{card}(G) = 2^{|\alpha|}$. Therefore the inequality in 2.4 is sharp.

3. Games with finite winning sets

In this section we discuss the games with finite winning sets.

THEOREM 3.1. *Let* $\mathcal{F}\subset[A]^{<\omega}$ *and suppose I has WS in* $G=(A, \mathcal{F})^{\alpha}$. *Then there is an* $\mathcal{F}_0\subset\mathcal{F}$, $|\mathcal{F}_0|<\omega$ *and* $n\in\omega$ *such that I wins the game* $(A, \mathcal{F}_0)^n$.

PROOF. Let S be a I-WS, and let $\beta<\alpha$ be an ordinal. The sequence

$$s = \langle a_\gamma\colon \gamma<\beta\rangle \in {}^{\underset{\sim}{\beta}}A$$

is a *partial play* according to S, if $a_\gamma=S(s\restriction\gamma)$ for every even ordinal γ. We identify S with the tree whose nodes are the partial plays (excluding the empty sequence), and $s<_S t$ iff s is a proper initial segment of t. We assume, that every branch in this tree has a highest node in some even level, and climbing on a branch I covers some element of $\mathcal{F}$ only at the last node (i.e., if I wins, the strategy ends).

Now let s be a node of S, $S\restriction s$ is the subtree above (and including) s. If the level of s is even then it induces a subgame $G\restriction s$ as follows. Let $s=\langle a_\gamma\colon\gamma<\beta\rangle$, β even,

$$A' = A-\{a_\gamma\colon \gamma<\beta\}$$

$$\mathcal{F}' = \{X-\{\alpha_\gamma\colon \gamma<\beta\}\colon X\in\mathcal{F} \text{ and } X\cap\{a_\gamma\colon \gamma<\beta \text{ and } \gamma \text{ is odd}\}=\emptyset\}.$$

Then $G\restriction s=(A', \mathcal{F}')^{\alpha-\beta}$. Obviously, $S\restriction s$ is a I-WS for $G\restriction s$, and if we replace $S\restriction s$ by any I-WS for $G\restriction s$, the resulting tree is a I-WS for the game G.

We shall prove the existence of a I-WS in which all of I's moves belong to the same finite subset of the board. The existence of this strategy implies the statement. The proof is by induction on the height of the tree S which will be denoted by h.

Case 1. $h=0$. The statement is true because I wins by his (unique) move.

Case 2. $0<h<\omega$. h is even, so $h\geqq 2$. Let $\{s_\gamma\}$ be the set of nodes at level 2. Because height $(S\restriction s_\gamma)\leqq h-2$, we may apply the induction assertion for the game $G\restriction s_\gamma$. Changing the subtrees $S\restriction s_\gamma$ to these strategies, we get that in every $S\restriction s_\gamma$ only finitely many points are engaged to I. In particular, let $B\subset A$, $|B|<\omega$ be the set of I-engaged points in $S\restriction s_0$. Now for every γ, if $s_\gamma(1)\notin B$ then replace $S\restriction s_\gamma$ by $S\restriction s_0$. The resulting tree S^* is good because B is finite and therefore only finitely many $S\restriction s_\gamma$ remain unchanged. S^* is not necessarily a I-WS because it may require I to pick the same element twice, but it can be turned into a strategy easily (Fig. 1).

Case 3. $h>\omega$ is a successor. Let $\beta<h$ be the maximal limit ordinal below h and let $s\in S$ be a node of height β. By Case 2 we may assume that $S\restriction s$ is a good strategy, i.e. there is a finite $\mathcal{F}_0\subset\mathcal{F}$ such that I covers one of them totally. $\cup\mathcal{F}_0$ is finite and there is a $\gamma<\beta$ such that the elements of this finite set picked by I during the first β moves, were picked before the γ-th move. Let $s\restriction\gamma=t$, and then we may replace $S\restriction t$ by $S\restriction s$ (Fig.2).

This transformation can be done for the remaining nodes of height β, and we get finally a tree S^* of height $\leqq\beta$ which is a I-WS, and the induction assertion can be applied.

Case 4. h is a limit. Let the node s be a *-node if the height of the tree $S\restriction s$ is h. For example the root of S is a *-node and the predecessors of a *-node are *-nodes,

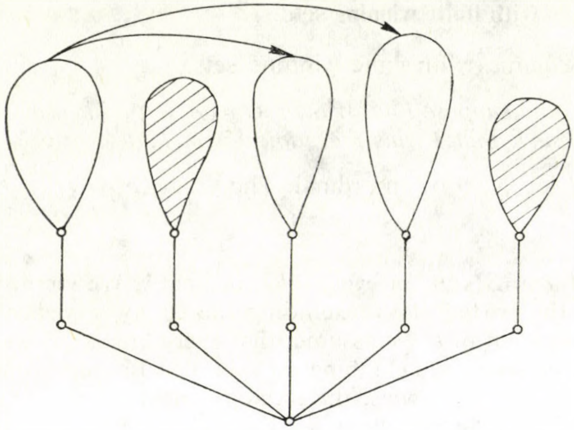


Fig. 1

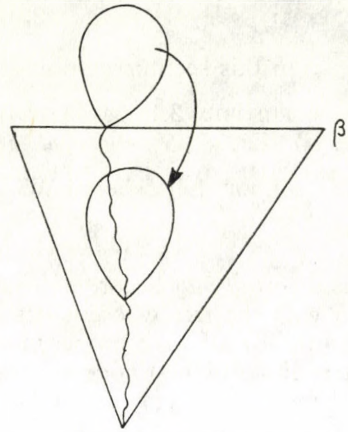


Fig. 2

too. Suppose that in some branch of S the limit s of *-nodes is not a *-node. Then height $(S{\restriction}s)<h$, and by the induction assertion, $S{\restriction}s$ can be supposed to be good. Then, as in Case 3, this $S{\restriction}s$ can be lowered in place of some *-node of this branch.

Similarly, if no successor of a *-node is a *-node, then we apply the induction assertion on these successors and just as in Case 2 we may replace $S{\restriction}s$ by a good subtree.

Therefore applying these steps sufficiently many times, we may achieve a I-WS tree S^* in which the limits of *-nodes are *-nodes and every *-node has a *-node successor. It means that if there are *-nodes in S^* then there is a cofinal branch of *-nodes. But this is impossible, because a I-WS has no cofinal branch (of limit lenght). Therefore the root of S^* is not a *-node, i.e. height $(S^*)<h$, and we are done.

THEOREM 3.2. *Let* $F\subset[A]^{<\omega}$, *then the game* $(A, \mathscr{F})^\alpha$ *is determined.*

PROOF. Suppose that I has no WS. Then II can make a move such that I still has no WS. This strategy is a WS for II. We have only to check that at limit moves I still has no WS. If he has, then, by the previous theorem, he has WS in some finite part of the game, therefore he has a WS *before* this move, a contradiction.

4. Games with countable boards

While we have a nice compactness theorem for finite winning sets, we cannot hope for one in general, as the following example shows.

EXAMPLE 4.1. Let $\mathscr{F}\subset[\omega]^\omega$, $|\mathscr{F}|=2^\omega$ be the (maximal) branches of a binary tree of height ω, such that $\bigcup\mathscr{F}=\omega$. Then I has a WS in $(\omega, \mathscr{F})$ but II wins the games $(\omega, \mathscr{F}')$ with $\mathscr{F}'\subsetneqq\mathscr{F}$, and $\big(A, P(A)\cap\mathscr{F}\big)$ with $A\subsetneqq\omega$.

There exist undetermined games. The following example is due to RALPH MCKENZIE.

THEOREM 4.2. *Let $U \subset P(\omega)$ be a non-trivial ultrafilter on ω. The game (ω, U) is undetermined.*

PROOF. Suppose first that II has a WS. At the end of the game either I or II (but not both) pick all the points of some element of U, therefore this strategy ensures II to cover an element of U. But also I can play by this strategy, and so he too covers an element of U, a contradiction.

Now assume that I has a WS, and let them play three instances of this game, see the figure. Let the first move of I be i_0. II may manage, that after the first ω moves every square is occupied and in every column (of three squares) except for the i_0-th one, at least one square belongs to him. If I has played by his strategy in each of the rows then the sets of squares occupied by I in the rows are elements of U, i.e. their intersection is infinite, a contradiction.

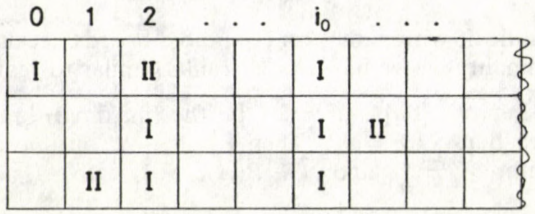


Fig. 3

In both examples the set of winning sets has cardinality 2^ω. This cannot be improved as the following theorem shows.

THEOREM 4.3. *Martin's axiom implies that if $\mathcal{F} \subset [\omega]^\omega$, $|\mathcal{F}| < 2^\omega$ then II has a WS in $(\omega, \mathcal{F})$.*

PROOF. In fact, we show that II may kill all of the winning sets in the first ω steps.

Let $\pi_k(x) = \min(x, k)$ for $x, k \in \omega$, and in general let

$$\pi_k(\langle x_0, \ldots, x_{i-1} \rangle) = \langle \pi_k(x_0), \ldots, \pi_k(x_{i-1}) \rangle.$$

Let ${}^{\underline{n}}\omega = \bigcup \{{}^i\omega \colon i < n\}$ and $f \colon {}^{\underline{n}}\omega \to \omega \; (n \in \omega)$ be a partial strategy for II, i.e. $f(\langle x_0, \ldots, x_{i-1} \rangle) \notin \{x_0, \ldots, x_{i-1}\}$, etc., such that $f = f \circ \pi_k$ for some $k \in \omega$. Obviously, the set P of these strategies is countable, therefore the partial ordering $f_1 \leqq f_2$ iff $f_1 \supset f_2$ for $f_1, f_2 \in P$ satisfies c.c.c. Given $F \in \mathcal{F}$, the subset

$$D_F = \{f \in P \colon (\forall g \in {}^\omega\omega)(\exists i \in \omega) f(g \restriction i) \in F\}$$

is dense, and by Martin's axiom there is a chain $G \subset P$ such that $G \cap D_F \neq \emptyset$ for every $F \in \mathcal{F}$. Now $\bigcup G$ is the strategy whose existence was stated.

On the other hand we have the following

THEOREM 4.4. Con $\bigl(ZFC + 2^\omega = \omega_2 +$ "*there is an $\mathcal{F} \subset [\omega]^\omega$, $|\mathcal{F}| = \omega_1$ such that II has no WS in the game $(\omega, \mathcal{F})$*"$\bigr)$.

PROOF. It is well-known that $2^\omega=\omega_2$ is consistent with the existence of a non-trivial ultrafilter generated by ω_1 elements [2]. Let $\mathcal{F}\subset[\omega]^\omega$, $|\mathcal{F}|=\omega_1$ be the set of generators, and suppose that $\mathcal{F}$ is closed under finite intersections. We claim that II has no WS in $(\omega, \mathcal{F})$. Indeed, otherwise after the game the set X of the points occupied by II intersects every element of $\mathcal{F}$. Similarly, I can also play by this strategy therefore the set $\omega-X$ of the points occupied by I intersects every element of $\mathcal{F}$ as well. But either X or $\omega-X$ is an element of the ultrafilter, i.e. contains some $F\in\mathcal{F}$ and then the other one cannot have a common point with F.

PROBLEM. Is it consistent that $2^\omega=\omega_2$ and for some $\mathcal{F}\subset[\omega]^\omega, |\mathcal{F}|=\omega_1$ I has a WS in $(\omega, \mathcal{F})$?

5. Large games

So far we have dealt with games on countable boards, let us take a step ahead. For boards of cardinality ω_1 we have some results similar to that of Section 4.

EXAMPLE 5.1. Let $\mathcal{F}\subset[\omega_1]^\omega$, $|\mathcal{F}|=2^\omega$ be the (maximal) branches of a normal Aronszajn tree such that $\bigcup\mathcal{F}=\omega_1$. Then I has a WS in $(\omega_1, \mathcal{F})$ but II wins the games $(\omega_1, \mathcal{F}')$ with $\mathcal{F}'\subsetneqq\mathcal{F}$ and $(A, P(A)\cap\mathcal{F})$ with $A\subsetneqq\omega_1$.

A version of Theorem 4.4 is true in this case.

THEOREM 5.2. Con $\left(ZFC+2^\omega=\omega_2+\text{“}\textit{there is an } \mathcal{F}\subset[\omega_1]^\omega, |\mathcal{F}|=\omega_1 \textit{ such that I wins the game } (\omega_1, \mathcal{F})\text{”}\right)$.

PROOF. The result Con $(ZFC+2^\omega=\omega_2+\clubsuit)$ is from [3], where $\clubsuit$ is the following combinatorial principle:

“There exists a sequence $\langle S_\alpha\colon \alpha<\omega_1, \alpha$ is limit$\rangle$ such that $\bigcup S_\alpha=\alpha$ and for every $X\in[\omega_1]^{\omega_1}$, there exists a limit $\alpha<\omega_1$ such that $S_\alpha\subset X$.”

The family $\mathcal{F}=\{S_\alpha\}$ evidently works.

PROBLEM. Is it true (in ZFC) that there is an $\mathcal{F}\subset[\omega_1]^\omega$, $|\mathcal{F}|=\omega_1$ such that $(\omega_1, \mathcal{F})$ is a win for I?

If the cardinality of the board and that of the family of the winning sets do not exceed $\varkappa$, and the length of the game is $<\varkappa^+$ then the strategies can be formulated in $V_\varkappa$, therefore we have

PROPOSITION 5.3. *Let $\varkappa$ be a weakly compact cardinal, $\mathcal{F}\subset[\varkappa]^\omega$, $|\mathcal{F}|=\varkappa$, $\alpha<\varkappa^+$. If I has a WS for $(\varkappa, \mathcal{F})^\alpha$ then there are $\mathcal{F}'\subset\mathcal{F}$, $|\mathcal{F}'|<\varkappa$ and $\lambda<\varkappa$ such that I wins $(\varkappa, \mathcal{F}')^\lambda$.*

The following construction is an unpublished result of A. HAJNAL.

THEOREM 5.4. *Let $V=L$, $\varkappa>\omega$, $\varkappa$ regular and not weakly compact. Then there exists an $\mathcal{F}\subset[\varkappa]^\omega$ such that I wins $(\varkappa, \mathcal{F})^\varkappa$, but*

(i) *II wins $(\varkappa, \mathcal{F}')$ if $\mathcal{F}'\subset\mathcal{F}, |\mathcal{F}'|<\varkappa$;*

(ii) *I has no WS in $(\varkappa, \mathcal{F})^\lambda$ if the regular cardinal $\lambda<\varkappa$.*

PROOF. (ii) follows from (i) and from the following lemma.

LEMMA 5.5. *Let $\lambda \geqq \omega$ be any regular cardinal, and suppose that I has WS in $(A, \mathscr{F})^{\lambda}$. Then there is a $B \subset A$, $|B| \leqq 2^{\underset{\smile}{\lambda}} = \sum_{\mu<\lambda} 2^{\mu}$ such that I still has WS in $(B, P(B) \cap \mathscr{F})^{\lambda}$.*

PROOF of the lemma. A I-WS is a function from ${}^{\underset{\smile}{\lambda}}A = \cup\{{}^{\alpha}A : \alpha < \lambda\}$ to A, choose B as the closure of any point of A with respect to this function.

As for (i) of Theorem 5.4, let $S \subset \varkappa$ be a stationary set consisting of ω-limits only such that for every limit $\xi < \varkappa$, $S \cap \xi$ is not stationary in ξ but S is stationary in $\varkappa$, and let $\langle X_\alpha : \alpha \in S \rangle$ be a $\diamondsuit_S$-sequence [4].

If we fix for every $\alpha \in S$ such that $\cup X_\alpha = \alpha$ a cofinal subset F_α of X_α of type ω, the family $\mathscr{F}$ of these F_α's will do.

Indeed, let I pick the elements of some closed unbounded set C of $\varkappa$. (He can do it by simply picking the smallest unpicked element above the set of previously picked ones.) Then, by $\diamondsuit_S$, the set $\{\alpha \in S : X_\alpha = C \cap S \cap \alpha\}$ is stationary in $\varkappa$, and so is its intersection with C, i.e.

$$A = \{\alpha \in S \cap C : \ X_\alpha = C \cap S \cap \alpha\}$$

is stationary, too. Now for some $\alpha \in A$, $\cup X_\alpha = \alpha$ otherwise there would be a regressive function on A which is impossible. But $\alpha \in S$, therefore α is an ω limit and $F_\alpha \subset X_\alpha \subset \subset C \cap S$. So F_α is covered by I, i.e. I wins the game.

Now let $\xi < \varkappa$ and $\mathscr{F}_\xi = \{F \in \mathscr{F} : F \subset \xi\}$. We claim that II wins the game $(\xi, \mathscr{F}_\xi)$, hence (i) follows. Instead of this we prove the following statement. Let $\alpha < \beta < \varkappa$, α and β be limit ordinals, $\alpha, \beta \notin S$, and $\mathscr{F}_{\alpha,\beta} = \{F \cap (\alpha, \beta) : F \in \mathscr{F}$ and $\alpha < \cup F < \beta\}$. Obviously, the elements of $\mathscr{F}_{\alpha,\beta}$ are countably infinite sets, and $\mathscr{F}_{0,\xi} = \mathscr{F}_\xi$. II wins the games $(\beta - \alpha, \mathscr{F}_{\alpha,\beta})$, this will be proven by induction on $\beta - \alpha$.

If $\beta - \alpha$ is countable, then $\mathscr{F}_{\alpha,\beta}$ is countable, too. (Different elements of $\mathscr{F}_{\alpha,\beta}$ have different suprema.) Let II sort the elements of $\mathscr{F}_{\alpha,\beta}$ in order type ω and at his i-th move pick an element of the i-th set.

If $\beta - \alpha \geqq \omega_1$ then $S \cap \beta$ is not stationary in β, i.e. there is a strictly increasing continuous sequence $\langle x_\xi : \xi \leqq \gamma \rangle$ such that $x_0 = \alpha$, $x_\gamma = \beta$, $x_{\xi+1} - x_\xi < \beta - \alpha$ for $\xi < \gamma$ and $x_\xi \notin S$. Now if $F \in \mathscr{F}_{\alpha,\beta}$ then $\cup F \in S$, i.e. $x_\xi < \cup F < x_{\xi+1}$ for some $\xi < \gamma$, therefore for some $F' \in \mathscr{F}_{x_\xi, x_{\xi+1}}$, $F' \subset F$ and $|F - F'| < \omega$. Because $x_{\xi+1} - x_\xi < \beta - \alpha$, II can play independently in each of these intervals by his previously defined strategy.

By this theorem, if $V = L$ then for every not weakly compact $\varkappa$ there exists an $\mathscr{F} \subset [\varkappa]^{\omega}$ such that card $((\varkappa, \mathscr{F})^{\varkappa}) = \varkappa$.

PROBLEM. Is it consistent that for every $\mathscr{F} \subset [\omega_2]^{\omega}$, card $((\omega_2, \mathscr{F})^{\omega_2}) < \omega_2$?

6. A topological game

Let (x, τ) be a Hausdorff topological space, $\tau \subset P(X)$ be the family of open sets in X. The game (X, τ) is the open-dense game. I wins if he covers an open subset of X, and II wins if he covers a dense subset of X. If X contains an isolated point then I has a WS, he has only to pick that point. However, I. JUHÁSZ made the following observation.

THEOREM 6.1. *Suppose X is a locally compact Hausdorff space with no isolated points. Then II has a WS in the game (X, τ).*

PROOF. Let $\mathcal{G}\subset\tau$ be the family of open sets G which have the following property. Every non-empty open subset of G has cardinality equal to that of G. $\mathcal{G}$ is evidently a π-base, take a maximal disjoint subfamily $\mathcal{G}^*$ of G. $\bigcup\mathcal{G}^*$ is dense in X therefore II can play independently in each element G of $\mathcal{G}^*$, because if II wins all of these games then he wins the whole game, too.

Now if $G\in\mathcal{G}^*$ then G is locally compact (endowed with the subspace topology), therefore the weight of G is at most $|G|$, see [5]. Let the base $\{G_\alpha: \alpha<|G|\}, \emptyset\neq G_\alpha\subset G$ witness this assertion, then $|G_\alpha|=|G|$ because $G\in\mathcal{G}^*\subset\mathcal{G}$. Then, at his α-th move (in G) II can choose an element of G_α, which ensures him the win.

On the other hand we have succeeded in proving the following theorems.

THEOREM 6.2. *Assume CH. There is a 0-dimensional Hausdorff topology τ on ω such that the game (ω, τ) is undetermined.*

PROOF. Let S_α^{I} and S_α^{II} for $\alpha<\omega_1$ be the possible strategies of I and II on ω respectively. We define the sets $B_\alpha, \tau_\alpha, X_\alpha$ and Y_α for $\alpha<\omega_1$ by induction on α such that

(i) $B_\alpha\subset[\omega]^\omega$, $|B_\alpha|=\omega$ is a base for a T_2-, 0-dimensional topology on ω, and $B_\beta\subset B_\gamma$ if $\beta<\gamma$.

(ii) τ_α is the topology induced by B_α.

(iii) $X_\alpha, Y_\alpha\subset\omega$, X_α is dense in τ_β if $\beta<\alpha$, and is open in τ_β if $\beta>\alpha$. Moreover $Y_\alpha\cap Z$ is infinite for every $Z\in B_\beta$ if $\alpha, \beta<\omega_1$.

(iv) If I plays by the strategy S_α^{I} then II has a counterplay such that he covers Y_α;

(v) if II plays by the strategy S_α^{II} then I has a counterplay such that he covers X_α.

By these conditions the topology $\tau=\bigcup\{\tau_\alpha: \alpha<\omega_1\}$ satisfies the requirements of the theorem.

Now let τ_{-1} be the topology of the dense linear ordering without endpoints on ω (i.e. the subspace topology of the rationals) and let B_{-1} be a countable base for it.

Suppose we have defined $B_\beta, \tau_\beta, X_\beta, Y_\beta$ for $\beta<\alpha$, and let $B=\bigcup\{B_\beta: \beta<\alpha\}$ and $\tau=\bigcup\{\tau_\beta: \beta<\alpha\}$. Observe that τ is generated by the countable base B, and there is no isolated point in τ. Therefore II has a counterplay against the α-th strategy S_α^{I} of I playing which he covers such a set Y_α which has the following property. For every $Z\in B, Z\cap Y_\alpha$ is infinite, in particular Y_α is dense in τ.

Similarly, I has a counterplay against the strategy S_α^{II} of II playing which he covers such a set X_α which has the following property. The sets $X_\alpha\cap Z\cap Y_\beta$ and $(\omega-X_\alpha)\cap Z\cap Y_\beta$ are infinite for each $Z\in B$ and $\beta\leqq\alpha$. Finally, let $B_\alpha=\{X_\alpha\cap Z, (\omega-X_\alpha)\cap Z: Z\in B\}$. The validity of the conditions (i)—(v) for α can be checked easily.

THEOREM 6.3. *There is a 0-dimensional Hausdorff topology τ on ω without isolated points such that I wins the game (ω, τ).*

PROOF. By Zorn's lemma there is a maximal 0-dimensional T_2-topology τ on A without isolated points where $|A|=\omega$. We claim that there is no $X\subset A$ such that both X and $A-X$ are dense in τ. Assuming the claim false, we find that $\tau \cup \{X, A-X\}$ constitutes a subbase for a 0-dimensional T_2-topology σ on A. Of course $\sigma \supset \tau$, $\sigma \neq \tau$ and τ is maximal, therefore σ contains an isolated point $p \in A$. It means that for some $G \in \tau$, either $G \cap X = \{p\}$ or $G \cap (A-X) = \{p\}$. Let $q \in G$ and $G' \in \tau$ be such that $q \in G'$ and $p \notin G'$. Then $\emptyset \neq G' \cap G \in \tau$ and either $G' \cap G \cap X$ or $G' \cap G \cap (A-X)$ is empty which contradicts to the denseness of X and $A-X$.

Let (A', τ') be a disjoint copy of (A, τ) and let I and II play on the topological sum of these spaces. By the previous remark, II cannot cover dense subsets in both of these spaces if I plays as follows. If II picked a point in A, I picks the same point in A', and if II picked a point in A', I picks the same point in A.

REFERENCES

[1] Ajtai, M., Csirmaz, L. and Nagy, Zs., On a generalization of the game Go-moku, I, *Studia Sci. Math. Hungar.* **14** (1979), 209—226.
[2] Baumgartner, J. and Laver, L., Iterated perfect set forcing, *Ann. Math. Logic.* **17** (1979), 271—288.
[3] Devlin, K. J., Variations on $\diamondsuit$, *J. Symbolic Logic* **44** (1979), 51—58. *MR* **80c**: 03050.
[4] Jensen, R. B., The fine structure of the constructible hierarchy. With a section by Jack Silver, *Ann. Math. Logic* **4** (1972), 229—308; erratum ibid. **4** (1972), 443. *MR* **46** #8834.
[5] Juhász, I., *Cardinal function in topology,* Amsterdam, 1975.
[6] Martin, D. A., Borel determinacy, *Ann. of Math.* (2) **102** (1975), 363—371. *MR* **53** #7785.

(Received July 2, 1980)

MTA MATEMATIKAI KUTATÓ INTÉZETE
REÁLTANODA U. 13—15.
H—1053 BUDAPEST
HUNGARY

PRINTED IN HUNGARY
Szegedi Nyomda, Szeged

Die *Studia Scientiarum Mathematicarum Hungarica* ist eine Halbjahrsschrift der Ungarischen Akademie der Wissenschaften. Sie veröffentlicht Originalbeiträge aus dem Bereich der Mathematik in deutscher, englischer, französischer oder russischer Sprache. Es erscheint jährlich ein Band.

Adresse der Redaktion: 1053 Budapest V., Reáltanoda u. 13—15, Ungarn.
Technischer Redaktor: J. Merza

Bestellbar bei Buch- und Zeitungs-Aussenhandelsunternehmen *Kultura* (Budapest 62, P. O. B. 149), oder bei den Vertretungen im Ausland.
Austauschabmachungen können mit der Bibliothek des Mathematischen Instituts (1053 Budapest V., Reáltanoda u. 13—15) getroffen werden.
Die zur Veröffentlichung bestimmten Manuskripte sind in zwei Exemplaren an die Redaktion zu schicken.

Studia Scientiarum Mathematicarum Hungarica est une revue biannuelle de l'Académie Hongroise des Sciences publiant des essais originaux, en francais, anglais, allemand ou russe, du domaine des mathématiques.

Rédaction: 1053 Budapest V., Reáltanoda u. 13—15, Hongrie.
Rédacteur technique: J. Merza

On s'abonne chez *Kultura*, Société pour le Commerce de Livres et Journaux (Budapest 62, P. O. B. 149) ou chez ses représentants à l'étranger.
Pour établir des relations d'échange on est prié de s'adresser à la Bibliothèque de l'Institut de Mathématique (1053 Budapest V., Reáltanoda u. 13—15).
On est prié d'envoyer les articles destinés à la publication en deux exemplaires à l'adresse de la Rédaction.

Studia Scientiarum Mathematicarum Hungarica — выходит два раза в год в Издании Академии Наук Венгрии. Журнал публикует оригинальные исследования в области математики на русском, немецком, английском, и французском языках. Отдельные выпуски составляют ежегодно один том.

Адрес редакции: 1053 Budapest V., Reáltanoda u. 13—15, Венгрия.
Технический редактор: J. Merza

Подписка на журнал принимается Внешнеторговым предприятием „Култура" (Budapest 62, P. O. B. 149) или его представительствами за границей.
По поводу отношения обмена просим обрашаться к Библиотеке Института Математики (1053 Budapest V., Reáltanoda u. 13—15).
Работы, преднасначенные для опубликования в журнале следует направлять по адресу редакции в двух экземплярах.

INDEX

HU ISSN 0081—6906

Index: 26 748