Studia Scientiarum Mathematicarum Hungarica

VOLUME 17

NUMBERS 1–4

1982

AKADÉMIAI KIADÓ, BUDAPEST

STUDIA SCIENTIARUM MATHEMATICARUM HUNGARICA

A QUARTERLY OF THE HUNGARIAN ACADEMY OF SCIENCES

Studia Scientiarum Mathematicarum Hungarica publishes original papers on mathematics mainly in English, but also in German, French and Russian.

Studia Scientiarum Mathematicarum Hungarica is published in yearly volumes of four issues (mostly double numbers published semiannually) by

AKADÉMIAI KIADÓ
Publishing House of the Hungarian Academy of Sciences
H-1054 Budapest, Alkotmány u. 21.

Manuscripts and editorial correspondence should be addressed to

J. Merza
Managing Editor
H-1053 Budapest, Reáltanoda u. 13—15

Subscription information

Orders should be addressed to

KULTURA Foreign Trading Company
H-1389 Budapest, P. O. Box 149

or to its representatives abroad.

Studia Scientiarum Mathematicarum Hungarica

Volume 17

Akadémiai Kiadó, Budapest

1982

STUDIA SCIENTIARUM MATHEMATICARUM HUNGARICA

Volume 17

CONTENTS

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 1—20

THE ALGEBRA OF FUNCTIONS ON A GRAPH

P. RIBENBOIM

Dedicated to the memory of Professor L. Rédei, illustrious algebraist

Möbius inversion formula for arithmetic functions is a classical tool for deriving explicit formulas. Weisner [10], Ward [9], Delsarte [1], Rédei [5], and Wiegandt [11] generalized this theory to new settings, indicating at the same time various applications to the theories of abelian groups and groups.

In recent years, the work of Rota [7] on the foundations of combinatorial theory brought to light the importance of incidence algebras and of Möbius inversion formula, which were studied in the context of locally finite partially ordered sets.

In this paper, we consider the situation of a directed multigraph, called a graph for simplicity. We establish the Möbius inversion formula in the algebra of functions on the paths of a locally finite graph. We give the natural matrix representation and generalize Stanley's theorem, assuming that the functions on the graph have values in a noetherian ring R (see Proposition 7). In Proposition 8, we characterize the algebras of matrices which represent algebras of functions with values on R, defined on the paths of a locally finite graph. In the last section, we consider locally finite semi-affine graphs Γ (see definition in § 4) and the subalgebra of invariant functions. We show that if Γ is division-closed, the algebra is commutative, and we observe that this is a reasonable generalization of Dirichlet's convolution of arithmetic functions.

Leroux and his collaborators, wrote at the same time, and independently, a series of papers ([2], [3], [4]) on the Möbius categories, which are, without any doubt, a very appropriate and all-embracing setting for these studies. Our basic constructions and some of our more elementary results may be obtained with Leroux's approach and are more or less apparent in the previous papers of Rota [7] and Stanley [8]. We have included their proofs here to make this paper more self-contained and not to give to the reader the impression that a more general point of view would be required to derive our limited results — without, on the other hand, disputing the interest of Leroux's work.

In order to avoid any misunderstanding concerning the terminology, we give explicitly the definitions of all the concepts used in this paper. To facilitate the task of the reader, we include all the details of the proofs, even though in some cases (as we indicate) similar proofs are sketched in other papers.

1980 *Mathematics Subject Classification.* Primary 05C20; Secondary 05C50.
Key words and phrases. Digraphs, incidence algebras, Möbius inversion.

1. Definition of the algebra of functions of any graph

Let $\Gamma=(\Gamma, V(\Gamma), o, t)$ be a (directed) graph, that is, Γ is a set, $V(\Gamma)$ a subset of Γ, o, t are mappings from Γ to $V(\Gamma)$ such that $o(x)=t(x)=x$ for every $x\in V(\Gamma)$. $V(\Gamma)$ is the set of vertices, $E(\Gamma)=\Gamma\setminus V(\Gamma)$ is the set of edges, $o(z)$ is the origin of z, $t(z)$ is the terminal of z, $o(z)$, $t(z)$ are the extremities of z.

Paths. Every vertex is a path of length 0. A path of length $n\geqq 1$ is an n-tuple $\eta=(y_n, \dots, y_1)$ where each y_i is an edge and $o(y_i)=t(y_{i-1})$ for $i=2, \dots, n$. We put $o(\eta)=o(y_1)$, $t(\eta)=t(y_n)$, and $l(\eta)=n$; $o(\eta)$ is the origin of η, $t(\eta)$ is the terminal of η and $o(\eta)$, $t(\eta)$ are the extremities of η.

A *circuit* is a path η of length $n\geqq 1$ such that $o(\eta)=t(\eta)$. A circuit of length 1 is a *loop*.

An *irreducible path* is a path of length 0 or 1, or any path $\eta=(y_n, \dots, y_1)$ with $n\geqq 2$, such that the vertices $o(y_1), o(y_2), \dots, o(y_n)$ are all distinct.

Thus, every path is irreducible if and only if Γ has no circuits.

Subpaths. The only subpath of $v\in V(\Gamma)$ is v.

If $\eta=(y_n, \dots, y_1)$ is a path of length $n\geqq 1$, a *subpath* of η is either $\eta'=o(y_i)$ $(i=1, \dots, n)$, or $\eta'=t(y_n)$, or $\eta'=(y_j, \dots, y_{i+1}, y_i)$ for $1\leqq i\leqq j\leqq n$. Any subpath of an irreducible path is an irreducible path.

If η is a path, and $\eta_1, \dots, \eta_m$ are subpaths of η, such that $o(\eta_1)=o(\eta)$, $t(\eta_m)=t(\eta)$, $o(\eta_i)=t(\eta_{i-1})$ for $i=2, \dots, m$, we write $\eta=\eta_m\circ\dots\circ\eta_1$. In particular, $v=v\circ v$ and if $z\in E(\Gamma)$, $o(z)=v$, $t(z)=w$ then $z=z\circ v=w\circ z$.

η' is an *initial* subpath of η when there exists a subpath η'' of η such that $\eta=\eta''\circ\eta'$. Similarly, η'' is a *terminal* subpath of η when there exists a subpath η' of η such that $\eta=\eta''\circ\eta'$. Thus $o(\eta)$ is an initial subpath and $t(\eta)$ is a terminal subpath of the path η.

We denote by $P(\Gamma)$ the set of irreducible paths of Γ. If $v, w\in V(\Gamma)$ let $\pi(v, w)$ be the cardinal number of the set $\{\eta\in P(\Gamma)|o(\eta)=v, t(\eta)=w\}$.

Γ is *locally finite* when Γ has no circuits and $\pi(v, w)<\infty$, for any $v, w\in V(\Gamma)$. Equivalently, between any two vertices there are only finitely many paths.

Let R be a commutative ring with unit element 1. Let $\mathscr{F}=\mathscr{F}(\Gamma, R)$ be the set of all functions from $P(\Gamma)$ to R.

If $f, g\in\mathscr{F}$ define $(f+g)(\eta)=f(\eta)+g(\eta)$. Thus $\mathscr{F}$ is an abelian additive group.

We define the convolution $*$ as follows:

$$(f*g)(\eta) = \sum_{\eta=\eta''\circ\eta'} f(\eta'')g(\eta').$$

In particular, for every $v\in V(\Gamma)$:

$$(f*g)(v) = f(v)g(v),$$

and for any edge z with $o(z)=v$, $t(z) = w$:

$$(f*g)(z) = f(z)g(v)+f(w)g(z).$$

It is easily seen that $\mathscr{F}$ is an associative ring with these operations, having unit element e:

$$\begin{cases} e(v) = 1 & \text{when } v\in V(\Gamma) \\ e(\eta) = 0 & \text{when } \eta \text{ has length at least 1.}\end{cases}$$

$\mathcal{F}$ is actually an algebra over R, defining $(rf)(\eta)=rf(\eta)$ for every $\eta\in P(\Gamma)$.

PROPOSITION 1. *Let* $f\in\mathcal{F}$. f *is invertible in the ring* $\mathcal{F}$ *if and only if* $f(v)$ *is invertible in* R, *for every* $v\in V(\Gamma)$.

PROOF. If $f*f'=e$ then $1=e(v)=f(v)f'(v)$ so $f(v)$ is invertible. Conversely, let $f'(v)=\frac{1}{f(v)}$ for every $v\in V(\Gamma)$. Suppose that $f'(\varrho)$ has been defined for every irreducible path ϱ of length less than n, in such a way that

$$\sum_{\varrho=\varrho''\circ\varrho'} f(\varrho'')f'(\varrho')=e(\varrho).$$

Let η be any irreducible path of length n, $o(\eta)=v$, $t(\eta)=w$. Put

$$f'(\eta)=-\frac{1}{f(w)}\sum_{\substack{\eta=\eta''\circ\eta'\\ l(\eta'')\geqq 1}} f(\eta'')f'(\eta')$$

(note that $l(\eta')\leqq n-1$). Then

$$\sum_{\eta=\eta''\circ\eta'} f(\eta'')f(\eta')=0=e(\eta).$$ ∎

The *Riemann function* ζ is defined by $\zeta(\eta)=1$ for every $\eta\in P(\Gamma)$.

The *Möbius function* μ is the inverse of the Riemann function: $\mu*\zeta=\zeta*\mu=e$.

PROPOSITION 2 (the inversion formulas).

a) *If* $f(\eta)=\Sigma\{g(\eta')|\eta'$ initial path of $\eta\}$ *for every* $\eta\in P(\Gamma)$, *then*

$$g(\eta)=\sum_{\eta=\eta''\circ\eta'}\mu(\eta'')f(\eta')$$

for every $\eta\in P(\Gamma)$.

b) *If* $f(\eta)=\Sigma\{g(\eta'')|\eta''$ terminal path of $\eta\}$ *for every* $\eta\in P(\Gamma)$, *then*

$$g(\eta)=\sum_{\eta=\eta''\circ\eta'} f(\eta'')\mu(\eta')$$

for every $\eta\in P(\Gamma)$.

PROOF. a) By hypothesis, $f=\zeta*g$, hence $g=\mu*f$. b) The proof is similar. ∎

We compute explicitly the Möbius function. Let $s=e-\zeta$ so

$$\begin{cases} s(v)=0 & \text{for every } v\in V(\Gamma)\\ s(\eta)=-1 & \text{for every } \eta\in P(\Gamma),\ l(\eta)\geqq 1.\end{cases}$$

LEMMA 1. *If* $l(\eta)=m\geqq 1$ *and* $r\geqq 1$ *then* $s^{*r}(\eta)=(-1)^r\binom{m-1}{r-1}$. *In particular* $s^{*(m+1)}(\eta)=0$.

PROOF. If $r=1$ it is true. We assume true for $r-1$. Then

$$s^{*r}(\eta) = (s^{*(r-1)} * s)(\eta) = \sum_{\eta=\eta''\circ\eta'} s^{*(r-1)}(\eta'')s(\eta') =$$

$$= s^{*(r-1)}(\eta)s(o(\eta)) + \sum_{\substack{\eta=\eta''\circ\eta' \\ l(\eta'')\leqq m-1}} s^{*(r-1)}(\eta'')s(\eta') =$$

$$= (-1)^r \sum_{k=1}^{m-1} \binom{k-1}{r-2} = (-1)^r \binom{m-1}{r-1}. \quad \blacksquare$$

PROPOSITION 3. *If* $l(\eta)=m\geqq 0$ *then*

$$\mu(\eta) = (e+s+s^{*2}+\ldots+s^{*m})(\eta).$$

PROOF.

$$[(e+s+s^{*2}+\ldots+s^{*m})*\zeta](\eta) = [(e+s+s^{*2}+\ldots+s^{*m})*(e-s)](\eta) =$$

$$= [e-s^{*(m+1)}](\eta) = e(\eta). \quad \blacksquare$$

Note that $\mu(\eta)$ depends only on the length of η. More explicitly, we have:

PROPOSITION 4. *If* $\eta\in P(\Gamma)$ *then*

$$\mu(\eta) = \begin{cases} 1 & \textit{when} \quad l(\eta)=0 \\ -1 & \textit{when} \quad l(\eta)=1 \\ 0 & \textit{when} \quad l(\eta)\geqq 2. \end{cases}$$

PROOF. It is trivial when $l(\eta)=0$ or 1. Let $l(\eta)=m\geqq 2$, $\eta=(y_m, \ldots, y_1)$. By Lemma 1, and Proposition 3,

$$\mu(\eta) = (e+s+s^{*2}+\ldots+s^{*m})(\eta) =$$

$$= -1+\binom{m-1}{1}-\binom{m-1}{2}+\ldots+(-1)^m\binom{m-1}{m-1} = -(1-1)^{m-1} = 0.$$

2. Comparison with the incidence algebra of partially ordered sets

Let $(S, \leqq)$ be a partially pre-ordered set, that is, $\leqq$ is a reflexive and transitive binary relation on S. Let $\Gamma(S)=\{(x, y)\in S\times S\,|\,x\leqq y\}$, $V(\Gamma(S))=\{(x, x)\,|\,x\in S\}$, and $o(x, y)=(x, x)$, $t(x, y)=(y, y)$. This defines a graph, denoted by $\Gamma(S)$. We identify S with $V(\Gamma(S))$.

Note that $\Gamma(S)$ is a *combinatorial* graph, that is, between two vertices there is at most one edge. In particular, $\Gamma(S)$ has no loops.

Actually, $S\to\Gamma(S)$ defines a covariant functor γ from the category of partially pre-ordered sets to the category of combinatorial graphs. Moreover, S is partially ordered if and only if $\Gamma(S)$ has no circuits.

Conversely, if Γ is any graph, if $x, y\in V(\Gamma)$, let $x\leqq y$ when there exists a path η of Γ such that $o(\eta)=x$, $t(\eta)=y$. Then $(V(\Gamma), \leqq)$ is a partially pre-ordered set. This defines a covariant functor γ' from the category of graphs to the category of partially pre-ordered sets. Moreover, Γ has no circuits if and only if $(V(\Gamma), \leqq)$ is a partially ordered set.

It is immediate that $\gamma' \circ \gamma$ is the identity functor. Hence, γ defines an embedding from partially pre-ordered (partially ordered) sets to combinatorial graphs (with no circuits).

Given the partially ordered set $(S, \leqq)$, let $\Gamma(S)$ be the associated graph (hence $\Gamma(S)$ has no circuits). We denote by $\mathrm{Segm}\,(S)$ the set of all segments $[x, y] = \{z \in S \mid x \leqq z \leqq y\}$ of S.

Let $\sigma\colon P(\Gamma(S)) \to \mathrm{Segm}\,(S)$ be defined by $\sigma(\eta) = [o(\eta), t(\eta)]$. Note that σ is surjective.

Now, we assume that the partially ordered set, $(S, \leqq)$ is *locally finite*, that is, for any $x, y \in S$ the segment $[x, y]$ is finite. The *incidence algebra* $\mathscr{I} = \mathscr{I}(S, R)$ is defined as follows. It is the set of all mappings from the set $\mathrm{Segm}\,(S)$ into R, with operations:

$$(f+g)[x, y] = f[x, y] + g[x, y]$$

$$(f * g)[x, y] = \sum_{x \leqq u \leqq y} f[u, y]\, g[x, u]$$

(this is a finite sum, since $[x, y]$ is finite).

$\mathscr{I}$ is an (associative) ring, with unit element e:

$$\begin{cases} e[x, x] = 1 \\ e[x, y] = 0 \quad \text{when} \quad x \neq y \end{cases}$$

and in fact, $\mathscr{I}$ is an R-algebra, with $(rf)[x, y] = rf[x, y]$.

Let $\tilde{\sigma}\colon \mathscr{I} \to \mathscr{F} = \mathscr{F}(\Gamma(S), R)$ be defined by $\tilde{\sigma}(f) = f \circ \sigma$. Then $\tilde{\sigma}$ is injective, $\tilde{\sigma}(f+g) = \tilde{\sigma}(f) + \tilde{\sigma}(g)$, $\tilde{\sigma}(rf) = r\tilde{\sigma}(f)$ for $f, g \in \mathscr{I}$, $r \in R$.

Since S is locally finite then $\Gamma(S)$ is also locally finite, so it defines a function $\pi\colon \mathrm{Segm}\,(S) \to \mathbf{N}$, namely $\pi[x, y]$ is the number of paths η in $\Gamma(S)$ with $o(\eta) = x$, $t(\eta) = y$.

We define $I\colon \mathscr{F} \to \mathscr{I}$ by $(If)[x, y] = \sum_{\substack{o(\eta)=x \\ t(\eta)=y}} f(\eta)$ for every $f \in \mathscr{F}$ and segment $[x, y]$ of S. In particular, $(If)[x, x] = f(x)$ for every $x \in S$.

We have $I(f+g) = I(f) + I(g)$, $I(rf) = rI(f)$ for $f, g \in \mathscr{F}$, $r \in R$.

$I \circ \tilde{\sigma}$ is the multiplication with the function π. Indeed

$$(I \circ \tilde{\sigma}(f))[x, y] = \sum_{\substack{o(\eta)=x \\ t(\eta)=y}} (\tilde{\sigma} f)(\eta) = \sum_{\substack{o(\eta)=x \\ t(\eta)=y}} f[x, y] = \pi[x, y] f[x, y].$$

I is surjective. Indeed, given $f \in \mathscr{I}$ let $f' \in \mathscr{F}$ be defined as follows. For every vertex x and every edge (x, y) of $\Gamma(S)$ let $f'(x) = f[x, x]$, $f'((x, y)) = f[x, y]$; for every path η of length at least 2, let $f'(\eta) = 0$. Then $If' = f$.

We have the exact sequence of R-modules

$$0 \to \mathrm{Ker}(I) \to \mathscr{F} \xrightarrow[I]{} \mathscr{I} \to 0.$$

Note that if $f \in \mathrm{Ker}\,(I)$ then $f(x) = 0$ for every $x \in S$.

Moreover

$$I(f * g) = I(f) * I(g).$$

Indeed, for every segment $[x, y]$ we have

$$(I(f*g))[x, y] = \sum_{\substack{o(\eta)=x\\ t(\eta)=y}} (f*g)(\eta) = \sum_{\substack{o(\eta)=x\\ t(\eta)=y}} \sum_{\eta=\eta''\circ\eta'} f(\eta'')g(\eta')$$

while

$$(I(f)*I(g))[x, y] = \sum_{x\leq u\leq y} I(f)[u, y]I(g)[x, u] =$$

$$= \sum_{x\leq u\leq y} \Big(\sum_{\substack{o(\eta'')=u\\ t(\eta'')=y}} f(\eta'')\Big)\Big(\sum_{\substack{o(\eta')=x\\ t(\eta')=y}} g(\eta')\Big)$$

hence the two expressions are equal.

It follows that Ker (I) is a two-sided ideal of the R-algebra $\mathscr{F}$.

If R is a **Q**-algebra, we define also the mapping $I'\colon \mathscr{I}\to\mathscr{F}$ as follows:

$$(I'f)(\eta) = \frac{1}{\pi[o(\eta), t(\eta)]} f[o(\eta), t(\eta)].$$

Clearly, I' is an R-module homomorphism and $I\circ I'$ is the identity map. Moreover

$$I'(If)(\eta) = \frac{1}{\pi[o(\eta), t(\eta)]} \sum_{\substack{\varrho\\ o(\varrho)=o(\eta)\\ t(\varrho)=t(\eta)}} f(\varrho).$$

We have also

$$\sum_{\substack{\varrho\\ o(\varrho)=x, t(\varrho)=y}} (I'f*I'g)(\varrho) = \sum_{\substack{\varrho\\ o(\varrho)=x, t(\varrho)=y}} \sum_{\varrho=\varrho''\circ\varrho'} (I'f)(\varrho'')(I'g)(\varrho') =$$

$$= \sum_{\substack{\varrho\\ o(\varrho)=x, t(\varrho)=y}} \sum_{u \text{ vertex of } \varrho} \frac{1}{\pi[u, y]} f[u, y] \frac{1}{\pi[x, u]} g[x, u].$$

Let $\pi[x, y; u]$ be the number of paths ϱ such that $o(\varrho)=x$, $t(\varrho)=y$ and u is a vertex of some edge of ϱ. The above sum is equal to

$$\sum_u \sum_{\substack{\varrho\\ o(\varrho)=x, t(\varrho)=y\\ u \text{ is a vertex of } \varrho}} \frac{1}{\pi[u, y]} f[u, y] \frac{1}{\pi[x, u]} g[x, u] =$$

$$= \sum_u f[u, y] g[x, u] \sum_{\substack{\varrho\\ u \text{ vertex of } \varrho}} \frac{1}{\pi[u, y]} \times \frac{1}{\pi[x, u]} =$$

$$= \sum_u f[u, y] g[x, u] \sum_{\substack{\varrho\\ u \text{ vertex of } \varrho}} \frac{1}{\pi[x, y; u]} =$$

$$= \sum_u f[u, y] g[x, u] = (f*g)[x, y] = \pi[x, y][I'(f*g)](\eta)$$

where η is any path such that $o(\eta)=x$, $t(\eta)=y$.

3. Matrix representation

Let Γ be a locally finite graph, R a ring with unit element, $\mathscr{F}=\mathscr{F}(\Gamma, R)$. With every $f\in\mathscr{F}$ we associate the matrix M_f, with entries in R, and rows and columns indexed by the set of vertices $V(\Gamma)$. Explicitly, if $x, y\in V(\Gamma)$, the entry at row y, column x, of M_f is

$$M_f(y, x) = \Sigma\{f(\eta)|\eta\in P(\Gamma),\ o(\eta) = x,\ t(\eta) = y\}$$

(note that this is a finite sum; it is equal to 0 if there exists no path η such that $o(\eta)=x$, $t(\eta)=y$).

Clearly, $M_{f+g}=M_f+M_g$. Moreover, $M_{f*g}=M_f M_g$. Indeed,

$$M_{f*g}(y, x) = \sum_{\substack{o(\eta)=x\\ t(\eta)=y}} (f*g)(\eta) = \sum_{\substack{o(\eta)=x\\ t(\eta)=y}} \Big(\sum_{\eta=\eta''\circ\eta'} f(\eta'')g(\eta')\Big).$$

On the other hand

$$(M_f M_g)(y, x) = \sum_{u\in V(\Gamma)} M_f(y, u) M_g(u, x) =$$

$$\sum_{u\in V(\Gamma)} \Big(\sum_{\substack{o(\varrho)=u\\ t(\varrho)=y}} f(\varrho)\Big)\Big(\sum_{\substack{o(\sigma)=x\\ t(\sigma)=u}} g(\sigma)\Big)$$

(note that there can only be finitely many vertices u for which the corresponding summand is not 0), hence the two expressions are equal.

Thus $f\mapsto M_f$ defines an R-algebra homomorphism $M^\Gamma: \mathscr{F}(\Gamma, R)\to\mathscr{M}(V(\Gamma), R)$, the R-algebra of matrices with entries in R and rows and columns indexed by $V(\Gamma)$. Note that $M_f(x, x)=f(x)$ for every $x\in V(\Gamma)$. Hence, if $f\in\operatorname{Ker}(M^\Gamma)$ then $f(x)=0$ for every $x\in V(\Gamma)$.

For example, taking the functions e, ζ, μ we have

$$M_e(y, x) = \begin{cases} 1 & \text{when } x = y \\ 0 & \text{when } x \neq y \end{cases}$$

$$M_\zeta(y, x) = \pi(x, y)$$

$$M_\mu(y, x) = \begin{cases} 1 & \text{when } x = y \\ -\#\ (\text{edges from } x \text{ to } y) & \text{when } x \neq y. \end{cases}$$

If $(S, \leqq)$ is a locally finite partially ordered set and $\mathscr{I}$ is the incidence algebra of S, we have the matrix representation $M^S: \mathscr{I}(S, R)\to\mathscr{M}(S, R)$, defined by $M^S_f(y, x)=f[x, y]$ (it is understood that $f[x, y]=0$ if $x\not\leqq y$). M^S is an injective homomorphism of R-algebras, whose image is the subalgebra of all matrices $A=(A(y, x))_{x, y\in S}$ such that if $x\not\leqq y$ then $A(y, x)=0$.

If $\Gamma(S)$ is the graph associated with S then $M^{\Gamma(S)}=M^S\circ I$, in particular $\operatorname{Ker}(M^{\Gamma(S)})=\operatorname{Ker}(I)$, $\operatorname{Im}(M^{\Gamma(S)})=\operatorname{Im}(M^S)$.

On the other hand, if Γ is a locally finite graph, let $S(\Gamma)$ be the partially ordered set associated with Γ, as indicated in § 2: if $x, y\in V(\Gamma)$ then $x\leqq y$ when there exists a path η of Γ such that $o(\eta)=x$, $t(\eta)=y$. Then $S(\Gamma)$ is locally finite, so we may compare the R-algebra $\mathscr{F}$, defined by Γ, and the incidence algebra $\mathscr{I}'$ defined by $S(\Gamma)$. Let I': $\mathscr{F} \to \mathscr{I}'$ be defined by

$$(I'f)[x, y] = \Sigma\{f(\eta) | \eta \in P(\Gamma), o(\eta) = x, t(\eta) = y\}.$$

Then I' is an R-algebra homomorphism. Moreover, I' is surjective. Indeed, if x, y are such that $x \leqq y$, let η_{xy} be a path such that $o(\eta_{xy}) = x$ $t(\eta_{xy}) = y$.

Given $g \in \mathscr{I}'$, let $h \in \mathscr{F}$ be defined as follows:

$$\begin{cases} h(\eta_{xy}) = g[x, y] \\ h(\eta) = 0 \quad \text{for any path} \quad \eta \neq \eta_{xy}, \quad o(\eta) = x, \quad t(\eta) = y. \end{cases}$$

Then $I'h = g$.

Let $M^\Gamma: \mathscr{F} \to \mathscr{M}(V(\Gamma), r)$, $M^{S(\Gamma)}: \mathscr{I}' \to \mathscr{M}(V(\Gamma), R)$ be the corresponding matrix representations. Then $M^{S(\Gamma)} \circ I' = M^\Gamma$. $M^{S(\Gamma)}$ is injective, $\operatorname{Ker}(I') = \operatorname{Ker}(M^\Gamma)$ and since I' is surjective then $\operatorname{Im}(M^\Gamma) = \operatorname{Im}(M^{S(\Gamma)})$.

We note also that there is a natural surjective mapping $\varphi: P(\Gamma) \to \Gamma(S(\Gamma))$ namely $\varphi(\eta) = (x, y)$ where $o(\eta) = x, t(\eta) = y$. φ is injective if and only if given any two vertices x, y of Γ there is at most one path η such that $o(\eta) = x, t(\eta) = y$.

If Γ is any graph, its *Hasse diagram* is the graph $H(\Gamma) = V(\Gamma) \cup E'$ where E' is the set of all pairs (x, y) with $x, y \in V(\Gamma), x \neq y$ and $\max\{l(\eta) | \eta \in P(\Gamma), o(\eta) = x, t(\eta) = y\} = 1$. Moreover $o'(x, y) = x, t'(x, y) = y$ for every $(x, y) \in E'$.

If S is a partially ordered set, the *Hasse diagram of* S is $H(S) = H(\Gamma(S))$.

First, we note:

Lemma 2. *If Γ, Δ are graphs with no circuits and if the associated partially ordered sets are isomorphic, $S(\Gamma) \cong S(\Delta)$, then so are the Hasse diagrams $H(\Gamma) \cong H(\Delta)$.*

Proof. Let $(x, y) \in H(\Gamma)$, so $x < y$ and there is an edge z of Γ, with $o(z) = x, t(z) = y$, but there is no path η of Γ, with $l(\eta) > 1, o(\eta) = x, t(\eta) = y$. For the corresponding elements in $S(\Delta)$, we have $x' < y'$, thus there is a path ϱ' in Δ such that $o(\varrho') = x', t(\varrho') = y'$. If $\varrho' = (z'_n, \ldots, z'_1)$, $n \geqq 1$, then $x' < t(z'_1) \leqq y'$ so $x < u \leqq y$, where u corresponds to $t(z'_1)$. Thus there exist paths η', η'' in Γ such that $o(\eta') = x, t(\eta') = o(\eta'') = u, t(\eta'') = y$. By hypothesis, $u = y, \eta'$ is an edge so $n = 1$ showing that there is an edge z' of Δ such that $o(z') = x', t(z') = y'$. The above proof shows also that $(x', y') \in H(\Delta)$.

It is easily seen that this correspondence is an isomorphism: $H(\Gamma) \cong H(\Delta)$. ∎

For later use, we note also the following

Lemma 3. *Let Γ be a graph. There is an isomorphism preserving the vertices between Γ and $H(\Gamma)$ if and only if Γ is a combinatorial graph, satisfying the following condition: $(*)$ if z is an edge of Γ, if $\eta \in P(\Gamma)$ and $o(\eta) = o(z), t(\eta) = t(z)$ then $\eta = (z)$.*

Proof. If there is an isomorphism as indicated, then Γ is combinatorial, because $H(\Gamma)$ is combinatorial. Also, if z is an edge of Γ, by the isomorphism z must correspond to the edge $(o(z), t(z))$ of $H(\Gamma)$. If there is a path $\eta = (z_n, \ldots, z_1)$ in Γ, with $o(\eta) = o(z), t(\eta) = t(z)$, then each z_i corresponds by the isomorphism to the edge $(o(z_i), t(z_i))$ of $H(\Gamma)$, so η corresponds to a path in $H(\Gamma)$, with $o(\eta) = o(z), t(\eta) = t(z)$. Thus η must have length 1, so $\eta = (z)$, proving condition $(*)$.

Conversely, let (x, y) be an edge of the Hasse diagram $H(\Gamma)$; thus $x, y \in V(\Gamma)$, $x \neq y$, there is one edge $z \in \Gamma$ with $o(z)=x$, $t(z)=y$ and if $\eta \in P(\Gamma)$, $o(\eta)=x$, $t(\eta)=y$, then η has length 1. So $\eta=(z)$, where z is the only edge of Γ such that $o(z)=x$, $t(z)=y$. With (x, y) we associate the edge z. This mapping is injective. It is also surjective, because if z is an edge of Γ, $o(z)=x$, $t(z)=y$, $x \neq y$, then $\eta=(z)$ is the only path such that $o(\eta)=x$, $t(\eta)=y$; hence $(x, y) \in H(\Gamma)$ and z is the corresponding edge. ▮

Stanley [8] announced the following result and sketched its proof.

Proposition 5. *Let R be a field, let S, T be locally finite partially ordered sets, let $\mathcal{I}(S, R)$, $\mathcal{I}(T, R)$ be the incidence algebras over R, let $M^S: \mathcal{I}(S, R) \to \mathcal{M}(S, R)$, $M^T: \mathcal{I}(T, R) \to \mathcal{M}(T, R)$ be the corresponding matrix representations. If $\mathrm{Im}(M^S) \cong \mathrm{Im}(M^T)$ then $S \cong T$.*

As a corollary, we have

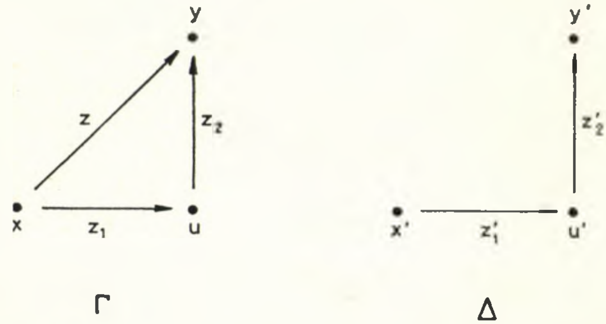
Proposition 6. *Let R be a field, let Γ, Δ be locally finite graphs, let $\mathcal{F}(\Gamma, R)$, $\mathcal{F}(\Delta, R)$ be the corresponding R-algebras of functions, let $M^\Gamma: \mathcal{F}(\Gamma, R) \to \mathcal{M}(\Gamma, R)$, $M^\Delta: \mathcal{F}(\Delta, R) \to \mathcal{M}(\Delta, R)$ be the respective matrix representations. If*

$$\mathrm{Im}(M^\Gamma) \cong \mathrm{Im}(M^\Delta) \text{ then } H(\Gamma) \cong H(\Delta).$$

Proof. Let S, T be the partially ordered sets associated with Γ, Δ, respectively, so S, T are locally finite. We have seen that $\mathrm{Im}(M^\Gamma) = \mathrm{Im}(M^S)$, $\mathrm{Im}(M^\Delta) = \mathrm{Im}(M^T)$. By Proposition 5, $S \cong T$, and by Lemma 2, $H(\Gamma) \cong H(\Delta)$. ▮

Corollary. *Under the hypothesis of the proposition if moreover Γ, Δ are combinatorial graphs satisfying the condition $(*)$ of Lemma 3, then $\Gamma \cong \Delta$.*

Note. In Proposition 6, it is not possible to conclude, in general, that $\Gamma \cong \Delta$. For example, let



For every $f \in \mathcal{F}(\Gamma, R)$, let $f' \in \mathcal{F}(\Delta, R)$ be defined by $f'(x') = f(x)$, $f'(u') = f(u)$, $f'(y') = f(y)$, $f'(z_1') = f(z_1)$, $f'(z_2') = f(z_2)$, $f'(z_2', z_1') = f(z_2, z_1) + f(z)$. Let $\varphi: \mathrm{Im}(M^\Gamma) \to \mathrm{Im}(M^\Delta)$ be defined by $\varphi(M_f^\Gamma) = M_{f'}^\Delta$. Then φ is an isomorphism, however Γ, Δ are not isomorphic.

We now indicate a generalization of Stanley's Proposition 5, for the case of an arbitrary (commutative) noetherian ring R.

First, we recall some well-known facts about rings. Let S be a ring (which is not necessarily commutative). Each representation of S as a finite direct product of rings corresponds to a decomposition of 1 as a sum of pairwise orthogonal (non-zero) central idempotents. Moreover, the factors in the decomposition are indecomposable rings if and only if the corresponding idempotents are primitive, that is, not equal to the sum of two central idempotents; note that if an idempotent is the sum of two central idempotents then it is the sum of two orthogonal central idempotents.

If R is a noetherian ring then every central idempotent is a sum of finitely many primitive orthogonal central idempotents.

For later use, we need also the following considerations. A graph Γ is *connected* if given any vertices x, y of Γ there are paths $\eta_1, \ldots, \eta_r$, such that x is an extremity of η_1, y is an extremity of η_r and η_i, η_{i+1} have a common extremity.

A subgraph Γ' of Γ is a *connected component* of Γ if it is a maximal connected subgraph of Γ. Every graph is the disjoint union of its connected components. If Γ has $m_i \geqq 1$ connected components isomorphic to Γ_i (for $i=1, \ldots, r$) and no other connected component, then we write $\Gamma = m_1\Gamma_1 + \ldots + m_r\Gamma_r$ (it is assumed that $\Gamma_i \not\cong \Gamma_j$ for $i \neq j$).

If Γ is any graph, if $\{1, 2, \ldots, n\}$ is the discrete graph with m vertices then the cartesian product graph $\Gamma \times \{1, 2, \ldots, n\}$ is the disjoint union of n copies of Γ. Thus, if $\Gamma = m_1\Gamma_1 + \ldots + m_r\Gamma_r$ where the graphs Γ_i are the connected components of Γ, then $\Gamma \times \{1, 2, \ldots, n\} = nm_1\Gamma_1 + \ldots + nm_r\Gamma_r$.

LEMMA 4. *If Γ, Δ are graphs with finitely many connected components and $\Gamma \times \{1, 2, \ldots, n\} \cong \Delta \times \{1, 2, \ldots, n\}$ then $\Gamma \cong \Delta$.*

PROOF. Let $\varphi: \Gamma \times \{1, 2, \ldots, n\} \to \Delta \times \{1, 2, \ldots, n\}$ be the given isomorphism, and let $\Gamma = m_1\Gamma_1 + \ldots + m_r\Gamma_r$, $\Delta = m_1'\Delta_1 + \ldots + m_s'\Delta_s$ be the decompositions as sums of connected components. Since Γ_i is a connected component of $\Gamma \times \{1, 2, \ldots, n\}$ then $\varphi(\Gamma_i)$ is a connected component of $\Delta \times \{1, 2, \ldots, n\}$, hence there exist $\pi(i)$, $1 \leqq \pi(i) \leqq s$ such that $\varphi(\Gamma_i) = \Delta_{\pi(i)}$. The mapping π is injective, hence $r \leqq s$. Also, all copies of Γ_i have image $\Delta_{\pi(i)}$, hence $nm_i \leqq nm'_{\pi(i)}$ so $m_i \leqq m'_{\pi(i)}$. In similar way, $s \leqq r$, $m'_{\pi(i)} \leqq m_i$, proving the result. ∎

PROPOSITION 7. *Let R be a noetherian commutative ring, let Γ, Δ be locally finite graphs, let $\mathscr{F}(\Gamma, R)$, $\mathscr{F}(\Delta, R)$ be the corresponding R-algebras of functions and M^Γ, M^Δ the respective matrix representations. If* $\mathrm{Im}(M^\Gamma) \cong \mathrm{Im}(M^\Delta)$ *then the Hasse diagrams of Γ, Δ are isomorphic: $H(\Gamma) \cong H(\Delta)$.*

PROOF. Let $1 = \sum_{i=1}^{n} e_i$ be a representation of 1 as the sum of (non-zero) primitive orthogonal idempotents of R. Let $\mathscr{F} = \mathscr{F}(\Gamma, R)$, $M = M^\Gamma$ and $\mathscr{M} = \mathrm{Im}(M^\Gamma)$ for simplicity of notation.

For every $\eta \in P(\Gamma)$ and $i = 1, \ldots, n$, let $k_{\eta, i} \in \mathscr{F}$ be defined by

$$\begin{cases} k_{\eta, i}(\eta) = e_i \\ k_{\eta, i}(\varrho) = 0 \quad \text{for every} \quad \varrho \in P(\Gamma), \quad \varrho \neq \eta. \end{cases}$$

Thus

$$M_{k_{\eta},i}(y,x)=\begin{cases} e_i & \text{if } \ o(\eta)=x, \quad t(\eta)=y \\ 0 & \text{otherwise.}\end{cases}$$

LEMMA 5. *$\{M_{k_{v,i}} | i=1, \ldots, n;\ v\in V(\Gamma)\}$ is a maximal set of (non-zero) primitive orthogonal central idempotents of $\mathcal{M}$.*

PROOF. Clearly, $M_{k_{v,i}}$ are orthogonal central idempotents.

If $M_f\in\mathcal{M}$ (with $f\in\mathcal{F}$) is such that $M_f M_{k_{v,i}}=0$ for every $v\in V(\Gamma)$, $i=1, \ldots, n$, then for every $x, y\in V(\Gamma)$:

$$0=(M_f M_{k_{x,i}})(y,x)=\sum_{u\in V(\Gamma)} M_f(y,u) M_{k_{x,i}}(u,x)=M_f(y,x)e_i.$$

Hence

$$M_f(y,x)=\sum_{i=1}^{n} M_f(y,x)e_i=0$$

so $M_f=0$.

If $M_{k_{v,i}}=M_{f_1}+\ldots+M_{f_s}$ where $M_{f_1}, \ldots, M_{f_s}$ are orthogonal central idempotents, then $M_{k_{v,i}}M_{f_j}=M_{f_j}$, $M_{f_j}M_{k_{v,i}}=M_{f_j}$ $(j=1, \ldots, s)$. So, if $y\neq v$ then

$$M_{f_j}(y,x)=\sum_{u\in V(\Gamma)} M_{k_{v,i}}(y,u) M_{f_j}(u,x)=0$$

and similarly, if $x\neq v$ then $M_{f_j}(y,x)=0$. Let $M_{f_j}(v,v)=r_j$, so $r_j\neq 0$ because $M_{f_j}\neq 0$. From $M_{f_j}M_{f_j}=M_{f_j}$ it follows that r_j is an idempotent.

We note that if $r_je_i\neq 0$ then $r_je_i=e_i$. Indeed, $e_i=r_je_i+(e_i-r_je_i)$, and since e_i is primitive then $r_je_i=e_i$. Since $r_j=r_j\left(\sum_{i=1}^{n} e_i\right)\neq 0$ there is at least one index i, $1\leqq i\leqq n$ such that $r_je_i\neq 0$. Moreover, if i' is any other index such that $r_je_{i'}\neq 0$ from $M_{f_j}M_{k_{v,i}}=M_{f_j}=M_{f_j}M_{k_{v,i'}}$ it follows that $e_i=r_je_i=r_j=r_je_{i'}=e_{i'}$, thus $i'=i$. We conclude that $r_j=r_je_i=e_i$ for all $j=1, \ldots, s$. Since $M_{f_j}M_{f_i}=0$ for $j\neq i$ then $s=1$. ∎

Let $\mathcal{J}=\{A\in\mathcal{M} | A(v,v)=0$ for every $v\in V(\Gamma)\}$ so $\mathcal{J}$ is an R-submodule of $\mathcal{M}$. We note:

LEMMA 6. *$\mathcal{J}$ is a two-sided ideal of $\mathcal{M}$ and $\bigcap_{m=1}^{\infty}\mathcal{J}^m=\{0\}$.*

PROOF. Let $M_f\in\mathcal{J}$, $M_g\in\mathcal{M}$. Then

$$(M_gM_f)(v,v)=\sum_{u\in V(\Gamma)} M_g(v,u)M_f(u,v)=$$

$$=\sum_{u\in V(\Gamma)}\Big(\sum_{\substack{\eta''\in P(\Gamma)\\ o(\eta'')=u,\, t(\eta'')=v}} g(\eta'')\Big)\Big(\sum_{\substack{\eta'\in P(\Gamma)\\ o(\eta')=v,\, t(\eta')=u}} f(\eta')\Big).$$

Since $\eta=\eta''\circ\eta'$ is a path of Γ with origin and terminal equal to v, and since there are no circuits in Γ, then $(M_gM_f)(v,v)=g(v)f(v)=M_g(v,v)M_f(v,v)=0$. Similarly, $(M_fM_g)(v,v)=0$ for every $v\in V(\Gamma)$. Therefore, $\mathcal{J}$ is a two-sided ideal of $\mathcal{M}$.

Now, let $M_f \in \mathscr{J}^m$ and let $x, y \in V(\Gamma)$ be such that the maximum of $l(\eta)$, for all paths $\eta \in P(\Gamma)$ with $o(\eta)=x$, $t(\eta)=y$, is at most $m-1$. We show that $M_f(y, x)=0$.

It suffices to prove it for matrices of the form $M_{f_1}M_{f_2}\dots M_{f_m}$, with $M_{f_i} \in \mathscr{J}$. The verification is immediate.

Hence, if $M_f \in \mathscr{J}^m$ for every $m \geqq 1$, if $x, y \in V(\Gamma)$ since Γ is locally finite, there exist s such that $l(\eta) \leqq s$ for every path η such that $o(\eta)=x$, $t(\eta)=y$.

We conclude that $M_f(y, x)=0$. Thus $M_f=0$. ∎

Let $\mathscr{V}$ be a maximal set of primitive orthogonal central idempotents of $\mathscr{M}$.

LEMMA 7. *There is a bijection* $\lambda\colon \mathscr{V} \to V(\Gamma) \times \{1, 2, \dots, n\}$ *such that* $A = M_{k_{\lambda(A)}} \in \mathscr{J}$ *for every* $A \in \mathscr{V}$. *Hence, if* $\lambda(A)=(v,i)$ *then* $A(v,v)=e_i$ *and* $A(w,w)=0$ *for every* $w \in V(\Gamma)$, $w \neq v$.

PROOF. Let $A=M_f \in \mathscr{V}$. From $A^2=A$ we deduce that

$$f(v) = M_f(v, v) = \sum_{u \in V(\Gamma)} M_f(v, u) M_f(u, v) = M_f(v, v)^2 = f(v)^2,$$

because there are no circuits in Γ. Thus $f(v) = \sum_{k \in X} e_k$ (for some subset X of $\{1, \dots, n\}$) by the argument of Lemma 5. If $A(v, v)=0$ for every $v \in V(\Gamma)$ then $A \in \mathscr{J}$, so $A=A^m \in \mathscr{J}^m$ for every $m \geqq 1$, hence $A=0$, by the preceding lemma, contrary to the hypothesis. Thus, there exists $v \in V(\Gamma)$ such that $f(v)=A(v, v) \neq 0$, hence $X \neq \emptyset$. Let $i \in X$. We have $(AM_{k_{v,i}}A)(y, x)=A(y, v)A(v, x)e_i$ so $(AM_{k_{v,i}}A)(v,v) = f(v)e_i=e_i$. Thus $AM_{k_{v,i}}A \neq 0$. We have also

$$(AM_{k_{v,i}}A)(AM_{k_{v,i}}A)(y, x) = \sum_{u \in V(\Gamma)} (AM_{k_{v,i}})(y, u)(AM_{k_{v,i}}A)(u, x) =$$

$$= \sum_{u \in V(\Gamma)} A(y, v) A(v, u) A(u, v) A(v, x) e_i = A(y, v) A(v, x) f(v) e_i =$$

$$= A(y, v) A(v, x) e_i.$$

Thus $AM_{k_{v,i}}A$ is a (non-zero) central idempotent. It follows that

$$(A - AM_{k_{v,i}}A)(A - AM_{k_{v,i}}A) = A - AM_{k_{v,i}}A - AM_{k_{v,i}}A + AM_{k_{v,i}}A = A - AM_{k_{v,i}}A$$

so $A - AM_{k_{v,i}}A$ is also a central idempotent.

From

$$A = AM_{k_{v,i}}A + (A - AM_{k_{v,i}}A)$$

and

$$AM_{k_{v,i}}A(A - AM_{k_{v,i}}A) = 0,$$

since A is primitive idempotent, then $A=AM_{k_{v,i}}A$. So

$$f(v) = A(v, v) = (AM_{k_{v,i}}A)(v, v) = e_i$$

thus $X = \{i\}$.

If $w \in V(\Gamma)$, $w \neq v$, then

$$A(w, w) = (AM_{k_{v,i}}A)(w, w) = A(w, v) A(v, w) e_i = 0$$

because there are no circuits.

We define $\lambda(A)=(v, i)$, hence $A - M_{k_{\lambda(A)}} \in \mathscr{J}$.

λ is injective: if $A, B\in\mathcal{V}$ and $\lambda(A)=\lambda(B)=(v,i)$ then $A-M_{k_{v,i}}\in\mathcal{J}$, $B-M_{k_{v,i}}\in\mathcal{J}$, hence $A-B\in\mathcal{J}$. But $AB=BA=0$, so $(A-B)^2=A+B$, hence

$$(A-B)^3=(A+B)(A-B)=A^2-B^2=A-B.$$

Therefore

$$A-B=(A-B)^3=(A-B)^9=\ldots\in\bigcap_{m\geq 1}\mathcal{J}^m=\{0\}.$$

Finally, λ is surjective. Indeed, let $(v,i)\neq\lambda(A)$ for every $A\in\mathcal{V}$. Then from $A=M_{k_{\lambda(A)}}+(A-M_{k_{\lambda(A)}})$ it follows that

$$M_{k_{v,i}}A=M_{k_{v,i}}(A-M_{k_{\lambda(A)}})\in\mathcal{J}.$$

But $(M_{k_{v,i}}A)^2=M_{k_{v,i}}A$ because $M_{k_{v,i}}$ is in the centre. So $M_{k_{v,i}}A\in\mathcal{J}^m$ for every $m\geq 1$ thus $M_{k_{v,i}}A=0$. Hence $\mathcal{V}\cup\{M_{k_{v,i}}\}$ is a set of primitive orthogonal central idempotents, properly containing $\mathcal{V}$, against the hypothesis. ▮

The following lemma, which was fully proved by Stanley, is included only for the sake of completeness.

LEMMA 8. *Let S be any ring (not necessarily commutative), let J be a two-sided ideal of S such that $\bigcap_{m\geq 1} J^m=\{0\}$. Let e,f,e',f' be idempotents such that $e'-e, f'-f\in J$. Then $eaf=0$ for every $a\in S$ if and only if $e'af'=0$ for every $a\in S$.*

PROOF. We assume that $eaf=0$ for every $a\in S$. Let $x=e'-e$, $y=f'-f$. Then $e'=e'^2=x^2+xe+ex+e$ hence $ex+xe+x^2=x$ and similarly $fy+yf+y^2=y$. If $a\in S$ then

$$e'af'=(e+x)a\,(f+y)=eaf+exaf+xeaf+x^2af+eafy+eayf+eay^2+$$
$$+exafy+exayf+exay^2+xeafy+xeayf+$$
$$+xeay^2+x^2afy+x^2ayf+x^2ay^2=$$
$$=x^2af+eay^2+exay^2+xeay^2+x^2afy+x^2ayf+x^2ay^2$$

so $e'af'\in J^2$. Replacing again x, y by the above expressions, it follows that $e'af'\in J^4$. By induction, for every $n\geq 1$, $e'af'\in J^{2^n}$. Hence $e'af'=0$ for every $a\in S$. ▮

PROOF of the proposition. Let $\mathcal{V}$ be any maximal set of (non-zero) primitive orthogonal central idempotents of $\mathcal{M}=\mathrm{Im}\,(M^\Gamma)$. Let

$$\mathcal{G}=\{(A,B)\in\mathcal{V}\times\mathcal{V}\mid \text{there exists } C\in\mathcal{M} \text{ such that } BCA\neq 0\},$$

$V(\mathcal{G})=\{(A,A)|A\in\mathcal{V}\}$, let $o,t\colon\mathcal{G}\to V(\mathcal{G})$ be defined by $o(A,B)=(A,A)$, $t(A,B)=(B,B)$. Thus, $(\mathcal{G}, V(\mathcal{G}), o, t)$ is a graph. We may identify $V(\mathcal{G})$ with $\mathcal{V}$.

It suffices to show that the Hasse diagrams of the graphs $\mathcal{G}$, $\Gamma\times\{1,2,\ldots,n\}$ are isomorphic $H(\mathcal{G})\cong H(\Gamma)\times\{1,2,\ldots,n\}$. Indeed, by means of the given isomorphism $\varphi\colon\mathrm{Im}\,(M^\Gamma)\to\mathrm{Im}\,(M^\Delta)$, it follows also that $H(\mathcal{G})\cong H(\Delta)\times\{1,2,\ldots,n\}$, hence from Lemma 4, $H(\Gamma)\cong H(\Delta)$. We show that $H(\mathcal{G})\cong H(\Gamma)\times\{1,2,\ldots,n\}$. By Lemma 7, there is a bijection $\lambda\colon\mathcal{V}\to V(\Gamma)\times\{1,\ldots,n\}$.

If (A, B) is an edge of $H(\mathcal{G})$, with $A, B\in\mathcal{V}$, $A\neq B$, there exists $C\in\mathcal{M}$ such that $BCA\neq 0$, but there exist no path $\eta\in P(\mathcal{G})$ of length $m>1$ such that $o(\eta)=A$, $t(\eta)=B$.

Let $\lambda(A)=(v, i)$, $\lambda(B)=(w, j)$. By Lemma 8, there exists $D=M_f\in\mathcal{M}$ such that $M_{k_{w,j}}DM_{k_{v,i}}\neq 0$. Since

$$(M_{k_{w,j}}DM_{k_{v,i}})(y, x) = \begin{cases} 0 & \text{if} \quad x\neq v \quad \text{or} \quad y\neq w \\ e_j D(w, v)e_i = e_j M_f(w, v)e_i = e_j\Big(\sum\limits_{\substack{\varrho\in P(\Gamma) \\ o(\varrho)=v,\, t(\varrho)=w}} f(\varrho)\Big)e_i \end{cases}$$

then $e_jD(w, v)e_i\neq 0$. In particular, $j=i$ and there is a path $\varrho\in P(\Gamma)$ such that $o(\varrho)=v$, $t(\varrho)=w$. Moreover, since $\lambda(A)\neq\lambda(B)$ then $v\neq w$.

Now let $\varrho=(z_n, \dots, z_1)$, with $o(z_1)=v$, $o(z_j)=u_j$ $(j=1, \dots, n)$, $t(z_n)=w$, $u_1=v$, $u_{n+1}=w$. Then $M_{k_{u_{j+1},i}}M_{k_{z_j,i}}M_{k_{u_j,i}}\neq 0$ (for $j=1, \dots, n$).

Let $A_j\in\mathcal{V}$ be such that $\lambda(A_j)=(u_j, i)$ (for $j=1, \dots, n+1$) so $A_1=A$, $A_{n+1}=B$. Then by Lemma 8, there exist $C_j\in\mathcal{M}$ such that $A_{j+1}C_jA_j\neq 0$. Thus $\eta=((A_n, B), \dots$ $\dots, (A_2, A_3), (A, A_2))$ is a path in $\mathcal{G}$ such that $o(\eta)=A$, $t(\eta)=B$. By hypothesis $n=1$, showing that $(v, w)\in H(\Gamma)$, hence $(\lambda(A), \lambda(B))=((v, i), (w, i))$ is identified with the edge $((v, w), i)$ of $H(\Gamma)\times\{1, \dots, n\}$.

Let $\bar\lambda(A, B)=((v, w), i)$. Then $\bar\lambda$ is injective. Finally, we show that $\bar\lambda$ is surjective. Let $((v, w), i)$ be an edge of $H(\Gamma)\times\{1, \dots, n\}$, with $v, w\in V(\Gamma)$, $v\neq w$; so there exists an edge z of Γ, but no path ϱ of length greater than 1, such that $o(z)=v$, $t(z)=w$, $o(\varrho)=v$, $t(\varrho)=w$. Let $A, B\in\mathcal{V}$ be such that $\lambda(A)=(v, i)$, $\lambda(B)=$ $=(w, i)$. From $M_{k_{w,i}}M_{k_{z,i}}M_{k_{v,i}}\neq 0$, it follows that there exist $C\in\mathcal{M}$ such that $BCA\neq 0$. In the same manner as before, we show that there exist no path $\eta\in P(\mathcal{G})$, of length greater than 1, such that $o(\eta)=A$, $t(\eta)=B$. This proves that (A, B) is an edge of $H(\mathcal{G})$ and also that $\bar\lambda(A, B)=((v, w), i)$, concluding the proof. ▮

COROLLARY. *Under the hypothesis of the proposition, if Γ, Δ are combinatorial graphs satisfying the condition* $(*)$ *of the Lemma 3, then* $\Gamma\cong\Delta$.

PROOF. We have $\Gamma\cong H(\Gamma)\cong H(\Delta)\cong\Delta$. ▮

We shall now determine which are the subalgebras, of matrices which may be obtained from graphs, by the above procedure.

Let R be, as before, a (commutative) noetherian ring, $1=\sum\limits_{i=1}^{n} e_i$, where $\{e_1, \dots, e_n\}$ is a maximal set of orthogonal primitive idempotents.

Let X be a set, $\mathcal{M}(X, R)$ the R-algebra of square matrices, with rows and columns labelled by X, with entries in R.

PROPOSITION 8. *Let $\mathcal{A}$ be a subalgebra of $\mathcal{M}(X, R)$. There is a locally finite graph Γ such that $V(\Gamma)=X$ and $\mathcal{A}\cong\operatorname{Im}(M^\Gamma)$ if and only if $\mathcal{A}$ satisfies the following conditions:*

1) *$\mathcal{A}$ has a maximal set $\mathcal{V}$ of central orthogonal primitive idempotents.*

2) *There is a bijection $\lambda\colon X\times\{1, \dots, n\}\to\mathcal{V}$ such that if $\lambda(x, i)=A$ then $A(x, x)=e_i$.*

3) *If* $A, B\in\mathscr{V}$ *there exist at most finitely many "paths" from* A *to* B, *that is, sequences* $(A, A_1, \ldots, A_{m-1}, B)$ *such that* $A_j\in\mathscr{V}$ *and there exist* $J_j\in\mathscr{A}$ *with* $A_1J_1A\neq 0,\ A_2J_2A_1\neq 0, \ldots, BJ_mA_{m-1}\neq 0$.

4) $\mathscr{A}$ *is equal to the set of all matrices* C *of* $\mathscr{M}(X, R)$ *such that: if* $x, y\in X$ *and* $C(y, x)\neq 0$ *then there exists* i, $1\leq i\leq n$, *and a "path" in* $\mathscr{A}$, *from* $A=\lambda(x, i)$ *to* $B=\lambda(y, i)$.

PROOF. Let Γ be a locally finite graph with $V(\Gamma)=X$. Let $\mathscr{A}=\operatorname{Im}(M^\Gamma)$, and for simplicity of notation, let $M=M^\Gamma$. The set $\mathscr{V}=\{M_{k_{v,i}}|v\in X,\ i=1, \ldots, n\}$ and the map λ defined by $\lambda(v, i)=M_{k_{v,i}}$ satisfy the above conditions (1) and (2) as seen in the proof of Proposition 7.

If $A=M_{k_{v,i}}, B=M_{k_{w,i}}$, if $(A, A_1, \ldots, A_{m-1}, B)$ with $A_j=M_{k_{v_j,l_j}}$ is a "path" of $\mathscr{A}$ from A to B, if $J_1, \ldots, J_m\in\mathscr{A}$ are such that $A_1J_1A\neq 0$, $A_2J_2A_1\neq 0, \ldots$ $\ldots, BJ_mA_{m-1}\neq 0$ then, by the proof of the proposition, $i=l_1=\ldots=l_{m-1}=l$ and since $J_j=M_{f_j}$ (for some $f_j\in\mathscr{F}$) there is a path $\varrho=(z_m, \ldots, z_2, z_1)$ in Γ with

$$o(z_1) = v,\ o(z_j) = v_{j-1} \qquad (j=2, \ldots, m), \qquad t(z_m) = w.$$

The correspondence $(A, A_1, \ldots, A_{m-1}, B)\to(\varrho, i)$ is injective, and since Γ is a locally finite graph, then condition (3) is satisfied.

Let $C\in\mathscr{M}(X, R)$ satisfy the condition indicated in (4). We shall define a function $f_C\colon P(\Gamma)\to R$. If $x, y\in X$ and $C(y, x)=0$, let $f_C(\eta)=0$ for every path η such that $o(\eta)=x$, $t(\eta)=y$. If $C(y, x)\neq 0$, by hypothesis there exists i, $1\leq i\leq n$ and a "path" in $\mathscr{A}=\operatorname{Im}(M^\Gamma)$ from $A=\lambda(x, i)$ to $B_i=\lambda(y, i)$. So there exist $g_j\in\mathscr{F}(\Gamma, R)$ (for $j=1, \ldots, m$) and $v_j\in X$ (for $j=2, \ldots, m$) such that $M_{k_{v_2,i}}M_{g_1}M_{k_{x,i}}\neq$ $\neq 0, \ldots, M_{k_{y,i}}M_{g_m}M_{k_{v_m,i}}\neq 0$. Hence, there exists a path from x to y in Γ. We choose one such path η_{xy}. Let $f_C(\eta_{xy})=C(y, x)$ and $f_C(\eta)=0$ for every path η from x to y, $\eta\neq\eta_{xy}$. It follows that

$$M_{f_C}(y, x) = \sum_{\substack{\varrho \\ o(\varrho)=x,\, t(\varrho)=y}} f_C(\varrho) = C(y, x)$$

hence $C\in\operatorname{Im}(M^\Gamma)$.

On the other hand, let $C\in\operatorname{Im}(M^\Gamma)$, so $C=M_f$ with $f\in\mathscr{M}(\Gamma, R)$. If $C(y, x)\neq 0$ there exists i, $1\leq i\leq n$, such that $C(y, x)e_i=M_f(y, x)e_i\neq 0$. Hence $\Big(\sum\limits_{\substack{o(\eta)=x\\ t(\eta)=y}} f(\eta)\Big)e_i\neq 0$.

Thus, there exists a path $\eta=(z_m, \ldots, z_1)$, with $o(\eta)=x$, $t(\eta)=y$.

Let $o(z_j)=v_j$ (for $j=2, \ldots, m$), so $M_{k_{v_2,i}}M_{k_{z_1,i}}M_{k_{x,i}}\neq 0, \ldots, M_{k_{y,i}}M_{k_{z_m,i}}M_{k_{v_m,i}}\neq 0$. Thus there exists a "path" in $\operatorname{Im}(M^\Gamma)$ from $A_i=M_{k_{x,i}}$ to $B_i=M_{k_{y,i}}$.

Now we prove the converse. Let $\mathscr{A}$ be a subalgebra of $\mathscr{M}(X, R)$ satisfying the conditions (1) to (4).

Let $\Gamma=\{(x, y)|x, y\in X$ and there exist i, $1\leq i\leq n$ and $C\in\mathscr{A}$ such that $\lambda(y, i)C\lambda(x, i)\neq 0\}$. Let $V(\Gamma)=\{(x, x)|x\in X\}\cong X$.

Let $o(x, y)=x$, $t(x, y)=y$. Thus (Γ, X, o, t) is a graph. We shall show that $\mathscr{A}=\operatorname{Im}(M^\Gamma)$.

Given $C\in\mathscr{A}$, let $f_C\colon P(\Gamma)\to R$ be defined as follows. If $x, y\in X$, choose a path η_{xy} such that $o(\eta)=x$, $t(\eta)=y$ (if one such path exists). Define $f_C(\eta_{xy})=$

$=C(y, x), f_C(\eta)=0$ for every $\eta \neq \eta_{xy}$ such that $o(\eta)=x$, $t(\eta)=y$. Then

$$M_{f_C}(y, x) = \sum_{\substack{o(\eta)=x \\ t(\eta)=y}} f_C(\eta) = C(y, x)$$

if there is a path from x to y. If, however, no such path exists then $M_{f_C}(y, x)=0$, but also $C(y, x)=0$ as follows from (4). Indeed, if $C(y, x)\neq 0$, since $C\in\mathscr{A}$ there exists i, $1\leq i\leq n$ and a "path" in $\mathscr{A}$, from $A=\lambda(x, i)$ to $B=\lambda(y, i)$, say $(A, A_2, \ldots, A_{m-1}, B)$, with $\lambda(A_j)=(v_j, l_j)$ for $j=2, \ldots, m-1$, it follows that $l_j=i$ (for every j) and that $\eta=\big((v_{m-1}, y), \ldots, (v_1, v_2), (x, v_1)\big)$ is a path in Γ from x to y, against the hypothesis. Thus, we have shown that $C=M_{f_C}\in \operatorname{Im}(M^\Gamma)$.

Conversely, given $f: P(\Gamma)\to R$, we show that $M_f\in\mathscr{A}$; it suffices to show that M_f satisfies the condition indicated in (4). Let

$$M_f(y, x) = \sum_{\substack{o(\eta)=x \\ t(\eta)=y}} f(\eta) \neq 0,$$

so there exists a path $\eta=\big((v_{m-1}, y), \ldots, (v_1, v_2), (x, v_1)\big)$ in Γ. Hence, as before there exists i, $1\leq i\leq n$, and a "path" $(A, A_2, \ldots, A_{m-1}, B)$ in $\mathscr{A}$, with $A=\lambda(x, i)$, $B=\lambda(y, i)$. By (4) $M_f\in\mathscr{A}$. ∎

4. The algebra of invariant functions on a semi-affine graph

We recall the definition of a semi-affine graph (see [1]). A *semi-affine graph* is a graph $(\Gamma, V(\Gamma), o, t)$ such that

1) $V(\Gamma)$ is a semigroup with unit element u;

2) $V(\Gamma)$ operates at the right and at the left of Γ, with the following properties: the operation extends the given operation in $V(\Gamma)$; $(vv')z=v(v'z)$, $z(vv')=(zv)v'$, $(vz)v'=v(zv')$, $o(vz)=vo(z)$, $o(zv)=o(z)v$, $t(vz)=vt(z)$, $t(zv)=t(z)v$ (where $v, v'\in V(\Gamma)$, $z\in\Gamma$).

These operations may be extended to paths, namely, if $\eta=(z_n, \ldots, z_1)\in P(\Gamma)$ then $v\eta=(vz_n, \ldots, vz_1)$, $\eta v=(z_n v, \ldots, z_1 v)$.

If moreover $V(\Gamma)$ is a group then the graph is an *affine graph*. In case $V(\Gamma)$ is abelian and $vz=zv$ for every $v\in V(\Gamma)$, $z\in\Gamma$, we call it an *abelian semi-affine* (or *affine*) *graph*.

Every semigroup graph is a semi-affine graph, every group-graph is an affine graph (see [1] for these definitions).

Let S be a partially ordered semigroup with unit element u, satisfying the following conditions.

1) $u\leq x$ for every $x\in S$;

2) if $x\leq y$ then $vx\leq vy$ and $xv\leq yv$ for every $v\in S$.

Then the associated graph $\Gamma(S)$ is a semi-affine graph, with $V(\Gamma(S))=S$. It is combinatorial, it has no circuits and for every $x\in S$ there is a path η such that $o(\eta)=u$, $t(\eta)=x$. If S is abelian then $\Gamma(S)$ is an abelian semi-affine graph. If S is a partially ordered group with above conditions then $\Gamma(S)$ is an affine graph.

On the other hand, if Γ is a semi-affine graph with no circuits and such that for every $x\in V(\Gamma)$ there is a path η, with $o(\eta)=u$, $t(\eta)=x$, then the associated partially ordered set $(S(\Gamma), \leqq)$ (see § 2) is a partially ordered semigroup (with properties (1), (2) above).

If Γ is a locally finite semi-affine graph, and R is a commutative ring, let $\mathscr{A}=\mathscr{A}(\Gamma, R)$ be the subset of $\mathscr{F}=\mathscr{F}(\Gamma, R)$ consisting of all functions $f: P(\Gamma)\to R$ such that $f(\eta v)=f(v\eta)=f(\eta)$ for every $v\in V(\Gamma)$, $\eta\in P(\Gamma)$. In particular, $f(v)=$ $=f(u)$ for every $v\in V(\Gamma)$.

It is easy to verify that $\mathscr{A}$ is closed under the algebra operations of scalar multiplication, $+$, $*$; thus $\mathscr{A}$ is an R-subalgebra of $\mathscr{F}$. Moreover the unit function e, the Riemann function and the Möbius function of Γ belong to $\mathscr{A}$ (see § 1).

Proposition 9. *If $f\in\mathscr{A}$ and f is invertible in $\mathscr{F}$ then the inverse f^{-1} belongs to $\mathscr{A}$.*

Proof. Let $g=f^{-1}$. We show by induction on $l(\eta)$ that $g(\eta v)=g(v\eta)=g(\eta)$ for every $v\in V(\Gamma)$. It is obvious when $l(\eta)=0$. If $\eta=(z_n, \ldots, z_1)$, with $n\geqq 1$, then

$$0=e(v\eta)=\sum_{v\eta=\varrho''\circ\varrho'} g(\varrho'')f(\varrho').$$

But $\varrho''=v\eta''$, $\varrho'=v\eta'$ with $\eta=\eta''\circ\eta'$. Hence

$$0=\sum_{\eta=\eta''\circ\eta'} g(v\eta'')f(\eta') \quad \text{and also} \quad 0=e(\eta)=\sum_{\eta=\eta''\circ\eta'} g(\eta'')f(\eta').$$

Since $g(v\eta'')=g(\eta'')$ when $l(\eta'')\leqq n-1$, then from

$$g(v\eta)f(o(\eta))+\sum_{\substack{\eta=\eta''\circ\eta'\\ l(\eta'')\leqq n-1}} g(v\eta'')f(\eta')=$$

$$=g(\eta)f(o(\eta))+\sum_{\substack{\eta=\eta''\circ\eta'\\ l(\eta'')\leqq n-1}} g(\eta'')f(\eta')$$

it follows that $g(v\eta)f(o(\eta))=g(\eta)f(o(\eta))$; since $f(o(\eta))$ is invertible, then $g(v\eta)=$ $=g(\eta)$. Similarly, $g(\eta v)=g(\eta)$. ∎

Now we shall indicate sufficient conditions for the algebra $\mathscr{A}$ to be commutative.

Let Γ be an abelian semi-affine graph. It is called *division-closed* when the following conditions are satisfied:

1) if $x, y\in V(\Gamma)$ and there exists a path η such that $o(\eta)=x$, $t(\eta)=y$, then there exists $v\in V(\Gamma)$ such that $vx=y$;

2) if $x, y, v\in V(\Gamma)$, if there exists a path η such that $o(\eta)=vx$, $t(\eta)=vy$ then there exists a unique path η', such that $o(\eta')=x$, $t(\eta')=y$, $v\eta'=\eta$.

Proposition 10. *If Γ is a division-closed abelian semi-affine graph then $\mathscr{A}$ is a commutative algebra.*

Proof. Let $f, g\in\mathscr{A}$, let η be a path of Γ, $x=o(\eta)$, $y=t(\eta)$. Then

$$(f*g)(\eta)=\sum_{\eta=\eta''\circ\eta'} f(\eta'')g(\eta').$$

For any decomposition $\eta=\eta''\circ\eta'$, let $v=o(\eta'')=t(\eta')$. By hypothesis there exists $y'\in V(\Gamma)$ such that $y=vy'$ and there exists a path ϱ_0'' such that $o(\varrho_0'')=u$,

$t(\varrho_0'')=y'$ and $\eta''=v\varrho_0''$. It follows that $\varrho''=x\varrho_0''$ is a path such that $o(\varrho'')=x$, $t(\varrho'')=xy'$. Since $f\in\mathscr{A}$ then $f(\varrho'')=f(\varrho_0'')=f(\eta)$. Similarly, $\varrho'=y'\eta'$ is a path such that $o(\varrho')=xy'$, $t(\varrho')=vy'=y$, and $g(\varrho')=g(\eta')$. This establishes a bijection between the set of pairs (η'',η') such that $o(\eta'')=t(\eta')$, $\eta=\eta''\circ\eta'$, and the set of pairs (ϱ',ϱ'') such that $o(\varrho')=t(\varrho'')$, $\eta=\varrho'\circ\varrho''$. Hence

$$(f*g)(\eta)=\sum_{\eta=\eta''\circ\eta'} f(\eta'')g(\eta')=\sum_{\eta=\varrho'\circ\varrho''} g(\varrho')f(\varrho'')=(g*f)(\eta). \quad \blacksquare$$

We note that if $\mathbf{N}$ is the set of positive integers, partially ordered by divisibility, then the associated graph $\Gamma(\mathbf{N})$ is a division-closed abelian semi-affine graph.

If R is a $\mathbf{Q}$-algebra, we may identify the R-algebra of arithmetic functions with the set $\mathscr{A}_0$ of all functions $f\in\mathscr{A}$ such that if ϱ,η are paths of $\Gamma(\mathbf{N})$ and $o(\varrho)=o(\eta)$, $t(\varrho)=t(\eta)$, then $f(\eta)=f(\varrho)$.

Indeed, given the arithmetic function $f:\mathbf{N}\to R$ let $\varphi(f)=\tilde f: P(\Gamma(\mathbf{N}))\to R$ be defined by $\tilde f(\eta)=\dfrac{f(m)}{\pi(m)}$ where $o(\eta)=k$, $t(\eta)=km$ and $\pi(m)$ denotes the number of paths in $\Gamma(\mathbf{N})$ with origin 1 and terminal m.

Then $\tilde f(r\eta)=\tilde f(\eta)$ for every path η and $r\in\mathbf{N}$; so $\tilde f\in\mathscr{A}$. Moreover, if ϱ,η are paths with the same extremities, then $\tilde f(\varrho)=\tilde f(\eta)$. Clearly $\widetilde{f+g}=\tilde f+\tilde g$, φ is injective. Also, φ is surjective: given $g\in\mathscr{A}_0$, let f be the arithmetic function $f(m)=g((1,m))\pi(m)$ (where $(1,m)$ denotes the edge with origin 1, terminal m). If $o(\eta)=k$, $t(\eta)=km$, then $\tilde f(\eta)=\dfrac{f(m)}{\pi(m)}=g((1,m))=g((k,km))=g(\eta)$.

Finally, we evaluate $\widetilde{f*g}$ and $\tilde f*\tilde g$. If $d|m$ let $\pi(m;d)$ denote the number of paths from 1 to m having a vertex equal to d. Thus $\pi(m;d)=\pi\left(\dfrac{m}{d}\right)\pi(d)$, because $\Gamma(\mathbf{N})$ is division-closed. We have:

$$\sum_{\substack{\varrho\\ o(\varrho)=k\\ t(\varrho)=km}} \widetilde{f*g}(\varrho)=\frac{1}{\pi(m)}\sum_{\varrho}(f*g)(m)=\frac{1}{\pi(m)}\sum_{\varrho}\sum_{d|m} f\left(\frac{m}{d}\right)g(d)=$$

$$=\sum_{d|m} f\left(\frac{m}{d}\right)g(d),$$

while

$$\sum_{\substack{\varrho\\ o(\varrho)=k\\ t(\varrho)=km}} (\tilde f*\tilde g)(\varrho)=\sum_{\varrho}\sum_{\varrho=\varrho''\circ\varrho'}\tilde f(\varrho'')\tilde g(\varrho')=$$

$$=\sum_{\varrho}\sum_{\substack{d\\ kd \text{ is a vertex of } \varrho}}\frac{f\left(\frac{m}{d}\right)}{\pi\left(\frac{m}{d}\right)}\times\frac{g(d)}{\pi(d)}=\sum_{d|m}\sum_{\substack{\varrho\\ kd \text{ is a vertex of } \varrho}}\frac{f\left(\frac{m}{d}\right)}{\pi\left(\frac{m}{d}\right)}\frac{g(d)}{\pi(d)}=$$

$$=\sum_{d|m}\frac{1}{\pi(m;d)}\sum_{\substack{\varrho\\ kd \text{ is a vertex of } \varrho}} f\left(\frac{m}{d}\right)g(d)=\sum_{d|m} f\left(\frac{m}{d}\right)g(d).$$

So

$$\sum_{\substack{\varrho \\ o(\varrho)=k \\ t(\varrho)=km}} \widetilde{f*g}(\varrho) = \sum_{\substack{\varrho \\ o(\varrho)=k \\ t(\varrho)=km}} (\tilde{f}*\tilde{g})(\varrho).$$

Hence if $o(\eta)=k$, $t(\eta)=km$, since $\widetilde{f*g}\in\mathscr{A}_0$ then

$$\widetilde{f*g}(\eta) = \frac{1}{\pi(m)} \sum_{\substack{\varrho \\ o(\varrho)=k \\ t(\varrho)=km}} (\tilde{f}*\tilde{g})(\varrho)$$

(note that, a priori, $\tilde{f}*\tilde{g}$ need not to belong to $\mathscr{A}_0$, but it belongs to $\mathscr{A}$).

$\mathscr{A}_0$ becomes an R-algebra, isomorphic to the algebra of arithmetic functions, if we define the averaged convolution $\tilde{f}*\tilde{g}$ by

$$(\tilde{f}*\tilde{g})(\eta) = \frac{1}{\pi(m)} \sum_{\substack{\varrho \\ o(\varrho)=k \\ t(\varrho)=km}} (\tilde{f}*\tilde{g})(\varrho).$$

We note also that if $I:\mathscr{F}\to\mathscr{I}$, as defined in § 2, then $I(\widetilde{f*g})=I(\tilde{f}*\tilde{g})$. Indeed

$$\big(I(\widetilde{f*g})\big)[k, km] = \sum_{\substack{\varrho \\ o(\varrho)=k \\ t(\varrho)=km}} (\widetilde{f*g})(\varrho) = \sum_{\substack{\varrho \\ o(\varrho)=k \\ t(\varrho)=km}} (\tilde{f}*\tilde{g})(\varrho) = \big(I(\tilde{f}*\tilde{g})\big)[k, km].$$

Hence $\widetilde{f*g}=\tilde{f}*\tilde{g}+\mathrm{Ker}\,(I)\cap\mathscr{A}$.

In general, $\mathrm{Ker}\,(I)\cap\mathscr{A}$ might be different from 0. But $\mathrm{Ker}\,(I)\cap\mathscr{A}_0=0$ because if $f\in\mathscr{A}_0$, $If=0$, then

$$0 = (If)[k, km] = \sum_{\substack{\varrho \\ o(\varrho)=k \\ t(\varrho)=km}} f(\varrho) = \pi(m) f\big((k, km)\big)$$

so $\big((k, km)\big)=0$ hence $f(\eta)=0$, showing that $f=0$.

REFERENCES

[1] Delsarte, S., Fonctions de Möbius sur les groupes abeliens finis, *Ann. of Math.* (2) **49** (1948), 600—609. *MR* **10**—9.

[2] Content, M., Lemay, F. and Leroux, P., Catégories de Möbius et fonctorialités: un cadre général pour l'inversion de Möbius, *J. Combin. Theory Ser. A* **28** (1980), 169—190. *MR* **82f**: 18002.

[3] Leroux, P., The isomorphism problem for incidence algebras of Möbius categories (preprint).

[4] Leroux, P. and Sarraillé, J., Structure of incidence algebras of graphs, *Comm. Algebra* (to appear).

[5] Rédei, L., Zetafunktionen in der Algebra, *Acta Math. Acad. Sci. Hungar.* **6** (1955), 5—25 *MR* **17**—344.

[6] Ribenboim, P., Algebraic structures on graphs, *Algebra Universalis* **16** (1983), 105—123.

[7] Rota, G.-C., On the foundations of combinatorial theory, I, Theory of Möbius functions, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **2** (1964), 340—368. *MR* **30** # 4688.

[8] STANLEY, R. P., Structure of incidence algebras and their automorphism groups, *Bull. Amer. Math. Soc.* **76** (1970), 1236—1239. *MR* **41** # 8319.
[9] WARD, M., The algebra of lattice functions, *Duke Math. J.* **5** (1939), 357—371. *Zbl* **22**, 123.
[10] WEISNER, L., Abstract theory of inversion of finite series, *Trans. Amer. Math. Soc.* **38** (1935), 474—484. *Zbl* **12**, 393.
[11] WIEGANDT, R., On the general theory of Möbius inversion formula and Möbius product, *Acta Sci. Math. (Szeged)* **20** (1959), 164—180. *MR* **21** # 7176.

(Received February 19, 1982)

DEPARTMENT OF MATHEMATICS AND STATISTICS
JEFFERY HALL
QUEEN'S UNIVERSITY
KINGSTON
K7L 3N6
CANADA

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 21—43

DENSITY RESULTS FOR THE 2-CLASSGROUPS AND FUNDAMENTAL UNITS OF REAL QUADRATIC FIELDS

PATRICK MORTON[1]

Dedicated to the memory of László Rédei

1. Introduction

In a previous paper [7] we have considered the 2-classgroup $\mathcal{C}$ of the imaginary quadratic field $Q(\sqrt{-Dq})$, where D is a fixed positive integer and q is a variable prime, and have shown that the structure of $\mathcal{C}/\mathcal{C}^8$ depends only on the splitting of q in a fixed normal extension Σ_D of Q.

In this paper we carry out a similar analysis for the *real* quadratic fields $Q(\sqrt{Dq})$, where D is a product of r primes p_i,

(1.1) $$D = p_1 \dots p_r,$$

satisfying the conditions

(1.2) $$p_i \equiv 1 \pmod 8, \quad \left(\frac{p_i}{p_j}\right) = +1 \quad \text{for} \quad i \neq j,$$

and q is a prime distinct from the p_i with

(1.3) $$q \equiv 1 \pmod 4).$$

This is a natural collection of fields to consider, since if D is fixed, all the fields $Q(\sqrt{Dq})$ share (roughly speaking) the same group of quadratic characters (see § 2).

Our analysis will show, as in the imaginary case, that the restricted (narrow) 2-classgroup $\mathcal{C}_q$ of $Q(\sqrt{Dq})$ has a quotient $\mathcal{C}_q/\mathcal{C}_q^8$ which depends only on the behavior of q in a suitable normal extension of Q. As a corollary we deduce the following theorem of Rédei [10].

THEOREM A. *Let $\mathcal{T}$ be a finite abelian 2-group of exponent at most 8. Then there are infinitely many real quadratic fields whose restricted 2-classgroups $\mathcal{C}$ satisfy the isomorphism*

$$\mathcal{C}/\mathcal{C}^8 \cong \mathcal{T}.$$

The principal motivation for considering this class of real fields in detail is the interest which attaches to the Pell equation

(1.4) $$x^2 - Dqy^2 = -1.$$

---

[1] Portions of this paper are taken from the author's Ph. D. thesis, written at the University of Michigan, 1979. The author gratefully acknowledges the support of an NSF fellowship during the period this article was written.

1980 *Mathematics Subject Classification.* Primary 12A25; Secondary 12A75.

Key words and phrases. Restricted 2-classgroup, genus characters, Frobenius density theorem, Pell equation.

As is well-known, this equation has an integer solution if and only if the fundamental unit η_q of $Q(\sqrt{Dq})$ has norm -1. There is an intimate connection between the value of Norm η_q and the structure of $\mathscr{C}_q$, which Rédei exploited to give an algorithm for computing Norm η_q without continued fractions (see [11], [6]). Using a simplified version of Rédei's approach and our analysis of $\mathscr{C}_q/\mathscr{C}_q^8$ we prove the following density result concerning η_q. (The symbol $\left(\frac{a}{p}\right)_4$ occurring below is the biquadratic residue symbol, defined for quadratic residues a of p by $\pm 1 = \left(\frac{a}{p}\right)_4 \equiv$ $\equiv a^{(p-1)/4} (\bmod p)$.)

THEOREM B. *Assume* D *is given by* (1.1), (1.2). *Let* $\mathscr{T}$ *be a finite abelian 2-group of rank* r *and exponent* 4, *and let* $\varepsilon = \pm 1$. *Then the primes* q, *for which*

$$\mathscr{C}_q \cong \mathscr{T} \quad \text{and} \quad \text{Norm } \eta_q = \varepsilon,$$

have positive density.

The densities occurring in this theorem can be given explicitly (see § 5 and the tables in § 6) and turn out to be rational numbers with denominator a power of 2. This result is an interesting counterpart to a theorem of Rédei [11] (see also [6]), which asserts that Norm $\eta_q = -1$ whenever $\mathscr{C}_q$ has exponent 2. Theorem B shows that in general the structure of $\mathscr{C}_q$ alone does not determine the signature of η_q.

From Theorem B we also deduce the following corollary concerning the set S_D of primes q for which the equation (1.4) is solvable, in case the prime factors of D satisfy the condition

$$\left(\frac{p_i}{p_j}\right)_4 = \left(\frac{p_j}{p_i}\right)_4, \quad \text{for } i \neq j. \tag{1.5}$$

THEOREM C. *Assume that* D *satisfies* (1.1), (1.2), *and* (1.5), *and let* $\underline{\delta}(D)$ *and* $\bar{\delta}(D)$ *denote the lower and upper densities of the set* S_D. *Then*

$$\frac{\theta}{2}\left(\frac{3}{4}\right)^r \lesssim \underline{\delta}(D) \leqq \bar{\delta}(D) \leqq \frac{1}{2}\left(\frac{3}{4}\right)^r, \quad \textit{as} \quad r \to \infty, \tag{1.6}$$

where

$$\theta = \prod_{k=1}^{\infty}\left(1 - \frac{1}{2^{2k+1}}\right) = .4194\ldots.$$

Hence (1.4) is not usually solvable if D has a large number of prime factors. This contrasts sharply with Rédei's result [9] that

$$x^2 - \Delta y^2 = -1$$

has a solution for the majority of discriminants Δ, none of whose prime factors is $\equiv 3 \pmod 4$.

It is interesting in this context that the correct order of magnitude of $\underline{\delta}(D)$ and $\bar{\delta}(D)$ in (1.6) has been gained through an investigation of $\mathscr{C}_q/\mathscr{C}_q^8$. It would be possible to sharpen (1.6) if the structure of $\mathscr{C}_q/\mathscr{C}_q^{2^n}$ could be determined theoretically for $n \geqq 4$. This also raises the question: does the set S_D have a well-defined density? The answer to this question seems to lie well beyond our present methods.

2. Character evaluation

Let e_{2^n} denote the 2^n-rank of the restricted 2-classgroup $\mathscr{C}_q$ of $\Omega = Q(\sqrt{Dq})$, i.e., the number of invariants of $\mathscr{C}_q$ divisible by 2^n. As in [7] we shall compute e_2, e_4 and e_8 using the quadratic character group X_2 of $\mathscr{C}_q$. As is well-known [3], the following r Hilbert symbols form a basis for X_2:

$$\chi_i(\mathfrak{a}) = \left(\frac{\text{Norm } \mathfrak{a}, Dq}{p_i}\right), \quad i = 1, \ldots, r, \tag{2.1}$$

where $\mathfrak{a}$ is an ideal of Ω, and where for technical reasons we assume

$$\left(\frac{p_i}{q}\right) = +1 \quad \text{for} \quad 1 \leqq i \leqq s, \qquad \left(\frac{p_i}{q}\right) = -1 \quad \text{for} \quad s < i \leqq r. \tag{2.2}$$

Since there are r basis characters, the 2-rank e_2 is given by

$$e_2 = r. \tag{2.3}$$

To compute e_4, we let $\mathfrak{p}_i$ (for $1 \leqq i \leqq r$) and $\mathfrak{q}$ denote the prime ideals of Ω lying above the primes p_i and q, respectively. Also, define the map ξ: $\{\pm 1\} \to \mathbf{F}_2$ by

$$\xi(1) = 0, \quad \xi(-1) = 1. \tag{2.4}$$

Then e_4 is r minus the rank over $\mathbf{F}_2$ of the $(r+1) \times r$ matrix $M = (\varepsilon_{ij})$, where

$$\varepsilon_{ij} = \begin{cases} \xi\chi_j(\mathfrak{p}_i), & i \leqq r, \\ \xi\chi_j(\mathfrak{q}), & i = r+1. \end{cases}$$

From (1.2) and (2.2) the rank of M is $r-s$, so that

$$e_4 = s. \tag{2.5}$$

Finally, e_8 is obtained as follows. The form of M shows that

$$\chi_j(\mathfrak{p}_i) = +1 \quad \text{for} \quad 1 \leqq i \leqq s \quad \text{and all} \quad j,$$

and

$$\chi_j(\mathfrak{p}_{s+1} \cdots \mathfrak{p}_r \mathfrak{q}) = +1 \quad \text{for all} \quad j.$$

Hence there are ideals $\mathfrak{z}_i$ of Ω $(1 \leqq i \leqq s+1)$ satisfying

$$\mathfrak{z}_i^2 \sim \mathfrak{p}_i, \quad \text{for} \quad 1 \leqq i \leqq s, \qquad \mathfrak{z}_{s+1}^2 \sim \mathfrak{p}_{s+1} \cdots \mathfrak{p}_r \mathfrak{q}. \tag{2.6}$$

(The symbol $\sim$ denotes restricted equivalence, so $\mathfrak{a} \sim \mathfrak{b}$ iff $\mathfrak{a}\mathfrak{b}^{-1} = (\eta)$ with Norm $\eta > 0$.) In terms of the $\mathfrak{z}_i$, it is not hard to see that

$$e_8 = s - \varrho, \tag{2.7}$$

where ϱ is the rank over $\mathbf{F}_2$ of the matrix

$$M' = (\xi\chi_j(\mathfrak{z}_i)) \qquad (1 \leqq i \leqq s+1,\ 1 \leqq j \leqq s)$$

(see [7], § 2).

In order to compute M' we need a simple preliminary lemma, whose proof can be found in [7].

Lemma 1. *Assume that the ideal*

$$\mathfrak{a} = \mathfrak{q}^{a_0} \prod_{i=1}^{r} \mathfrak{p}_i^{a_i}, \quad a_i = 0 \quad or \quad 1,$$

is equivalent to a square in Ω. *If* $a = \text{Norm}\, \mathfrak{a}$, *and* (x, y, z) *is a positive primitive solution of*

(2.8) $$x^2 - Dqy^2 - 4az^2 = 0,$$

then there is an ideal $\mathfrak{b}$ *for which* $\mathfrak{b}^2 \sim \mathfrak{a}$ *and* Norm $\mathfrak{b} = z$.

The computation of M' is contained in the following three lemmas.

Lemma 2. *For* $1 \leqq i \leqq s$ *we have*

(2.9) $$\chi_i(\mathfrak{z}_i) = \left(\frac{D/p_i}{p_i}\right)_4 \left(\frac{q}{p_i}\right)_4.$$

Proof. The conditions (1.2) and (2.2) show that the field $Q(\sqrt{p_i})$ contains an ideal $\mathfrak{a}$ of norm $\frac{D}{p_i} q$. Since the fundamental unit of $Q(\sqrt{p_i})$ has norm -1, it follows that the ideal $\mathfrak{a}^{h_i} = (\eta)$ has a generator η with norm $-\left(\frac{D}{p_i} q\right)^{h_i}$, where h_i is the class number of $Q(\sqrt{p_i})$. We deduce that

(2.10) $$-\left(\frac{D}{p_i} q\right)^{h_i} = z'^2 - p_i x'^2$$

for some $z', x' \in \mathbf{Z}$, by a simple congruence argument modulo 8, and then that z' is even, by considering (2.10) modulo 4. Thus we have a solution to equation (2.8) with $a = p_i$, $x = p_i x'$, $y = \left(\frac{D}{p_i} q\right)^{(h_i-1)/2}$ and $z = z'/2$. (Recall that h_i is odd.) Hence by Lemma 1, and the fact that $\chi_i(\mathfrak{z}_i)$ is independent of the choice of $\mathfrak{z}_i$, we see that

$$\chi_i(\mathfrak{z}_i) = \left(\frac{z}{p_i}\right) = \left(\frac{2}{p_i}\right)\left(\frac{z'^2}{p_i}\right)_4 = \left(\frac{qD/p_i}{p_i}\right)_4^{h_i} = \left(\frac{qD/p_i}{p_i}\right)_4,$$

using $p_i \equiv 1 \pmod 8$. Q.E.D.

Lemma 3. *For* $1 \leqq i, j \leqq s$ *and* $i \neq j$, *let* $\mathfrak{d}_{ij}$ *and* $\mathfrak{q}_{ij}$ *be integral ideals of* $Q(\sqrt{p_i p_j})$ *satisfying*

(2.11) $$\text{Norm}\ \mathfrak{d}_{ij} = \frac{D}{p_i p_j}, \quad \text{Norm}\ \mathfrak{q}_{ij} = q.$$

Let $\mathfrak{r}_{ij}$ *be an integral ideal of* $Q(\sqrt{p_i p_j})$ *for which* $(\mathfrak{r}_{ij}, 2Dq\mathfrak{r}'_{ij})=1$ *and*

(2.12) $$\mathfrak{d}_{ij}\mathfrak{q}_{ij}\mathfrak{r}_{ij}^2 \sim (\sqrt{p_i p_j}) \quad in \quad Q(\sqrt{p_i p_j}).$$

(As usual, $\mathfrak{a}'$ *denotes the conjugate of the ideal* $\mathfrak{a}$.*) Then if* Norm $\mathfrak{r}_{ij}=r_{ij}$, *we have*

(2.13) $$\chi_j(\mathfrak{z}_i) = \left(\frac{p_i}{p_j}\right)_4 \left(\frac{r_{ij}}{p_j}\right).$$

PROOF. It is clear from the conditions (1.2) and (2.2) that each prime divisor of $Dq/p_i p_j$ splits in $Q(\sqrt{p_i p_j})$, so there exist ideals $\mathfrak{d}_{ij}$ and $\mathfrak{q}_{ij}$ satisfying (2.11). Moreover, by the genus theory in $Q(\sqrt{p_i p_j})$ and the fact that

$$\left(\frac{D/p_i p_j}{p_i}\right) = \left(\frac{q}{p_i}\right) = 1 \quad (\text{for } 1 \leqq i \leqq s),$$

we see that $\mathfrak{d}_{ij}$ and $\mathfrak{q}_{ij}$ are equivalent to squares in $Q(\sqrt{p_i p_j})$. This is also true of the ideal $(\sqrt{p_i p_j})$ since

$$\left(\frac{\text{Norm}(\sqrt{p_i p_j}),\, p_i p_j}{p_i}\right) = \left(\frac{p_i p_j,\, -p_i p_j}{p_i}\right)\left(\frac{-1}{p_i}\right) = +1.$$

Hence (2.12) is solvable, and the ideal $\mathfrak{d}_{ij}\mathfrak{q}_{ij}\mathfrak{r}_{ij}^2$ has a generator $\eta=z'+x'\sqrt{p_i p_j}$ with negative norm. (Note that η has the given form since its norm is odd and $p_i p_j \equiv 1 \pmod 8$.) Thus

$$-\frac{D}{p_i p_j} q r_{ij}^2 = z'^2 - p_i p_j x'^2,$$

where as before, $z'=2z$ is even. Consequently, $(x, y, z)=(p_i p_j x', r_{ij}, z)$ is a primitive solution to the equation

$$x^2 - Dqy^2 - 4p_i p_j z^2 = 0,$$

and Lemma 1 shows that $\mathfrak{p}_i\mathfrak{p}_j \sim \mathfrak{b}^2$, where Norm $\mathfrak{b}=z$. We deduce that

$$\chi_j(\mathfrak{z}_i) = \chi_j(\mathfrak{z}_i\mathfrak{z}_j)\chi_j(\mathfrak{z}_j) = \chi_j(\mathfrak{b})\chi_j(\mathfrak{z}_j) =$$

$$= \left(\frac{z}{p_j}\right)\chi_j(\mathfrak{z}_j) = \left(\frac{z^2}{p_j}\right)_4 \chi_j(\mathfrak{z}_j) =$$

$$= \left(\frac{Dq/p_i p_j}{p_j}\right)_4 \left(\frac{r_{ij}}{p_j}\right)\left(\frac{D/p_j}{p_j}\right)_4 \left(\frac{q}{p_j}\right)_4 = \left(\frac{p_i}{p_j}\right)_4 \left(\frac{r_{ij}}{p_j}\right).$$ Q.E.D.

The computation of M' is completed in

LEMMA 4. *Let* $d=p_1 \dots p_s$, *and let* $\bar{\mathfrak{d}}$ *and* $\bar{\mathfrak{q}}$ *denote integral ideals of* $Q(\sqrt{-d})$ *with norms*

(2.14) $$\text{Norm}\,\bar{\mathfrak{d}} = D/d, \quad \text{Norm}\,\bar{\mathfrak{q}} = q.$$

If $\bar{\mathfrak{r}}$ *is an integral ideal of* $Q(\sqrt{-d})$ *satisfying*

(2.15) $$\bar{\mathfrak{d}}\bar{\mathfrak{q}}\bar{\mathfrak{r}}^2 \sim 1 \quad in \quad Q(\sqrt{-d}),$$

where $(\bar{\mathfrak{r}}, 2Dq\bar{\mathfrak{r}}')=1$, *and if* $\bar{r}=\text{Norm } \bar{\mathfrak{r}}$, *then*

(2.16) $$\chi_j(\mathfrak{z}_{s+1}) = \left(\frac{qD/d}{p_j}\right)_4 \left(\frac{\bar{r}}{p_j}\right), \quad for \quad 1 \leqq j \leqq s.$$

PROOF. The existence of $\bar{\mathfrak{d}}, \bar{\mathfrak{q}}$ and $\bar{\mathfrak{r}}$ is proved as in Lemma 3, using the fact that the Hilbert symbols

(2.17) $$\bar{\chi}_j(\mathfrak{a}) = \left(\frac{\text{Norm}\,\mathfrak{a}, -4d}{p_j}\right), \qquad 1 \leqq j \leqq s,$$

form a basis for the quadratic characters on the classgroup of $Q(\sqrt{-d})$. We conclude that

$$\bar{\mathfrak{d}}\bar{\mathfrak{q}}\bar{\mathfrak{r}}^2 = (z+y\sqrt{-d})$$

for some z, y in $\mathbf{Z}$, and rearranging the resulting norm equation gives

$$\left(\frac{D}{d}q\bar{r}\right)^2 - Dqy^2 - \frac{D}{d}qz^2 = 0.$$

Now $\dfrac{D}{d}q$ is the norm of the ideal $\mathfrak{p}_{s+1} \dots \mathfrak{p}_r\mathfrak{q}$ of Ω, so by Lemma 1 and (2.6) we may assume that $\mathfrak{z}_{s+1}$ satisfies

$$\text{Norm } \mathfrak{z}_{s+1} = \begin{cases} z/2, & \text{if } z \text{ is even} \\ z, & \text{if } z \text{ is odd.} \end{cases}$$

Since all p_j are $\equiv 1 \pmod 8$, we see as in our previous computations that

$$\chi_j(\mathfrak{z}_{s+1}) = \left(\frac{z}{p_j}\right) = \left(\frac{qD/d}{p_j}\right)_4 \left(\frac{\bar{r}}{p_j}\right), \qquad 1 \leqq j \leqq s,$$

and this proves the lemma.

3. The field $\bar{\Sigma}_d$

In this section we show that the dependence of the matrix M' on q can be described by the splitting of q in a normal extension of Q. We first introduce the following fields. (See the corresponding discussion in [7].)

For each i between 1 and s, we let K_i be the quartic subfield of the field of p_i-th roots of unity. Then K_i is the unique cyclic quartic extension of Q with conductor p_i, and corresponds to a quartic character $\psi_i \pmod{p_i}$ satisfying

(3.1) $$\psi_i(a) = \left(\frac{a}{p_i}\right)_4 \quad \text{if} \quad \left(\frac{a}{p_i}\right) = +1.$$

Moreover, K_i contains the subfield $Q(\sqrt{p_i})$; we denote the non-trivial automorphism of $K_i/Q(\sqrt{p_i})$ by σ_i.

Also, for each pair (i, j) with $1 \leqq i, j \leqq s$ and $i \neq j$, we let L_{ij} be the classfield over $Q(\sqrt{p_i p_j})$ corresponding to the subgroup of 4-th powers in the restricted ideal classgroup of $Q(\sqrt{p_i p_j})$. Since this classgroup has a cyclic 2-Sylow subgroup with order divisible by 4 (by the Rédei—Reichardt theorem [8]), L_{ij} is the unique cyclic quartic extension of $Q(\sqrt{p_i p_j})$ unramified at finite primes. Moreover, L_{ij} contains the field $Q(\sqrt{p_i}, \sqrt{p_j})$, which is the genus field of $Q(\sqrt{p_i p_j})$. We let λ_{ij} denote the non-trivial automorphism of $L_{ij}/Q(\sqrt{p_i}, \sqrt{p_j})$.

For later use we also note the following explicit representation of L_{ij} (for a proof see [7], Lemma 11). Let h_i be the classnumber of $Q(\sqrt{p_i})$, and let $\mathfrak{p}_{ij}$ be a prime ideal of $Q(\sqrt{p_i})$ lying over p_j. If

$$\mathfrak{p}_{ij}^{3h_i} = (\beta_{ij}), \quad \beta_{ij} = x + y\sqrt{p_i}, \tag{3.2}$$

where $\operatorname{Norm} \beta_{ij} = p_j^{3h_i} > 0$ and x is chosen so that

$$x \equiv \begin{cases} +1 \pmod 4, & \text{if } 4|y \\ -1 \pmod 4, & \text{if } 2|y \text{ but } 4 \nmid y, \end{cases}$$

then

$$L_{ij} = Q(\sqrt{p_i p_j}, \sqrt{\beta_{ij}}). \tag{3.3}$$

In terms of the splitting of q in K_i and L_{ij}, Lemmas 2 and 3 translate as follows. (The unexplained symbols are Artin symbols.)

LEMMA 5. *For $1 \leqq i, j \leqq s$ and $i \neq j$, we have*

$$\chi_i(\mathfrak{z}_i) = \left(\frac{D/p_i}{p_i}\right)_4 (-1)^{a_i}, \quad \text{where} \quad \left(\frac{K_i}{q}\right) = \sigma_i^{a_i}, \tag{3.4}$$

and

$$\chi_j(\mathfrak{z}_i) = \left(\frac{p_i}{p_j}\right)_4 (-1)^{b_{ij}}, \tag{3.5}$$

where

$$\left(\frac{L_{ij}/Q(\sqrt{p_i p_j})}{\mathfrak{q}_{ij}}\right) = \lambda_{ij}^{b_{ij}} \tilde{\delta}_{ij} \tag{3.6}$$

and $\tilde{\delta}_{ij}$ is the automorphism

$$\tilde{\delta}_{ij} = \left(\frac{L_{ij}/Q(\sqrt{p_i p_j})}{\mathfrak{d}_{ij}(\sqrt{p_i p_j})}\right). \tag{3.7}$$

In addition, $b_{ij} = b_{ji}$ for $i \neq j$.

PROOF. Formula (3.4) follows from (3.1) and (2.9) and the fact that the character value $\psi_i(q)$ is determined by the splitting of q in K_i. Formula (3.5) follows from (2.13) and the fact that the term $\left(\frac{\mathfrak{r}_{ij}}{p_j}\right)$ in that equation equals $+1$ if and only if

$$\mathfrak{r}_{ij} \sim \mathfrak{s}^2 \quad \text{for some ideal } \mathfrak{s} \text{ in } Q(\sqrt{p_i p_j}).$$

But this condition is equivalent to

$$\left(\frac{L_{ij}/Q(\sqrt{p_i p_j})}{\mathfrak{q}_{ij}}\right) = \tilde{\delta}_{ij}$$

by (2.12) and the definition of L_{ij}. (Note that the Artin symbols in (3.6) and (3.7) fix the field $Q(\sqrt{p_i}, \sqrt{p_j})$, and so are equal to 1 or λ_{ij}.)

Lemma 5 shows that the values $\chi_j(\mathfrak{z}_i)$, for $1 \leqq i, j \leqq s$, depend on the splitting of q in the field Σ_d which is the compositum of the K_i for $1 \leqq i \leqq s$ and the L_{ij} for $1 \leqq i < j \leqq s$. In order to prove a similar result for the values $\chi_j(\mathfrak{z}_{s+1})$, we make the following remarks concerning the classgroup $\mathscr{C}(d)$ in $Q(\sqrt{-d})$, where as before $d = p_1 \dots p_s$.

We note first that $\bar{\mathscr{C}}(d)$ has full 4-rank, i.e., has s invariants divisible by 4, by (1.2) and the Rédei—Reichardt theorem. If $\bar{X}$ is the group of characters on $\bar{\mathscr{C}}(d)$, it follows that every quadratic character in $\bar{X}$ has a square-root. If $\bar{\chi}_j$ is the character defined by (2.17), let ω_j be an element of $\bar{X}$ for which

$$\bar{\chi}_j = \omega_j^2 \qquad (\text{for } 1 \leqq j \leqq s). \tag{3.8}$$

Then ω_j corresponds (in the sense of classfield theory) to a cyclic quartic extension $\bar{L}_j$ of $Q(\sqrt{-d})$ inside the Hilbert classfield of $Q(\sqrt{-d})$. This field contains the field corresponding to $\bar{\chi}_j$, which is $Q(\sqrt{-d}, \sqrt{p_j})$ (by the transference theorem). Let the non-trivial automorphism of the quadratic extension $\bar{L}_j/Q(\sqrt{-d}, \sqrt{p_j})$ be denoted by $\bar{\lambda}_j$.

We may now state

LEMMA 6. *Let* $\bar{\delta}_j$ *be the automorphism*

$$\bar{\delta}_j = \left(\frac{\bar{L}_j/Q(\sqrt{-d})}{\bar{\mathfrak{d}}}\right), \tag{3.9}$$

where $\bar{\mathfrak{d}}$ *satisfies* (2.14). *Then* $\bar{\delta}_j = 1$ *or* $\bar{\lambda}_j$, *and for* $1 \leqq j \leqq s$ *we have*

$$\chi_j(\mathfrak{z}_{s+1}) = \left(\frac{D/d}{p_j}\right)_4 (-1)^{a_j + c_j}, \tag{3.10}$$

where

$$\left(\frac{\bar{L}_j/Q(\sqrt{-d})}{\bar{\mathfrak{q}}}\right) = \bar{\lambda}_j^{c_j} \bar{\delta}_j, \quad c_j = 0 \quad \text{or} \quad 1, \tag{3.11}$$

and a_j *is defined by* (3.4).

PROOF. We only need to consider the term $\left(\frac{\bar{r}}{p_j}\right)$ in (2.16). From (2.17), (3.8) and (2.15) we have

$$\left(\frac{\bar{r}}{p_j}\right) = \bar{\chi}_j(\bar{\mathfrak{r}}) = \omega_j(\bar{\mathfrak{r}}^2) = \omega_j(\bar{\mathfrak{d}}\bar{\mathfrak{q}}).$$

Hence $\left(\frac{\bar{r}}{p_j}\right)=1$ if and only if

$$\left(\frac{\bar{L}_j/Q(\sqrt{-d})}{\bar{\mathfrak{q}}}\right)=\bar{\delta}_j,$$

by classfield theory. Also, both automorphisms in this last equation fix $Q(\sqrt{-d}, \sqrt{p_j})$, since

$$\bar{\chi}_j(\bar{\mathfrak{d}})=\bar{\chi}_j(\bar{\mathfrak{q}})=+1.$$

This proves the lemma.

REMARK. It is easy to see that the ideal group in $Q(\sqrt{-d})$ corresponding to $\bar{L}_j$ is invariant under the automorphism $(\sqrt{-d}\to-\sqrt{-d})$. Hence a theorem of Hasse ([4], II, p. 24) implies that $\bar{L}_j$ is normal over Q.

We note that the field $\bar{L}_j$ of Lemma 6 may be chosen in several ways. Any cyclic quartic unramified extension of $Q(\sqrt{-d})$, which contains $Q(\sqrt{-d}, \sqrt{p_j})$, corresponds to a character ω_j satisfying (3.8) and may therefore serve as $\bar{L}_j$. However, the compositum $\bar{\Lambda}_d$ of the $\bar{L}_j$ (for $1\leqq j\leqq s$) is unique and depends only on d. In fact, $\bar{\Lambda}_d$ is the classfield over $Q(\sqrt{-d})$ corresponding to the subgroup $\mathscr{C}^4(d)$ of fourth powers in $\mathscr{C}(d)$, by (3.8) and the fact that the $\bar{\chi}_j$ generate the quadratic characters of $\mathscr{C}(d)$. Thus $\bar{\Lambda}_d$ is the counterpart to the field L_{ij}. We also note that $\bar{\Lambda}_d$ is the independent compositum over $Q(\sqrt{-d})$ of the fields $\bar{L}_j$, since the characters $\bar{\chi}_j$, and therefore also the ω_j, are independent in $\bar{X}$. Hence

$$[\bar{\Lambda}_d\colon Q(\sqrt{-d})]=4^s. \tag{3.12}$$

We fix a particular choice for $\bar{L}_j$ in

LEMMA 7. *For* $1\leqq i, j\leqq s$, $i\neq j$, *let* h_i *and* $\mathfrak{p}_{ij}$ *be as in* (3.2), *and let* $\mathfrak{l}_i$ *be a prime ideal of* $Q(\sqrt{p_i})$ *lying over* 2. *Set*

$$(\gamma_i)=\prod_{\substack{1\leqq j\leqq s\\ j\neq i}}\mathfrak{p}_{ij}^{h_i}, \tag{3.13}$$

where γ_i *is subject to the conditions*

$$\gamma_i\equiv 1 \pmod{\mathfrak{l}_i^2},\qquad \text{Norm}\,\gamma_i=-\left(\frac{d}{p_i}\right)^{h_i}. \tag{3.14}$$

Then, for $1\leqq i\leqq s$, *the field*

$$\bar{L}_i=Q(\sqrt{-d}, \sqrt{\gamma_i}) \tag{3.15}$$

is a cyclic quartic unramified extension of $Q(\sqrt{-d})$ *containing* $Q(\sqrt{-d}, \sqrt{p_i})$.

This lemma is an immediate consequence of Rédei's Satz 6 in [11]. It may also be proved in the same way as the corresponding Lemma 11 in [7]. Note that Norm $\mathfrak{l}_i=2$, so either γ_i or $-\gamma_i$ is $\equiv 1 \pmod{\mathfrak{l}_i^2}$.

We now state our main result, which is immediate from Lemmas 5 and 6.

THEOREM 1. *Let* D *and* q *have the form* (1.1)—(1.3), *and let* $\mathscr{C}_q$ *be the restricted 2-classgroup of* $Q(\sqrt{Dq})$. *Then the structure of* $\mathscr{C}_q/\mathscr{C}_q^8$ *is completely determined by* *the Frobenius symbol* $\left(\frac{\bar{\Sigma}_D/Q}{q}\right)$, *where* $\bar{\Sigma}_D$ *is a suitable normal extension of* Q. *In fact, we may take* $\bar{\Sigma}_D=\Sigma_D\bar{\Lambda}_D$, *where* Σ_D *is the compositum of all the fields* K_i *and* L_{ij} *(for* $1\leqq i, j\leqq r$*) and* $\bar{\Lambda}_D$ *is the compositum of all* $\bar{\Lambda}_d$ *for* $d|D$.

If we restrict ourselves to primes q for which $d=p_1 \dots p_s$ is fixed, then clearly we may replace $\bar{\Sigma}_D$ in this theorem by the subfield $\bar{\Sigma}_d=\Sigma_d\bar{\Lambda}_d$. This greatly simplifies the computation of densities. Our computations will then require us to know something about the Galois group of $\bar{\Sigma}_d/Q$.

Before computing the degree $[\bar{\Sigma}_d\colon Q]$, we recall the following fact from [7] concerning Σ_d. Σ_d is the independent compositum over $\Omega_d=Q(\sqrt{p_1}, \dots, \sqrt{p_s})$ of the fields $K_i\ (1\leqq i\leqq s)$ and $L_{ij}\ (1\leqq i<j\leqq s)$ and so has degree

(3.16) $$[\Sigma_d\colon Q]=2^{\binom{s}{2}+2s}.$$

For notational convenience we let K_d be the compositum of the K_i, Λ_d be the compositum of the L_{ij}, and $\bar{\Omega}_d=\Omega_d(\sqrt{-1})$.

We now prove

LEMMA 8. *For the field* $\bar{\Sigma}_d=\Sigma_d\bar{\Lambda}_d$ *we have*

(3.17) $$[\bar{\Sigma}_d\colon \bar{\Omega}_d]=[\Sigma_d\colon \Omega_d][\bar{\Lambda}_d\colon \bar{\Omega}_d]=2^{\binom{s}{2}+2s}.$$

REMARK. Since $\bar{\Lambda}_d$ contains $\sqrt{-d}$ and $\sqrt{p_i}$, for each $p_i|d$, it is clear that $\bar{\Omega}_d\subseteqq\bar{\Lambda}_d$.

PROOF. It will suffice to prove the first equality in (3.17), since the second is a consequence of (3.12) and (3.16). First observe that $\Lambda_d\bar{\Lambda}_d$ is unramified over $\bar{\Omega}_d$, since L_{ij} is unramified over $Q(\sqrt{p_ip_j})$ and $\bar{L}_j$ is unramified over $Q(\sqrt{-d})$. Hence by Leopoldt's genus theory [5], and the fact that $\bar{\Omega}_d$ is its own genus field, we have

$$K_d(\sqrt{-1})\cap\Lambda_d\bar{\Lambda}_d=\bar{\Omega}_d$$

(this intersection is abelian over Q and unramified over $\bar{\Omega}_d$). Therefore

$$[\bar{\Sigma}_d\colon \bar{\Omega}_d]=[K_d(\sqrt{-1})\colon \bar{\Omega}_d][\Lambda_d\bar{\Lambda}_d\colon \bar{\Omega}_d]=$$
$$=[K_d\colon \Omega_d][\Lambda_d\bar{\Lambda}_d\colon \bar{\Omega}_d],$$

and to prove (3.17) it suffices to show that

$$[\Lambda_d\bar{\Lambda}_d\colon \bar{\Omega}_d]=[\Lambda_d\colon \Omega_d][\bar{\Lambda}_d\colon \bar{\Omega}_d].$$

We prove the equivalent statement

(3.18) $$[\Lambda_d\bar{\Lambda}_d\colon \Omega_d]=[\Lambda_d\colon \Omega_d][\bar{\Lambda}_d\colon \Omega_d].$$

Now $\Lambda_d\bar{\Lambda}_d$ is a Kummer extension of Ω_d which is obtained by adjoining the radicals $\sqrt{\beta_{ij}}$ (for $1\leqq i<j\leqq s$), $\sqrt{\gamma_i}$ (for $1\leqq i\leqq s$), and $\sqrt{-1}$. (See (3.3) and

(3.15).) To prove (3.18) we only need to show that a relation of the form

$$(3.19)\qquad \Big(\prod_{1\le i<j\le s} \beta_{ij}^{a_{ij}} \prod_{1\le i\le s} \gamma_i^{b_i}\Big)(-1)^c = \eta^2, \quad \eta \text{ in } \Omega_d,$$

implies that $a_{ij}\equiv b_i\equiv c\equiv 0 \pmod 2$ for all i and j. We fix a k between 1 and s and apply to (3.19) the automorphism of Ω_d which takes $\sqrt{p_k}$ to $-\sqrt{p_k}$ and fixes the other $\sqrt{p_i}$. Multiplying the resulting equation by (3.19) and using (3.2) and (3.14) gives that

$$\Big(\prod_{k<j\le s} p_j^{3h_k a_{kj}}\Big)(-1)^{b_k}\left(\frac{d}{p_k}\right)^{h_k b_k} = \eta_k^2,$$

where $\eta_k\in\Omega_d$. But each p_j is a square in Ω_d, while $\sqrt{-1}\notin\Omega_d$. Hence the last equation shows $2|b_k$ for $1\le k\le s$, and so (3.19) gives

$$\Big(\prod_{1\le i<j\le s} \beta_{ij}^{a_{ij}}\Big)(-1)^c = \zeta^2, \quad \zeta \text{ in } \Omega_d.$$

But $2|c$ since $\Lambda_d\cap\bar{\Omega}_d=\Omega_d$, and we conclude from Lemma 13 of [7] that $2|a_{ij}$ for all i and j. This completes the proof.

Using this lemma and what we already know about Σ_d, we conclude that $\bar{\Sigma}_d$ is the independent compositum over $\bar{\Omega}_d$ of the fields $K_i\bar{\Omega}_d$, $L_{ij}\bar{\Omega}_d$ and $\bar{L}_i\bar{\Omega}_d$. Thus the Galois group

$$(3.20)\qquad \bar{\Theta}_d = \mathrm{Gal}\,(\bar{\Sigma}_d/\bar{\Omega}_d)$$

has the decomposition

$$(3.21)\qquad \bar{\Theta}_d = \prod_{1\le i\le s}\langle\sigma_i\rangle\times\prod_{1\le i<j\le s}\langle\lambda_{ij}\rangle\times\prod_{1\le i\le s}\langle\bar{\lambda}_i\rangle,$$

where the automorphisms $\sigma_i, \lambda_{ij}, \bar{\lambda}_i$ are extended to $\bar{\Sigma}_d$ so as to fix all but one of the component fields $K_i\bar{\Omega}_d, L_{ij}\bar{\Omega}_d, \bar{L}_i\bar{\Omega}_d$. Since the fields $K_i, L_{ij}, \bar{L}_i$ are all normal over Q, and since $\bar{\Theta}_d$ has exponent 2, it is easy to see that $\bar{\Theta}_d$ is contained in the center of $\mathrm{Gal}\,(\bar{\Sigma}_d/Q)$. (See the argument in [7], Lemma 15.) This is important for our computations in the next section.

4. Density results

Before discussing the fundamental unit in $Q(\sqrt{Dq})$, it is convenient to make the dependence of M' on $\bar{\Sigma}_d$ explicit. Observe that the conjugacy class $\left(\frac{\bar{\Sigma}_d/Q}{q}\right)$ fixes $\bar{\Omega}_d$, and therefore lies in $\bar{\Theta}_d$, by virtue of (2.2). Our remarks in § 3 imply that this conjugacy class contains a single element. Hence we may set

$$(4.1)\qquad \left(\frac{\bar{\Sigma}_d/Q}{q}\right) = \prod_{i=1}^{s}\sigma_i^{a_i}\prod_{1\le i<j\le s}\lambda_{ij}^{b_{ij}\bar{\delta}_{ij}}\prod_{1\le i\le s}\bar{\lambda}_i^{c_i\bar{\delta}_i},$$

where $\bar{\delta}_{ij}$ and $\bar{\delta}_i$ are defined by (3.7) and (3.9), and the integers a_i, b_{ij}, c_i are unique mod 2. The character values $\chi_j(\mathfrak{z}_i)$ are then given by

LEMMA 9. *If* (4.1) *holds, then*

$$(4.2)\qquad \chi_j(\mathfrak{z}_i)=\begin{cases}\left(\dfrac{D/p_i}{p_i}\right)_4(-1)^{a_i}, & 1\leqq i\leqq s,\quad j=i,\\ \left(\dfrac{p_i}{p_j}\right)_4(-1)^{b_{ij}}, & 1\leqq i,j\leqq s,\quad i\neq j,\\ \left(\dfrac{D/d}{p_j}\right)_4(-1)^{a_j+c_j}, & 1\leqq j\leqq s,\quad i=s+1,\end{cases}$$

where b_{ij} *is defined to equal* b_{ji} *if* $i>j$.

This lemma is immediate from Lemmas 5 and 6, by well-known properties of the Frobenius symbol. It gives a 1—1 correspondence between the automorphisms σ in $\bar{\Theta}_d$ and the $(s+1)\times s$ matrices $\bar{R}=(\varepsilon_{ij})$ over $\mathbf{F}_2$ satisfying

$$(4.3)\qquad (-1)^{\varepsilon_{ij}+\varepsilon_{ji}}=\left(\frac{p_i}{p_j}\right)_4\left(\frac{p_j}{p_i}\right)_4,\qquad 1\leqq i,j\leqq s,\quad i\neq j.$$

Our density results will be a consequence of this correspondence and the following technical lemma, which is easily proved using (3.17) and the Frobenius density theorem. (For exact details see the analogous Lemma 17 in [7].)

LEMMA 10. *Let* $d=p_1\dots p_s$ *be a fixed divisor of* D, *and set* $D/d=\tilde{p}_1\dots\tilde{p}_l$ $(l=r-s)$. *If* $\sigma\in\bar{\Theta}_d$, *then the set of primes* q *which satisfy the conditions*

$$\left(\frac{\tilde{p}_i}{q}\right)=-1\quad for\quad 1\leqq i\leqq l,$$

$$\left(\frac{\bar{\Sigma}_d/Q}{q}\right)=\sigma,$$

is non-empty, and has Dirichlet density equal to $2^{-\binom{s}{2}-2s-r-1}$.

The statement of our results requires the following definition.

DEFINITION. Let d be an integer with s prime factors satisfying (1.2). We define $\bar{N}(d,\varrho)$ to be the number of $(s+1)\times s$ matrices $\bar{R}=(\varepsilon_{ij})$ over $\mathbf{F}_2$ which have rank ϱ and satisfy (4.3). Also, $\bar{N}(1,0)=1$.

It is not hard to see that $\bar{N}(d,\varrho)$ does not depend on the ordering of the prime factors of d. Since there is no condition on the last row of $\bar{R}$ in this definition, it is possible to express $\bar{N}(d,\varrho)$ in terms of the number $N(d,\varrho)$ of $s\times s$ matrices over $\mathbf{F}_2$ which have rank ϱ and satisfy (4.3) (see [7], § 6). Clearly,

$$(4.4)\qquad \bar{N}(d,\varrho)=2^{\varrho}N(d,\varrho)+(2^s-2^{\varrho-1})N(d,\varrho-1).$$

This holds for any ϱ between 0 and s, as long as we define $N(d,-1)=0$.

Now fix a divisor d of D and consider the primes q, for which $\left(\frac{q}{\tilde{p}}\right)=-1$ for $\tilde{p}|D/d$ and

$$(4.5)\qquad \mathscr{C}_q/\mathscr{C}_q^8\cong C_2^{(r-s)}\times C_4^{(\varrho)}\times C_8^{(s-\varrho)};$$

here $0 \leqq \varrho \leqq s \leqq r$ and $C_n^{(m)}$ denotes a product of m cyclic groups of order n. Then by the above correspondence, Lemma 10, and (2.7), we see that this set of primes has density $2^{-\binom{s}{2}-2s-r-1}\bar{N}(d, \varrho)$. Summing over d shows that the following theorem holds.

THEOREM 2. *Let* D *be given by* (1.1), (1.2), *and let* s, ϱ *satisfy* $0 \leqq \varrho \leqq s \leqq r$. *Then the primes* q, *for which the restricted* 2-*classgroup* $\mathscr{C}_q$ *of* $Q(\sqrt{Dq})$ *satisfies* (4.5), *have a density* $\bar{\partial}(D, s, \varrho)$ *given by the formula*

$$\bar{\partial}(D, s, \varrho) = 2^{-\binom{s}{2}-2s-r-1} \sum_{\substack{d \mid D \\ v(d)=s}} \bar{N}(d, \varrho); \tag{4.6}$$

here $v(d)$ *is the number of prime factors of* d.

As in [7], this expression simplifies in case the prime factors of D satisfy

$$\left(\frac{p_i}{p_j}\right)_4 = \left(\frac{p_j}{p_i}\right)_4 \quad \text{for} \quad i \neq j. \tag{4.7}$$

In this case $\bar{N}(d, \varrho) = \bar{N}_s(\varrho)$ is independent of d, and we have

THEOREM 3. *If the prime factors of* D *satisfy the additional condition* (4.7), *then*

$$\bar{\partial}(D, s, \varrho) = \binom{r}{s} 2^{-\binom{s}{2}-2s-r-1} \bar{N}_s(\varrho),$$

where $\bar{N}_s(\varrho)$ *is given as follows:*

$$\bar{N}_s(\varrho) = \bar{N}_\varrho(\varrho) \frac{(2^s-1)(2^{s-1}-1)\dots(2^{s-\varrho+1}-1)}{(2^\varrho-1)(2^{\varrho-1}-1)\dots(2-1)}, \tag{4.8}$$

with

$$\bar{N}_\varrho(\varrho) = u(\varrho) \prod_{k=0}^{\left[\frac{\varrho-1}{2}\right]} (2^{2k+1}-1), \tag{4.9}$$

and

$$u(\varrho) = \begin{cases} 3 \cdot 2^{(\varrho^2+4\varrho-5)/4}, & \varrho \text{ odd}, \\ (3 \cdot 2^\varrho - 1) 2^{(\varrho^2+2\varrho-4)/4}, & \varrho \text{ even}. \end{cases} \tag{4.10}$$

This calculation follows from (4.4) and the fact that for integers D satisfying (4.7), $N(d, \varrho) = N_s(\varrho)$ is just the number of *symmetric* $s \times s$ matrices over $\mathbf{F}_2$ with rank ϱ. The computation of $N_s(\varrho)$ is given in [7] (see Lemma 18); we quote the result here since it will occur again in connection with the Pell equation. We have

$$N_s(\varrho) = N_\varrho(\varrho) \frac{(2^s-1)(2^{s-1}-1)\dots(2^{s-\varrho+1}-1)}{(2^\varrho-1)(2^{\varrho-1}-1)\dots(2-1)}, \tag{4.11}$$

with

$$N_\varrho(\varrho) = 2^{\varepsilon(\varrho)} \prod_{k=0}^{\left[\frac{\varrho-1}{2}\right]} (2^{2k+1}-1), \tag{4.12}$$

and

$$(4.13) \qquad \varepsilon(\varrho) = \begin{cases} (\varrho^2-1)/4, & \varrho \text{ odd}, \\ (\varrho^2+2\varrho)/4, & \varrho \text{ even}. \end{cases}$$

We also note that the quantity $\bar{\partial}(D, s, \varrho)$ in Theorem 3 is *positive*. This implies the following result, first proved by Rédei [10].

THEOREM 4. *Let $\mathcal{T}$ be any finite abelian group whose exponent divides* 8. *Then there are infinitely many real quadratic fields with a restricted classgroup $\mathcal{C}$ satisfying*

$$\mathcal{C}/\mathcal{C}^8 \cong \mathcal{T}.$$

If the exponent of $\mathcal{T}$ divides 4, *infinitely many real quadratic fields have restricted 2-classgroups isomorphic to $\mathcal{T}$.*

The only point to be checked here is that for any $r \geqq 1$, there are integers D satisfying (1.1), (1.2) and (4.7). But this is easy to prove using induction and familiar principles from algebraic number theory.

5. The Pell equation

We are now ready to apply the preceding analysis to the investigation of the fundamental unit η_q of $Q(\sqrt{Dq})$. We start with the following simple observation: the norm of η_q is -1 if and only if the ideal $\mathfrak{u} = (\sqrt{Dq}) = \mathfrak{p}_1 \dots \mathfrak{p}_r\mathfrak{q}$ has a generator with positive norm. Hence we have the criterion

$$(5.1) \qquad \text{Norm } \eta_q = -1 \quad \text{iff} \quad \mathfrak{u} = \mathfrak{p}_1 \dots \mathfrak{p}_r \mathfrak{q} \sim 1.$$

We now consider a somewhat weaker condition on the ideal $\mathfrak{u}$.

LEMMA 11. *Assume that* (2.2) *holds and that the Frobenius symbol* $\left(\frac{\Sigma_d/Q}{q}\right)$ *is given by* (4.1). *Then the ideal* $\mathfrak{u}$ *satisfies*

$$(5.2) \qquad \mathfrak{u} \sim \mathfrak{c}^4 \quad \textit{for some } \mathfrak{c}$$

if and only if we have

$$(5.3) \qquad c_j \equiv \sum_{\substack{1 \leqq k \leqq s \\ k \neq j}} b_{kj} \pmod 2, \quad \textit{for} \quad 1 \leqq j \leqq s.$$

PROOF. From (2.6) we see first of all that (5.2) holds iff

$$(5.4) \qquad \mathfrak{z}_1 \cdots \mathfrak{z}_s \mathfrak{z}_{s+1} \sim \mathfrak{c}^2 \mathfrak{a}, \quad \text{where } \mathfrak{a}^2 \sim 1.$$

Let A be the group of ambiguous classes in $\mathcal{C}_q$. From the form of the matrix M in § 2 it is clear that the characters $\chi_1, \dots, \chi_s$ generate the quadratic characters of X_2 which are trivial on A, and are therefore a basis for the quadratic characters of the quotient group $\mathcal{C}_q/A$. Hence (5.4) holds iff

$$(5.5) \qquad \chi_j(\mathfrak{z}_1 \cdots \mathfrak{z}_{s+1}) = +1 \quad \text{for} \quad 1 \leqq j \leqq s.$$

But from Lemma 9 and (4.1),

$$\chi_j(\mathfrak{z}_1 \cdots \mathfrak{z}_{s+1}) = \left(\frac{D/p_j}{p_j}\right)_4 (-1)^{a_j} \left(\prod_{\substack{1 \le i \le s \\ i \ne j}} \left(\frac{p_i}{p_j}\right)_4 (-1)^{b_{ij}} \right) \left(\frac{D/d}{p_j}\right)_4 (-1)^{a_j+c_j} =$$

$$= (-1)^{c_j + \sum_{i \ne j} b_{ij}},$$

and this proves the assertion.

Now suppose that the invariant e_8 of $Q(\sqrt{Dq})$ is zero, so that $\mathscr{C}_q$ has exponent 2 or 4. Then (5.2) is equivalent to $u \sim 1$, and (5.3) (or (5.5)) becomes necessary and sufficient for Norm $\eta_q = -1$. Thus in the presence of the condition $e_8 = 0$, Norm η_q depends only on the behavior of q in Σ_d. This observation allows us to prove

THEOREM 5. *Let D satisfy* (1.1) *and* (1.2), *and let s be any integer with* $0 \le s \le r$. *Then the primes q, for which*

(5.6) $$\mathscr{C}_q \cong C_2^{(r-s)} \times C_4^{(s)} \quad \textit{and} \quad \text{Norm } \eta_q = -1,$$

have a positive density, given by the formula

(5.7) $$\partial_-(D, s) = 2^{-\binom{s}{2} - 2s - r - 1} \sum_{\substack{d \mid D \\ \nu(d) = s}} N(d, s);$$

as in § 4, *$N(d, s)$ is the number of nonsingular $s \times s$ matrices over $\mathbf{F}_2$ satisfying* (4.3) *for all the prime divisors p_i, p_j of d.*

If $s > 0$, then the primes q, for which

(5.8) $$\mathscr{C}_q \cong C_2^{(r-s)} \times C_4^{(s)} \quad \textit{and} \quad \text{Norm } \eta_q = +1,$$

also have positive density, the value of which is

(5.9) $$\partial_+(D, s) = 2^{-\binom{s}{2} - 2s - r - 1} \sum_{\substack{d \mid D \\ \nu(d) = s}} \left(\bar{N}(d, s) - N(d, s)\right).$$

PROOF. Fix a divisor d of D, and let B_d denote the set of automorphisms of $\bar{\Theta}_d$ which satisfy the condition (5.3) and

(5.10) $$\text{rank } M' = \text{rank}\left(\chi_j(\mathfrak{z}_i)\right) = s,$$

where the $\chi_j(\mathfrak{z}_i)$ are obtained as before from Lemma 9. If $\left(\frac{q}{\bar{p}}\right) = -1$ for $\bar{p} | D/d$, then we see from (2.7) that (5.10) is equivalent to $e_8 = 0$. Hence the above discussion shows that (5.6) holds iff $\left(\frac{\Sigma_d/Q}{q}\right) \in B_d$.

But from (5.5) (which is equivalent to (5.3)), it is clear that rank $M' = s$ iff the first s rows of M' form a nonsingular $s \times s$ matrix $\tilde{M}$ over $\mathbf{F}_2$. Note that $\tilde{M}$ satisfies (4.3), and that such matrices $\tilde{M}$ are in 1—1 correspondence with the elements of the set B_d. Hence $|B_d| = N(d, s)$ and (5.7) is a consequence of Lemma 10. Also, (5.9) follows by subtracting (5.7) from (4.6).

Finally, we note $\partial_-(D,s)>0$, because for any divisor d of D there is certainly a nonsingular upper-triangular $s\times s$ matrix $\bar{M}=(\varepsilon_{ij})$ satisfying (4.3). Furthermore, if $s>0$, then setting $\varrho=s$ in (4.4) gives that

$$\bar{N}(d,s)\geqq 2^s N(d,s) > N(d,s),$$

and so $\partial_+(D,s)>0$. This proves the theorem.

As a corollary we obtain

THEOREM 6. *Let* $\mathcal{T}$ *be any finite abelian group of exponent* 4, *and let* $\varepsilon\in\{\pm 1\}$. *Then infinitely many real quadratic fields have a restricted* 2*-classgroup isomorphic to* $\mathcal{T}$ *and a fundamental unit with norm* ε. *In particular, infinitely many real quadratic fields have an absolute* 2*-classgroup isomorphic to* $\mathcal{T}$.

The last assertion of Theorem 6 is clear from the fact that the absolute and restricted classgroups coincide whenever $\text{Norm}\,\eta_q=-1$.

We now consider the subclass of integers D satisfying (4.7), and we prove

THEOREM 7. *For a fixed* D *satisfying* (1.1), (1.2) *and* (4.7), *let* S_D *denote the set of primes* q *for which* $\text{Norm}\,\eta_q=-1$, *and let* $\underline{\delta}(D)$ *and* $\bar{\delta}(D)$ *denote the lower and upper densities of* S_D. *Then*

$$0<a(r)\leqq\underline{\delta}(D)\leqq\bar{\delta}(D)\leqq\frac{1}{2}\left(\frac{3}{4}\right)^r, \tag{5.11}$$

where

$$a(r)=\sum_{s=0}^{r}\frac{\binom{r}{s}N_s(s)}{2^{\binom{s}{2}+2s+r+1}}, \tag{5.12}$$

and $N_s(s)$ *is the number of* $s\times s$ *matrices over* $\mathbf{F}_2$ *which are symmetric and nonsingular. The function* $a(r)$ *satisfies the asymptotic formula*

$$a(r)\sim\frac{\theta}{2}\left(\frac{3}{4}\right)^r,\quad \textit{as}\quad r\to\infty, \tag{5.13}$$

where

$$\theta=\prod_{k=0}^{\infty}\left(1-\frac{1}{2^{2k+1}}\right)=.4194\ldots. \tag{5.14}$$

PROOF. The lower bound in (5.11) is a consequence of (5.7). For summing on s and noting $N(d,s)=N_s(s)$ (see (4.11)—(4.13)) gives that $\text{Norm}\,\eta_q=-1$ for a set of primes having the density (5.12). For the upper bound we use Lemma 11 and the fact that (5.2) is a necessary condition for $\text{Norm}\,\eta_q=-1$ (regardless of the value of e_8). Exactly $2^{\binom{s}{2}+s}$ of the automorphisms in $\bar{\Theta}_d$ satisfy (5.3), for a fixed d, and so the set of q for which (5.2) holds has the density

$$\sum_{s=0}^{r}\binom{r}{s}\frac{2^{\binom{s}{2}+s}}{2^{\binom{s}{2}+2s+r+1}}=\frac{1}{2^{r+1}}\sum_{s=0}^{r}\binom{r}{s}2^{-s}=\frac{1}{2}\left(\frac{3}{4}\right)^r.$$

It only remains to prove (5.13). We do this as follows. It is not difficult to see from equations (4.12) and (4.13) that

$$N_s(s) \sim \theta 2^{(s^2+s)/2}, \quad \text{as} \quad s \to \infty,$$

where θ is defined by (5.14). Setting

$$w_s = \frac{N_s(s)}{2^{(s^2+s)/2}}$$

we have

(5.15) $$\lim_{s\to\infty} w_s = \theta,$$

and

(5.16) $$a(r) = \sum_{s=0}^{r} \frac{\binom{r}{s}}{2^{s+r+1}} w_s.$$

If we put

$$b(r) = \sum_{s=0}^{r} \frac{\binom{r}{s}}{2^{s+r+1}} = \frac{1}{2}\left(\frac{3}{4}\right)^r,$$

and

(5.17) $$c_{rs} = \begin{cases} \binom{r}{s} 2^{-s-r-1} b^{-1}(r), & \text{if} \quad r \geqq s, \\ 0, & \text{if} \quad r < s, \end{cases}$$

then (5.16) and (5.17) imply

(5.18) $$\frac{a(r)}{b(r)} = \sum_{s=0}^{\infty} c_{rs} w_s.$$

Now the infinite matrix (c_{rs}) $(r, s = 0, 1, \ldots)$ satisfies the Toeplitz condition for regularity (see [2], p. 43), since

$$\sum_{s=0}^{\infty} |c_{rs}| = \sum_{s=0}^{\infty} c_{rs} = 1,$$

and

$$|c_{rs}| \leqq \frac{r^s}{s!} 2^{-s-r-1} 2\left(\frac{4}{3}\right)^r = \frac{2^{-s}}{s!} r^s \left(\frac{2}{3}\right)^r \to 0,$$

as $r \to \infty$, for fixed s. Hence by (5.18) and (5.15) we have

$$\lim_{r\to\infty} \frac{a(r)}{b(r)} = \theta,$$

which proves (5.13).

REMARK. The upper bound in (5.11) holds whether or not D satisfies (4.7). Thus the equation

$$x^2 - Dqy^2 = -1$$

is more likely to be unsolvable than solvable, for large $r = \nu(D)$. This is in sharp contrast to the results proved by Rédei [9]. (See his Satz IX and Satz X.)

We note that the proof of the upper bound in (5.11) may be adapted to show that for a fixed factorization $D=D'D''$, and a fixed choice of $\varepsilon=\pm 1$, the equation

$$D'x^2-D''qy^2=\varepsilon z^4,\quad (x,y)=1,\quad 2\nmid z,$$

is solvable for a set of primes q having the density $3^{\nu(D'')}/2^{2r+1}$. (Also see [7], Theorem 6.) In addition, our analysis of η_q may be extended to the Diophantine equations

$$D'x^2-D''qy^2=\pm 1, \tag{5.19}$$

by replacing the ideal $\mathfrak{u}$ in (5.1) and in Lemma 11 by an ideal $\tilde{\mathfrak{u}}$ for which $\operatorname{Norm}\tilde{\mathfrak{u}}=D'$ or $D''q$, according as the upper or lower sign is taken in (5.19). For the sake of brevity we leave further discussion of these results to the interested reader.

6. Numerical results

The explicit values of the densities $\bar{\partial}(D,s,\varrho)$, $\partial_+(D,s)$ and $\partial_-(D,s)$ in the cases $r=1,2,3$ have been worked out and listed in Tables 1—3 opposite the isomorphism types to which they correspond. When $r=2$ and $D=p_1p_2$, the densities depend on the value of $\left(\frac{p_1}{p_2}\right)_4\left(\frac{p_2}{p_1}\right)_4$. When $r=3$ and $D=p_1p_2p_3$, there are 4 possible cases, corresponding to the number of the values

$$\mu_{ij}=\left(\frac{p_i}{p_j}\right)_4\left(\frac{p_j}{p_i}\right)_4,\qquad 1\leqq i<j\leqq 3,$$

which are $+1$. The columns in Table 3 correspond to these 4 cases, and are indexed by the vector

$$\underline{\mu}=(\mu_{12},\mu_{13},\mu_{23}).$$

For completeness we have also included the values of $N(d,\varrho)$ and $\bar{N}(d,\varrho)$ in Tables 4—7.

Note that ∂_+ and ∂_- are only given for groups having exponent 2 or 4.

In connection with the analogous results in [7], § 6, the following remark is of interest. Let D be given by (1.1) and (1.2) and let $\partial(D,s,\varrho)$ denote the density of primes q, for which the classgroup $\mathscr{C}$ of the *imaginary* quadratic field $Q(\sqrt{-Dq})$ satisfies

$$\mathscr{C}/\mathscr{C}^8\cong C_2^{(r-s)}\times C_4^{(\varrho)}\times C_8^{(s-\varrho)}.$$

(All such q are $\equiv 3 \pmod 4$.) Then ∂_- and $\bar{\partial}$ are related to ∂ by the formulas

$$\partial_-(D,s)=2^{-s}\partial(D,s,s),$$

and

$$\bar{\partial}(D,s,\varrho)=2^{\varrho-s}\partial(D,s,\varrho)+(1-2^{\varrho-1-s})\partial(D,s,\varrho-1),\quad \varrho\geqq 1.$$

The first relation follows from (5.7) and formula (6.7) of [7], while the second is a consequence of (4.6) and (4.4). Setting $\varrho=s$ and subtracting gives the further relation

$$\partial_+(D,s)=(1-2^{-s})\partial(D,s,s)+\frac{1}{2}\partial(D,s,s-1),\quad s\geqq 1.$$

Thus the densities in Tables 1—3 may be computed directly from the results in Tables 1—3 of [7].

Table 1 $(r=1)$

$\mathscr{C}/\mathscr{C}^8$	$\bar{\partial}$	∂_+	∂_-
C_2	1/4	0	1/4
C_4	3/16	1/8	1/16
C_8	1/16		

Table 2 $(r=2)$

	$\left(\frac{p_1}{p_2}\right)_4=\left(\frac{p_2}{p_1}\right)_4$			$\left(\frac{p_1}{p_2}\right)_4\neq\left(\frac{p_2}{p_1}\right)_4$		
$\mathscr{C}/\mathscr{C}^8$	$\bar{\partial}$	∂_+	∂_-	$\bar{\partial}$	∂_+	∂_-
$C_2\times C_2$	1/8	0	1/8	1/8	0	1/8
$C_2\times C_4$	3/16	1/8	1/16	3/16	1/8	1/16
$C_2\times C_8$	1/16			1/16		
$C_4\times C_4$	11/128	9/128	1/64	5/64	9/128	1/128
$C_4\times C_8$	9/256			3/64		
$C_8\times C_8$	1/256			0		

Table 3 $(r=3)$

$\mathscr{C}/\mathscr{C}^8$	$\bar{\partial}$	∂_+	∂_-	$\bar{\partial}$	∂_+	∂_-
		(1, 1, 1)			(−1, 1, 1)	
$C_2\times C_2\times C_2$	1/16	0	1/16	1/16	0	1/16
$C_2\times C_2\times C_4$	9/64	3/32	3/64	9/64	3/32	3/64
$C_2\times C_2\times C_8$	3/64			3/64		
$C_2\times C_4\times C_4$	33/256	27/256	3/128	1/8	27/256	5/256
$C_2\times C_4\times C_8$	27/512			15/512		
$C_2\times C_8\times C_8$	3/512			1/256		
$C_4\times C_4\times C_4$	21/512	77/2048	7/2048	79/2048	37/1024	5/2048
$C_4\times C_4\times C_8$	77/4096			93/4096		
$C_4\times C_8\times C_8$	21/8192			5/4096		
$C_8\times C_8\times C_8$	1/8192			0		
		(−1, −1, 1)			(−1, −1, −1)	
$C_2\times C_2\times C_2$	1/16	0	1/16	1/16	0	1/16
$C_2\times C_2\times C_4$	9/64	3/32	3/64	9/64	3/32	3/64
$C_2\times C_2\times C_8$	3/64			3/64		
$C_2\times C_4\times C_4$	31/256	27/256	1/64	15/128	27/256	3/256
$C_2\times C_4\times C_8$	33/512			9/128		
$C_2\times C_8\times C_8$	1/512			0		
$C_4\times C_4\times C_4$	39/1024	73/2048	5/2048	39/1024	73/2048	5/2048
$C_4\times C_4\times C_8$	47/2048			47/2048		
$C_4\times C_8\times C_8$	3/2048			3/2048		
$C_8\times C_8\times C_8$	0			0		

Table 4

ϱ	$N(p_1p_2, \varrho)$: $\left(\frac{p_1}{p_2}\right)_4 = \left(\frac{p_2}{p_1}\right)_4$	$N(p_1p_2, \varrho)$: $\left(\frac{p_1}{p_2}\right)_4 \neq \left(\frac{p_2}{p_1}\right)_4$
0	1	0
1	3	6
2	4	2

Table 5

	$N(p_1p_2p_3, \varrho)$			
ϱ	(1, 1, 1)	(−1, 1, 1)	(−1, −1, 1)	(−1, −1, −1)
0	1	0	0	0
1	7	5	6	6
2	28	39	38	38
3	28	20	20	20

Table 6

ϱ	$\bar{N}(p_1p_2, \varrho)$: $\left(\frac{p_1}{p_2}\right)_4 = \left(\frac{p_2}{p_1}\right)$	$\bar{N}(p_1p_2, \varrho)$: $\left(\frac{p_1}{p_2}\right)_4 \neq \left(\frac{p_2}{p_1}\right)_4$
0	1	0
1	9	12
2	22	20

Table 7

	$\bar{N}(p_1p_2p_3, \varrho)$			
ϱ	(1, 1, 1)	(−1, 1, 1)	(−1, −1, 1)	(−1, −1, −1)
0	1	0	0	0
1	21	10	12	12
2	154	186	188	188
3	336	316	312	312

7. An example

We shall illustrate the preceding discussion with the following example. Let $r=3$ and take

$$D=p_1p_1p_3=41\cdot 241\cdot 617=6\,096\,577.$$

Note that

$$(7.1)\qquad \begin{gathered}\left(\frac{41}{241}\right)_4 = \left(\frac{241}{41}\right)_4 = -1, \quad \left(\frac{41}{617}\right)_4 = \left(\frac{617}{41}\right)_4 = -1,\\ \left(\frac{241}{617}\right)_4 = \left(\frac{617}{241}\right)_4 = +1,\end{gathered}$$

so D satisfies (4.7). In Table 8 we have given generators for the field $\bar{\Sigma}_D$. These numbers satisfy the norm conditions

$$\text{Norm}\,\tilde{\beta}_{ij} = p_j, \quad \text{Norm}\,\tilde{\gamma}_i = -4\frac{D}{p_i},$$

and are easily seen to generate the same fields as do the numbers β_{ij} and γ_i given by (3.2) and Lemma 7. (For this note $h_i = 1$ in each field $Q(\sqrt{p_i})$. See [1].)

Table 8

Generators for $\Sigma_{6\,096\,577}$			
(i, j)	$\tilde{\beta}_{ij}$	i	$\tilde{\gamma}_i$
(1, 2)	$129+20\sqrt{41}$	1	$\frac{1617+349\sqrt{41}}{2}$
(1, 3)	$181+28\sqrt{41}$	2	$\frac{417+49\sqrt{241}}{2}$
(2, 3)	$34029+2192\sqrt{241}$	3	$\frac{477+25\sqrt{617}}{2}$

We next note that $\delta_i = 1$ for each i, in the case $d = D$, by (2.14) and (3.9). We claim that $\tilde{\delta}_{ij} = 1$ as well. For D/p_ip_j splits in L_{ij}, for each pair (i, j), and the fundamental units in fields $Q(\sqrt{41\cdot 241})$ and $Q(\sqrt{41\cdot 617})$ have norm -1 (see (7.1) and [6], Theorem 5). Thus we see from (2.11) and (3.7) that $\tilde{\delta}_{12} = \tilde{\delta}_{13} = 1$. However, the fundamental unit in $Q(\sqrt{241\cdot 617})$ has norm $+1$. (This is because $617\cdot 5^2 - 241\cdot 8^2 = 1$, which implies that the unique ambiguous principal ideal in $Q(\sqrt{241\cdot 617})$ is the prime divisor of 617, and not $(\sqrt{241\cdot 617})$.) To compute

$$\tilde{\delta}_{23} = \left(\frac{L_{23}/Q(\sqrt{241\cdot 617})}{(\sqrt{241\cdot 617})}\right)$$

we may set

$$\tilde{\delta}_{23} = \left(\frac{L_{23}/Q(\sqrt{241\cdot 617})}{\mathfrak{p}_{347}}\right),$$

where the prime ideal $\mathfrak{p}_{347} = (771 + 2\sqrt{241\cdot 617})$ is generated by an element with negative norm. Since $\beta_{23} \equiv 157 \pmod{\mathfrak{p}_{347}}$ and $\left(\frac{157}{347}\right) = +1$, $\mathfrak{p}_{347}$ splits in L_{23}, giving $\tilde{\delta}_{23} = 1$.

Now if $\left(\frac{q}{p_i}\right) = +1$ for $1 \leqq i \leqq 3$ and

$$\left(\frac{\Sigma_D/Q}{q}\right) = \sigma_1^{a_1}\sigma_2^{a_2}\sigma_3^{a_3}\lambda_{12}^{b_{12}}\lambda_{13}^{b_{13}}\lambda_{23}^{b_{23}}\bar{\lambda}_1^{c_1}\bar{\lambda}_2^{c_2}\bar{\lambda}_3^{c_3},$$

then the matrix M' is given by

$$M' = \begin{pmatrix} a_1 & b_{12}+1 & b_{13}+1 \\ b_{12}+1 & a_2+1 & b_{23} \\ b_{13}+1 & b_{23} & a_3+1 \\ a_1+c_1 & a_2+c_2 & a_3+c_3 \end{pmatrix}.$$

In particular, M' is the zero matrix if and only if

$$\left(\frac{\Sigma_D/Q}{q}\right) = \sigma_2\sigma_3\lambda_{12}\lambda_{13}\bar{\lambda}_2\bar{\lambda}_3. \tag{7.2}$$

Equation (7.2) gives the necessary and sufficient condition for the isomorphism

$$\mathscr{C}_q/\mathscr{C}_q^8 \cong C_8 \times C_8 \times C_8$$

and is satisfied by a set of primes having density 1/8192. Also, if

$$\left(\frac{\Sigma_D/Q}{q}\right) = \sigma_1\lambda_{12}\lambda_{13}\bar{\lambda}_2\bar{\lambda}_3,$$

then

$$M' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

and by Lemma 11 (and the discussion following) the equation

$$x^2 - 6\,096\,577\,qy^2 = -1$$

has a solution.

We now work out M' in the case $q = 1321$. The residues modulo 1321 of the $\tilde{\beta}_{ij}$ and $\tilde{\gamma}_i$ are listed below, where we have taken the square-roots $\sqrt{41}, \sqrt{241}, \sqrt{617}$ to be respectively congruent to 103, 617, 324 (mod 1321).

	$\tilde{\beta}_{12}$	$\tilde{\beta}_{13}$	$\tilde{\beta}_{23}$	$\tilde{\gamma}_1$	$\tilde{\gamma}_2$	$\tilde{\gamma}_3$
Residue	868	423	764	288	794	986
Quadratic character (mod 1321)	−1	−1	−1	+1	−1	−1

Since $\left(\frac{1321}{p_i}\right)_4 = -1$ for each i, we see that

$$\left(\frac{\Sigma_D/Q}{1321}\right) = \sigma_1\sigma_2\sigma_3\lambda_{12}\lambda_{13}\lambda_{23}\bar{\lambda}_2\bar{\lambda}_3,$$

and

$$M' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

The rank of M' is 3, and so

$$\mathscr{C}_q \cong C_4 \times C_4 \times C_4.$$

From Lemma 11 it follows that $x^2 - D\cdot 1321 y^2 = -1$ is not solvable. However, the sum of the first and last rows of M' is zero, and using this fact it is not difficult to see that the equation

$$41\cdot 1321x^2 - 241\cdot 617y^2 = 1$$

does have a solution. Thus the unique ambiguous principal ideal in $Q(\sqrt{8\,053\,578\,217})$ is the ideal of norm $41\cdot 1321 = 54\,161$.

REFERENCES

[1] BOREVIČ, Z. I. and ŠAFAREVIČ, I. R., *Number theory*, Pure and Applied Mathematics, Vol. 20, Academic Press, New York—London, 1966, p. 424. *MR* **33** # 4001.

[2] HARDY, G. H., *Divergent series*, Oxford, at the Clarendon Press, 1949. *MR* **11**—25.

[3] HASSE, H., An algorithm for determining the structure of the 2-Sylow-subgroup of the divisor class group of a quadratic number field, *Symposia Mathematica*, Vol. XV (Convegno di Strutture in Corpi Algebrici, INDAM, Rome, 1973), Academic Press, London, 1975, 341—352. *MR* **52** # 8082.

[4] HASSE, H., *Bericht über neuere Untersuchungen und Probleme aus der Theorie der algebraischen Zahlkörper, Teil I: Klassenkörpertheorie, Teil Ia: Beweise zu Teil I, Teil II: Reziprozitätsgesetz*, Dritte Auflage, Physica-Verlag, Würzburg—Wien, 1970. *MR* **42** # 1795.

[5] LEOPOLDT, H., Zur Geschlechtertheorie in abelschen Zahlkörpern, *Math. Nachr.* **9** (1953), 351—362. *MR* **15**—14.

[6] MORTON, P., On Rédei's theory of the Pell equation, *J. Reine Angew. Math.* **307/308** (1979), 373—398. *MR* **81f**: 12005.

[7] MORTON, P., Density results for the 2-classgroups of imaginary quadratic fields, *J. Reine Angew. Math.* **332** (1982), 156—187.

[8] RÉDEI, L. and REICHARDT, H., Die Anzahl der durch 4 teilbaren Invarianten der Klassengruppe eines beliebigen quadratischen Zahlkörpers, *J. Reine Angew. Math.* **170** (1933), 69—74. *Zbl* **7**. 396.

[9] RÉDEI, L. Über einige Mittelwertfragen im quadratischen Zahlkörper, *J. Reine Angew. Math.* **174** (1935), 15—55. *Zbl* **13**. 003.

[10] RÉDEI, L., Ein neues zahlentheoretisches Symbol mit Anwendungen auf die Theorie der quadratischen Zahlkörper I, *J. Reine Angew. Math.* **180** (1939), 1—43. *Zbl* **21**.007.

[11] RÉDEI, L., Die 2-Ringklassengruppe des quadratischen Zahlkörpers und die Theorie der Pellschen Gleichung, *Acta Math. Acad. Sci. Hungar.* **4** (1953), 31—87. *MR* **16**—450.

(Received March 23, 1982)

DEPARTMENT OF MATHEMATICS
HARVARD UNIVERSITY
CAMBRIDGE, MA 02138
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 45—50

FIXED POINT THEOREMS FOR ORDERED SETS

M. KOLIBIAR

Dedicated to the memory of László Rédei

The aim of this paper is to investigate conditions under which each isotone mapping of an ordered set P into itself has a fixed point. In the case P has a least element and every chain in P has a l.u.b. the structure of the set of all fixed points of an isotone mapping from P to P is considered. Most of the results given here were published earlier without proofs [8; 9].

1. Preliminaries. Throughout this paper P denotes an ordered set (i. e. a set with a reflexive, transitive and antisymmetric relation $\leqq_P$ or, more simply, $\leqq$), 0 and I its least and greatest elements (if exist), respectively. The terms least upper bound and greatest lower bound will be abbreviated to l. u. b. and g. l. b., respectively. Given a subset A of P, denote $A^* = \{x \in P\colon a \in A$ implies $a \leqq x\}$, $A^+ = \{x \in P\colon a \in A$ implies $x \leqq a\}$. In particular $\emptyset^* = P$. A subset A of P is said to be upper-bordered in P whenever $A^* \subset A$. In particular a chain is upper-bordered in P iff it is cofinal with a maximal chain in P. A subset S of P is said to satisfy the finite minimal condition if to each $x \in S$ a minimal element m of S exists such that $m \leqq x$, and the set of all minimal elements of S is finite. The interval topology in P is that one whose subbasis of closed sets consists of the sets $(a] = \{x \in P\colon x \leqq a\}$ and $[a) = \{x \in P\colon x \geqq a\}$ (for all $a \in P$). A mapping $f\colon P \to P$ is called isotone (antitone) if $x \leqq y$ implies $f(x) \leqq f(y)$ ($f(x) \geqq f(y)$, respectively). An element $x \in P$ is called a fixed point of f if $f(x) = x$. $P(f)$ will denote the set of all fixed points of f in P. P is said to have the fixed point property (f. p. p.) if every isotone mapping from P to P has a fixed point. A mapping $f\colon P \to P$ is called directable [17] if it is isotone and $H(f) = \{x \in P\colon x \leqq f(x)\}$ is an up-directed subset of P. The set of all positive integers will be denoted by N.

The following lemma will be useful (see e.g. [17, Lemma 2]).

(C) *If C is any chain in P there exists a well-ordered chain $W \subset C$ with $W^* = C^*$.*

2. Statement of the results. A classical theorem by A. Tarski [15] and A. C. Davis [5] states that a lattice has the f. p. p. iff it is complete. An analogous theorem for general ordered sets is not known. The present paper deals with some weaker forms of completeness in ordered sets which are necessary or sufficient for the f. p. p. The most familiar result is the following [1; 2].

(AB) *Let P satisfy the condition*

1980 *Mathematics Subject Classification.* Primary 06A10; Secondary 90D12.
Key words and phrases. Fixed points, isotone mappings.

(a) *l. u. b. of every non-empty well-ordered chain in P exists.*
Let f: $P \to P$ be an isotone mapping such that $a \in P$ exists with $a \leqq f(a)$. Then a fixed point of f exists. In particular, if P has the element 0 *and fulfils* (a) *then P has the f. p. p.*

In a specific class of ordered sets the condition (a) is also necessary; there holds the following theorem (stated in [8, Th. 3] without proof, the proof was given in [14, Theorem 1.2]).

Let P satisfy the condition
($*$) *For each (well-ordered) chain C in P, C^* is down-directed.*
Then P has the f. p. p. iff every (well-ordered) chain (including $\emptyset$) in P has a l. u. b.

In the general case we can prove the following theorems 1 and 2.

THEOREM 1 [8, Theorem 2]. *Let each directable mapping from P to P have a fixed point. Then for every (well-ordered) chain C in P (including $\emptyset$), the set C^* is not empty and to each element $x \in C^*$ a minimal element m of C^* exists with $m \leqq x$. In particular, the conclusion holds if P has the f. p. p.*

COROLLARY. *Let P have the f. p. p. Then to every element $x \in P$ there is a minimal element u and a maximal element v of P with $u \leqq x \leqq v$.*

THEOREM 2 [8, Theorem 5]. *Let for each subset A of P (including $\emptyset$) the set A^* be not empty and satisfy the finite minimal condition. Then P has the f. p. p.*

The condition of Theorem 2 is not necessary as the following example shows. Let P consist of the chain C of all negative integer in its natural order, and of the elements o, a, b where $o < a$, $o < b$, and for each $n \in C$, $a < n$ and $b < n$.
Simple examples show that in Theorem 2 the phrase "each subset A" cannot be replaced by "each chain A". Such a modification is possible in a special case:

THEOREM 3 [8, Corollary of Theorem 4]. *Let for every well-ordered chain C in P (including $\emptyset$) the set C^* be not empty and satisfy the finite minimal condition. Then every entire isotone mapping from P to P has a fixed point.*

(f is said to be entire if $x \leqq f(x)$ whenever $x \leqq f^n(x)$ for some $n \in N$ (depending on x). In particular f is entire if $f^2 = f$.)

REMARK. M. Benado [2] proved, without using the Axiom of choice, a theorem which arises from Theorem 3 by substituting "every set" for "every well-ordered chain".

THEOREM 4 [8, Theorem 1]. *Let P be an ordered set with* 0, *satisfying the condition*
(a′) *If l.u.b. of a well-ordered chain C in P does not exist then C^* is not empty and every inversely well-ordered chain (including $\emptyset$) in C^* has a g.l.b.*
Then P has the f.p.p.

Theorem (AB) is a corollary of Theorem 4.

THEOREM 5[1]. *Suppose to any $x \in P$ there is an element m in* max (P), *the set of*

[1] For the case P is finite this was proved using retracts in [6].

all maximal elements of P, with $x \leqq m$, and let $\max(P)$ *be finite. If for each non-empty subset S of* $\max(P)$, *S^+ has the f. p. p. then P has, too.*

THEOREM 6. *Let P be down-directed and compact in its interval topology. Then P has the f. p. p.*

THEOREM 7. *Let a $\vee$-semilattice P satisfy the condition $(*)$. Then P has the f. p. p. iff it is compact in its interval topology.*

REMARK. This theorem gives a correction of a theorem by L. E. Ward, Jr. [16, Theorem 3]. Counterexamples to the last theorem were given in [9] and [11]. In [11, 12] another theorems correcting Ward's result are given.

THEOREM 8[2]. *Let P contain 0 and satisfy* (a), *and let $f\colon P \to P$ be an isotone mapping. Then the set $P(f)$ (is not empty and) has the following properties:*

(i) *$P(f)$ has a least element.*

(ii) *If l. u. b. of a subset $A \subset P(f)$ in P exists then l. u. b. of A exists in $P(f)$. In particular, l. u. b. of every chain $C \subset P(f)$ exists in $P(f)$.*

(iii) *If $A \subset P(f)$ is upper-bordered in $P(f)$ then l. u. b. of A in $P(f)$ coincides with that in P. In particular, if a chain $C \subset P(f)$ is cofinal with a maximal chain in $P(f)$ then l. u. b. of C in $P(f)$ coincides with that in P.*

COROLLARY 1. *Suppose the hypothesis of Theorem 8 is satisfied. Then*

(i′) *Given $x \in P(f)$, a maximal element m of $P(f)$ exists with $x \leqq m$. (A greatest element in $P(f)$ need not exist even in the case P has a greatest element.)*

(ii′) *Let $\emptyset \neq A \subset P(f)$. The set $A^* = \{x \in P(f)\colon x \leqq a \text{ for all } a \in A\}$ is not empty and for every $x \in A^*$ there is a maximal element m of A^* with $x \leqq m$.*

(Using the terminology of M. Benado [3], $P(f)$ is a complete lower semimultilattice with a least element. But $P(f)$ need not be a multilattice.)

COROLLARY 2. *Let P contain 0 and satisfy* (a), *and let $f\colon P \to P$ be an antitone mapping. Then there is $s \in P$ such that $f^2(s) = s$ and $s \leqq f(s)$.*

REMARK. As a special case we get the theorem in [13] which has applications in game theory (see [13]). In [13] there is supposed that P is a complete lattice and f satisfies $f(\vee A) = \wedge f(A)$ for each $A \subset P$.

3. Proofs of the theorems. The following lemmas will be useful.

LEMMA 1. *Let Y be a well-ordered chain in P and let an element $x \in Y^*$ exist such that there is no minimal element m of Y^* with $m \leqq x$. Then there exists an inversely well-ordered chain Z in P such that*

(1) *$y \in Y$ and $z \in Z$ imply $y \leqq z$.*

(2) *$Y^* \cap Z^+ = \emptyset$.*

PROOF. If $Y^* = \emptyset$ we can set $Z = \emptyset$. Otherwise a maximal chain C in Y^* (containing x) exists with no g. l. b. Let Z be an inversely well-ordered chain with $Z \subset C$

---

[2] Part (i) and the particular assertions of (ii) and (iii) concerning chains were published (without proofs) in [8, Theorem 6]. Part (i) and the particular assertion of (ii) concerning chains are contained in [10, Theorem 9]. But the proof of (i) is simple and is therefore included.

and $Z^+=C^+$ (see Lemma (C)). Then (1) and (2) hold ($u\in Y^*\cap Z^+$ would imply that u is a least element of C).

LEMMA 2. *If P contains a well-ordered chain Y and an inversely well-ordered chain Z (one of them can be empty) such that* (1) *and* (2) *of Lemma 1 hold then a directable mapping from P to P exists with no fixed point.*

The proof can easily be obtained from the proof of Theorem 1.2 in [14].

LEMMA 3. *Let C be a chain in P and M a finite subset of P such that to each $x\in C$, an element y in M exists with $y\leqq x$. Then there is an element y_0 in M such that $y_0\leqq x$ for all $x\in C$.*

PROOF. Given $y\in M$, let $C(y)=\{x\in C\colon y\leqq x\}$. Then $C=\cup\{C(y)\colon y\in M\}=$ $=C(y_0)$ for some $y_0\in M$.

LEMMA 4. *Let $f\colon P\to P$ be an isotone mapping, and let S be a non-empty subset of P satisfying the finite minimal condition. Suppose to each $x\in S$ a positive integer n exists with $f^n(x)\in S$. Then there is a minimal element x of S and $k\in N$ such that $f^k(x)\in S$ and $x\leqq f^k(x)$.*

PROOF. Let $M=\{a_1, \ldots, a_n\}$ be the set of all minimal elements of S and let t_i be positive integers with $f^{t_i}(a_i)\in S$ $(i=1, \ldots, n)$. We shall define a sequence $b_1, b_2, \ldots$ as follows. Put $b_1=a_1$ and let b_2 be such $a_i\in M$ that $a_i\leqq f^{t_1}(b_1)$. Suppose $b_1, \ldots, b_j$ were found. Then $b_j=a_m$ for some $a_m\in M$. We let b_{j+1} be such an element a_p of M that $a_p\leqq f^{t_m}(b_j)$. Then $b_{r+1}\leqq f^{t(r)}(b_r)$ holds for $r=1, 2, \ldots$ $(t(r)\in N)$. In the sequence $b_1, \ldots, b_{n+1}$ some two members are equal. Hence $i, j\in N$ exist such that $i+j\leqq n+1$ and $b_i=b_{i+j}$. This follows $b_{i+1}\leqq f^{t(i)}(b_i)$, $b_{i+2}\leqq f^{t(i+1)}(b_{i+1})\leqq f^{t(i)+t(i+1)}(b_i), \ldots$ $\ldots, b_i=b_{i+j}\leqq f^s(b_i)$, where $s=\sum_{k=0}^{j-1} t(i+k)$, which proves the lemma.

PROOF OF THEOREM 1. Suppose a well-ordered chain Y in P exists such that Y^* fails to have the property stated in the theorem. According to Lemma 1 there exists an inversely well-ordered chain Z in P satisfying (1) and (2). By Lemma 2 a directable mapping $f\colon P\to P$ exists with no fixed point.

PROOF OF THEOREM 2. By the hypothesis, $P^*\neq\emptyset$, hence P has a greatest element. To prove the theorem it suffices to show that every chain in P has a g. l. b. and use the dual of the theorem (AB). Let C be a chain in P and let M be the (finite) set of all minimal elements of C^{++}. Since $C\subset C^{++}$, according to Lemma 3 an element $y\in M$ exists which belongs to C^+. Hence y is the greatest element of C^+, and the g. l. b. of C.

PROOF OF THEOREM 3. According to Lemma 4 there exist $x\in P$ and $k\in N$ such that $x\leqq f^k(x)$, hence $x\leqq f(x)$. Let $M=\{x\in P\colon x\leqq f(x)\}$ and let C be a maximal chain in M. If $y\in C^*$ then $x\leqq y$ hence $x\leqq f(x)\leqq f(y)$ for all $x\in C$ so that $f(y)\in C^*$. According to Lemma 4 there is an element $t\in C^*$ such that $t\leqq f^k(t)$ for some $k\in N$, hence $t\leqq f(t)$. If $t<f(t)$ then $C\cup\{f(t)\}$ is a chain in M greater than C which yields a contradiction. Hence $t=f(t)$.

PROOF OF THEOREM 4. Let $f\colon P\to P$ be isotone and $Q=\{x\in P\colon x\leqq f(x)\}$. Since $0\in Q$, Q is not empty. Let C be a maximal chain in Q. If l. u. b. of C exists (say c), then for each $x \in C$, $x \leqq c$, hence $x \leqq f(x) \leqq f(c)$. This follows $c \leqq f(c)$, $f(c) \leqq f(f(c))$, hence $f(c) \in Q$ and $f(c) \in C$ in virtue of the maximality of C. This yields $f(c) \leqq c$ hence $f(c) = c$.

Let l. u. b. of C do not exist. By the hypothesis, C^* has a greatest element ($=\inf \emptyset$). Moreover $u \in C^*$ implies $x \leqq u$ for each $x \in C$, hence $x \leqq f(x) \leqq f(u)$, and $f(u) \in C^*$. Applying the dual of the theorem (AB) to the partial mapping $f|C^*$ we get the assertion.

Proof of Theorem 5. Let $f: P \to P$ be isotone. Obviously, there is a sequence $(a_i: i \in N)$, $a_i \in \max(P)$, such that $f(a_i) \leqq a_{i+1}$ for each $i \in N$. According to the finiteness of $\max(P)$ there are $i, j \in N$ such that $i < j$ and $a_i = a_j$. Denote $S = \{a_i, a_{i+1}, \ldots, a_{j-1}\}$. Then $f(S^+) \subset S^+$ hence there is $x \in S^+$ with $f(x) = x$.

Proof of Theorem 6. From the hypothesis of the theorem it follows immediately that P has a least element. By [7, Theorem 2] the condition (a) is satisfied, hence theorem (AB) is applicable.

Proof of Theorem 7. Let P have the f. p. p. In [12, Theorem 7] there is shown that the f. p. p. implies compactness if P satisfies the condition

(i) For each $a \in P$, the subset $[a)$ is a lattice. But (i) is a consequence of (*). Indeed, let $u, v \in [a)$. Let C be a maximal chain in $\{u, v\}^+$, containing a. The elements u, v are in C^* and there is $z \in C^*$, $z \leqq u$, $z \leqq v$. $C \cup \{z\}$ is a chain contained in $\{u, v\}^+$, hence $C \cup \{z\} = C$ and $z \in C$. Hence z is a greatest element in C, thus maximal in the set $\{u, v\}^+$. Since P is a $\vee$-semilattice, z is the greatest element of $\{u, v\}^+$. This proves P to be compact. The converse implication follows by the dual of Theorem 6.

Proof of Theorem 8. The set $P(f)^+$ is not empty and satisfies (a). Moreover if $x \in P(f)^+$ then $f(x) \leqq f(c) = c$ for each $c \in P(f)$ so that $f(x) \in P(f)^+$. Using the theorem (AB) to the partial mapping $f|P(f)^+$ we get that f has a fixed point in $P(f)^+$ which is obviously a least element of $P(f)$. This proves (i). Suppose a subset A of $P(f)$ has a l. u. b. in P. If $y \in A^*$ then for each $x \in A$, $x \leqq y$ thus $x = f(x) \leqq f(y)$ hence $f(y) \in A^*$. According to (i), the set $A^*(f|A^*)$ has a least element which is a l. u. b. of A in $P(f)$. This proves (ii). The proof of (iii) is straightforward.

Proof of Corollary 2. According (i), $P(f^2)$ has a least element s. Since $f^2(f(s)) = f(s)$, $s \leqq f(s)$.

REFERENCES

[1] Abian, S. and Brown, A. B., A theorem on partially ordered sets, with applications to fixed point theorems, *Canad. J. Math.* **13** (1961), 78—82. *MR* **23** # A817.

[2] Benado, M., A propos des théorèmes de point fixe dans la théorie des ensembles partiellement ordonnés, preprint, 1961.

[3] Benado, M., Bemerkungen zur Theorie der Vielverbände. IV. Über die Möbius'sche Funktion, *Proc. Cambridge Philos. Soc.* **56** (1960), 291—317. *MR* **22** # 10929.

[4] Birkhoff, G., *Lattice theory*, Third edition, American Mathematical Society Colloquium Publications, Vol. XXV, American Mathematical Society, Providence, R. I., 1967. *MR* **37** # 2638.

[5] Davis, A. C., A characterization of complete lattices, *Pacific J. Math.* **5** (1955), 311—319. *MR* **17**—574.

[6] Duffus, D., Poguntke, W. and Rival, I., Retracts and the fixed point problem for finite partially ordered sets, *Canad. Math. Bull.* **23** (1980), 231—236.

[7] KOGALOVSKIĬ, S. R., On a theorem of Frink, *Uspehi Mat. Nauk* **19** (1964), no. 2 (116), 143—145 (in Russian). *MR* **28** # 5012.

[8] KOLIBIAR, M., Über Fixpunktsätze in geordneten Mengen, *Spisy Přírod. Fak. Univ. Brno* **457** (1964), 469—472.

[9] KOLIBIAR, M., Bemerkungen über Intervalltopologie in halbgeordneten Mengen, *General Topology and its Relations to Modern Analysis and Algebra* (Proc. Sympos., Prague, 1961), Academic Press, New York; Publ. House Czech. Acad. Sci., Prague, 1962, 252—253. *MR* **26** # 4940.

[10] MARKOWSKY, G., Chain-complete posets and directed sets with applications, *Algebra Universalis* **6** (1976), 53—68. *MR* **53** # 2764.

[11] MUENZENBERGER, T. B. and SMITHSON, R. E., Fixed point structures, *Trans. Amer. Math. Soc.* **184** (1973), 153—173. *MR* **48** # 7242.

[12] MUENZENBERGER, T. B. and SMITHSON R. E., Characterizations of compactness of the interval topology in semilattices, *Proc. Amer. Math. Soc.* **46** (1974), 133—136. *MR* **49** # 10615.

[13] ROTH, A. E., A lattice fixed-point theorem with constraints, *Bull. Amer. Math. Soc.* **81** (1975), 136—138. *MR* **50** # 12839.

[14] SMITHSON, R. E., Fixed points in partially ordered sets, *Pacific J. Math.* **45** (1973), 363—367. *MR* **47** # 4871.

[15] TARSKI, A., A lattice-theoretical fixpoint theorem and its applications, *Pacific J. Math.* **5** (1955), 285—309. *MR* **17**—574.

[16] WARD JR, L. E., Completeness in semi-lattices, *Canad. J. Math.* **9** (1957), 578—582. *MR* **19**—938.

[17] WOLK, E. S., Dedekind completeness and a fixed-point theorem, *Canad. J. Math.* **9** (1957), 400—405. *MR* **19**—243.

(Received January 11, 1982)

KATEDRA ALGEBRY A TEÓRIE ČÍSEL
MATEMATICKO-FYZIKÁLNEJ FAKULTY
UNIVERZITY KOMENSKÉHO
MATEMATICKY PAVILÓN
MLYNSKÁ DOLINA
ČS—842 15 BRATISLAVA
CZECHOSLOVAKIA

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 51—67

INTEGRABLE SOLUTIONS OF FUNCTIONAL EQUATIONS OF A GENERAL TYPE

J. ACZÉL and J. K. CHUNG

Dedicated to the memory of L. Rédei, a great mathematician, teacher, and friend

1. Introduction

We deal here with functional equations of the form

$$\sum_{i=1}^{m} F_i(\alpha_i x+\beta_i y) = \sum_{k=1}^{n} p_k(x) q_k(y), \tag{1}$$

special cases of which were considered in [1—12] (see also the references in these works). We will assume (1) to be valid for $x \in]A, B[$, $y \in]C, D[$ (say; real rectangles). We may suppose that

$$\alpha_i : \beta_i \neq \alpha_j : \beta_j \quad \text{if} \quad i \neq j, \tag{2}$$

for else the respective terms on the left-hand side of (1) could be combined into single terms. Similarly, we may suppose

$$\alpha_i \beta_i \neq 0 \quad \text{for all} \quad i = 1, 2, \ldots, m, \tag{3}$$

for else, if some α_i or β_i is 0, the corresponding $\tilde{F}_i(y)=F_i(\beta_i y)$ or $\bar{F}_i(x)=F_i(\alpha_i x)$ can be transferred to the right-hand side of (1).

It is also natural to assume that $p_1, p_2, \ldots, p_n$ are linearly independent and so are $q_1, q_2, \ldots, q_n$, for else the right-hand side of (1) may be replaced by a similar sum of less than n terms. We will suppose a bit more, that the p_k and also the q_k are *L-independent* on $]A, B[$ or $]C, D[$, respectively, which means, for instance for the q_k's, that $\sum_{k=1}^{n} a_k q_k(y)=0$ *almost everywhere* on $]C, D[$ *implies* $a_k=0$ $(k=1, 2, \ldots, n)$. Clearly, *L-independence implies linear independence*.

In view of (2) and (3), we can write (1) as

$$\sum_{i=1}^{m} f_i(x+\lambda_i y) = \sum_{k=1}^{n} p_k(x) q_k(y), \tag{4}$$

where $\lambda_i \neq 0$ and $\lambda_i \neq \lambda_j$ for $j \neq i = 1, 2, \ldots, m$. We will first prove, in Section 2, that f_i, p_k, q_k $(i=1, 2, \ldots, m;\ k=1,2, \ldots, n)$ have derivatives of all orders if, in addition to the above assumptions, $f_1, f_2, \ldots, f_m$ are supposed to be Lebesgue integrable on all closed intervals (really f_i is supposed to be just locally integrable

---

1980 *Mathematics Subject Classification*. Primary 39B30; Secondary 28A20.

Key words and phrases. Lebesgue integrable, functional equations, L-independent, functional determinant, differentiable, C^∞, linear homogeneous differential equations and systems, eigenvalues, polynomials.

on $]A, B[+\lambda_i]C, D[$). — Then we will show, in Section 3, on the example of the equation

$$f(x+y)+g(x-y) = \sum_{i=1}^{n} p_i(x)q_i(y), \tag{5}$$

how *all* solutions can explicitly be determined under the above conditions (including the Lebesgue integrability of f and g). As a particular case, we will enumerate for $n=2$ the solutions of (5) in Section 4.

General results on continuous or measurable solutions of (5) on topological groups are contained a.o. in [12]. The present results, on the other hand, are local and explicit, the proofs elementary. Further, it could be possible to weaken the integrability condition in this paper to, say, measurability by applying results and methods of A. Járai (cf. [4]).

2. A differentiability theorem

Generalizing a method worked out in [5] and more completely in [2], we prove the following

THEOREM 1. *If, in*

$$\sum_{i=1}^{m} f_i(x+\lambda_i y) = \sum_{k=1}^{n} p_k(x)q_k(y) \qquad (x\in]A, B[,\ y\in]C, D[), \tag{4}$$

the functions $p_1, p_2, ..., p_n$ *and* $q_1, q_2, ..., q_n$ *are L-independent on* $]A, B[$ *or* $]C, D[$, *respectively (cf. Section 1),*

$$\lambda_i \neq 0 \quad \textit{and} \quad \lambda_i \neq \lambda_j \quad \textit{for} \quad j \neq i = 1, 2, ..., m$$

and the f_i *are locally Lebesgue-integrable on* $]A, B[+\lambda_i]C, D[$ *(cf. Section 1), then the functions* f_i, p_k *and* q_k *have derivatives of all orders (are* C^∞*) on their respective intervals. Moreover,* $p_1, p_2, ..., p_n$ *and also* $q_1, q_2, ..., q_n$ *satisfy systems of explicit homogeneous linear differential equations of m-th order with constant coefficients and, if* 0 *is both in* $]A, B[$ *and in* $]C, D[$, *then the* $f_i^{(m-1)}(t)$ *are linear combinations of the* $p_k^{(s)}(t)$ *and also of the* $q_k^{(s)}\left(\frac{t}{\lambda_i}\right)$ $(i=1, 2, ..., m;\ k=1, 2, ..., n;\ s=0, 1, ..., m-1)$.

PROOF. We first prove that the p_k's and, similarly, the q_k's are Lebesgue integrable on $]A, B[$ or $]C, D[$, respectively. Since the q_k's are (L-independent and thus) linearly independent, there exist $y_1, y_2, ..., y_n$ $(\in]C, D[)$ such that $\det q_k(y_j)\neq 0$ (see e.g. [1], pp. 201—203). Substituting these y_j for y in (4), we get n equations which can be uniquely solved with respect to $p_1(x), p_2(x), ..., p_n(x)$. By the Cramer rule we get $p_k(x)=\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ijk}f_i(x+\lambda_i y_j)$. Since the f_i are Lebesgue integrable, so are the p_k and, similarly, the q_k.

We now integrate (4) with respect to y (from $C^*\in]C, D[$):

$$\sum_{i=1}^{m} \int_{C^*}^{t} f_i(x+\lambda_i y)\,dy = \sum_{k=1}^{n} p_k(x) \int_{C^*}^{t} q_k(y)\,dy,$$

then introduce into each integrand on the left-hand side individually new variables $s=x+\lambda_i y$ and obtain

$$(6)\qquad \sum_{i=1}^{m}\frac{1}{\lambda_i}\int_{x+\lambda_i C^*}^{x+\lambda_i t} f_i(s)\,ds=\sum_{k=1}^{n} p_k(x)Q_k(t)$$

where

$$Q_k(t)=\int_{C^*}^{t} q_k(y)\,dy.$$

The functions Q_k are linearly independent for else there would exist n constants a_k $(k=1, 2, \ldots, n)$, not all zero, such that

$$\sum_{k=1}^{n} a_k Q_k(t)\equiv 0,\quad \text{i.e.}\quad \int_{C^*}^{t}\left[\sum_{k=1}^{n} a_k q_k(y)\,dy\right]\equiv 0\quad \text{for all}\quad t\in C, D.$$

But then $\sum_{k=1}^{n} a_k q_k(y)=0$ almost everywhere on $]C, D[$ which, by the L-independence of the q_k, would imply $a_k=0$ $(k=1, 2, \ldots, n)$, contrary to the supposition.

Similarly to our previous argument, the linear independence of the Q_k and (6) imply that the p_k are linear combinations of the $x\mapsto \int_{x+\lambda_i C^*}^{x+\lambda_i t} f_i(s)ds$ and, since the latter are continuous, so are the p_k (on $]A, B[$) and, similarly, the q_k (on $]C, D[$).

We now prove that the f_i (for instance f_1) are continuous, too. For this purpose, we substitute $u=x+\lambda_1 y$ into (4) and integrate the equation

$$f_1(u)=\sum_{k=1}^{n} p_k(u-\lambda_1 y)q_k(y)-\sum_{i=2}^{m} f_i[u+(\lambda_i-\lambda_1)y],$$

thus obtained, with respect to y:

$$(t-C^*)f_1(u)=\sum_{k=1}^{n}\int_{C^*}^{t} p_k(u-\lambda_1 y)q_k(y)\,dy-\sum_{i=2}^{m}\int_{C^*}^{t} f_i[u+(\lambda_i-\lambda_1)y]\,dy.$$

Introducing, into each integrand in the second sum on the right-hand side individually, the new variables $v=u+(\lambda_i-\lambda_1)y$ we get

$$(7)\qquad (t-C^*)f_1(u)=\sum_{k=1}^{n}\int_{C^*}^{t} p_k(u-\lambda_1 y)\,q_k(y)\,dy-\sum_{i=2}^{m}\frac{1}{\lambda_i-\lambda_1}\int_{u+(\lambda_i-\lambda_1)C^*}^{u+(\lambda_i-\lambda_1)t} f_i(v)\,dv.$$

The right-hand side of this equation is a continuous function of u, thus so is f_1 and, similarly, all f_i.

However, if all f_i are continuous then $x\mapsto \int_{x+\lambda_i C^*}^{x+\lambda_i t} f_i(s)ds$ is differentiable.

Thus, applying the above argument again to (6), all p_k (on $]A, B[$) and similarly, all q_k (on $]C, D[$) are differentiable.

Moreover, the p_k and q_k are C^1. Indeed, differentiate (6) with respect to x and obtain

$$\sum_{i=1}^{m} [f_i(x+\lambda_i t)-f_i(x+\lambda_i C^*)]/\lambda_i = \sum_{k=1}^{n} p_k'(x)Q_k(t). \tag{8}$$

Since the left-hand side is continuous and the functions Q_k are linearly independent, we get, as before, that the p_k' and, similarly, the q_k' are continuous.

Since all integrands are continuous on the right-hand side of (7) and the p_k are C^1, f_1 is differentiable, and so are all f_i. Now, from (8) and again from the possibility to choose y_j so that $\det Q_k(y_j)\neq 0$ (the Q_k being linearly independent), the p_k are twice differentiable and, similarly, so are the q_k. From (7) again f_1 and, similarly, all f_i are twice differentiable, etc.

We differentiate now (4) $(m-j)$ times with respect to x and j times with respect to y $(j=0, 1, \ldots, m)$ and get

$$\sum_{i=1}^{m} f_i^{(m)}(x+\lambda_i y) = \sum_{k=1}^{m} p_k^{(m)}(x)\,q_k(y),$$

$$\sum_{i=1}^{m} \lambda_i f_i^{(m)}(x+\lambda_i y) = \sum_{k=1}^{n} p_k^{(m-1)}(x)\,q_k'(y),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\sum_{i=1}^{m} \lambda_i^{m-1} f_i^{(m)}(x+\lambda_i y) = \sum_{k=1}^{n} p_k'(x)\,q_k^{(m-1)}(y),$$

$$\sum_{i=1}^{m} \lambda_i^{m} f_i^{(m)}(x+\lambda_i y) = \sum_{k=1}^{n} p_k(x)\,q_k^{(m)}(y).$$

It follows from this system of $m+1$ homogeneous linear equations in $f_1^{(m)}(x+\lambda_1 y), \ldots$ $\ldots, f_m^{(m)}(x+\lambda_m y)$, 1 that

$$\begin{vmatrix} 1 & 1 & \ldots\ 1 & \sum_{k=1}^{n} p_k^{(m)}(x)\,q_k(y) \\ \lambda_1 & \lambda_2 & \ldots\ \lambda_m & \sum_{k=1}^{n} p_k^{(m-1)}(x)\,q_k'(y) \\ \vdots & & \ldots & \\ \lambda_1^{m-1} & \lambda_2^{m-1} & \ldots\ \lambda_m^{m-1} & \sum_{k=1}^{n} p_k'(x)\,q_k^{(m-1)}(y) \\ \lambda_1^{m} & \lambda_2^{m} & \ldots\ \lambda_m^{m} & \sum_{k=1}^{n} p_k(x)\,q_k^{(m)}(y) \end{vmatrix} = 0,$$

that is,

$$(9)\qquad A_1 \sum_{k=1}^{n} p_k^{(m)}(x) q_k(y) - A_2 \sum_{k=1}^{n} p_k^{(m-1)}(x) q_k'(y) + \dots$$
$$\dots + (-1)^{m-1} A_m \sum_{k=1}^{n} p_k(x) q_k^{(m)}(y) = 0,$$

where each A_i is the determinant of the $(m+1)\times m$ matrix obtained from

$$\begin{pmatrix} 1 & 1 & \dots 1 \\ \lambda_1 & \lambda_2 & \dots \lambda_m \\ \vdots & \vdots & \vdots \\ \lambda_1^m & \lambda_2^m & \dots \lambda_m^m \end{pmatrix}$$

by deleting the i-th row, and hence

$$A_1 = \begin{vmatrix} \lambda_1 & \lambda_2 & \dots \lambda_m \\ \lambda_1^2 & \lambda_2^2 & \dots \lambda_m^2 \\ \vdots & \vdots & \vdots \\ \lambda_1^m & \lambda_2^m & \dots \lambda_m^m \end{vmatrix} = \lambda_1\lambda_2\dots\lambda_m \begin{vmatrix} 1 & 1 & \dots 1 \\ \lambda_1 & \lambda_2 & \dots \lambda_m \\ \vdots & \vdots & \vdots \\ \lambda_1^{m-1} & \lambda_2^{m-1} & \dots \lambda_m^{m-1} \end{vmatrix} =$$
$$= \lambda_1\lambda_2\dots\lambda_m A_m = \lambda_1\lambda_2\dots\lambda_m \prod_{1\le i<j\le m} (\lambda_i - \lambda_j) \ne 0.$$

Since $q_1, q_2, \dots, q_n$ are (L-independent and thus) linearly independent, there again exist $y_1, y_2, \dots, y_n$ $(\in]C, D[)$, such that $\det q_k(y_j) \ne 0$ $(k, j = 1, \dots, n)$. Substituting these y_j into (9) we indeed obtain

$$(10)\qquad p_k^{(m)}(x) = \sum_{j=1}^{n} \sum_{s=0}^{m-1} a_{kjs} p_j^{(s)}(x) \quad (k = 1, \dots, n).$$

Similarly, the $q_k(y)$ satisfy

$$(11)\qquad q_k^{(m)}(y) = \sum_{j=1}^{n} \sum_{s=0}^{m-1} b_{kjs} q_j^{(s)}(y) \quad (k = 1, \dots, n).$$

REMARK. Substituting (10) and (11) into (9) we obtain

$$A_1 \sum_{k=1}^{n} \sum_{j=1}^{n} \sum_{s=0}^{m-1} a_{kjs} p_j^{(s)}(x) q_k(y) - A_2 \sum_{k=1}^{n} p_k^{(m-1)}(x) q_k'(y) + \dots$$
$$\dots + (-1)^{m-2} A_{m-1} \sum_{k=1}^{n} p_k'(x) q_k^{(m-1)}(y) + (-1)^{m-1} A_m \sum_{k=1}^{n} \sum_{j=1}^{n} \sum_{s=0}^{m-1} b_{kjs} p_k(x) q_j^{(s)}(y) = 0,$$

which can modify the systems (10) or (11).

In order to express $f_i^{(m-1)}$ with the aid of $p_k^{(s)}$ or $q_k^{(s)}$ $(s=0, 1, \dots, m-1; i=1, \dots, m; k=1, \dots, n)$, we differentiate (4) $(m-1-l)$ times with respect to x and l times with respect to y $(l=0, 1, \ldots, m-1)$ and get

$$\sum_{i=1}^{m} f_i^{(m-1)}(x+\lambda_i y) = \sum_{k=1}^{n} p_k^{(m-1)}(x) q_k(y),$$

$$\sum_{i=1}^{m} \lambda_i f_i^{(m-1)}(x+\lambda_i y) = \sum_{k=1}^{n} p_k^{(m-2)}(x) q_k'(y),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\sum_{i=1}^{m} \lambda_i^{m-2} f_i^{(m-1)}(x+\lambda_i y) = \sum_{k=1}^{n} p_k'(x) q_k^{(m-2)}(y),$$

$$\sum_{i=1}^{m} \lambda_i^{m-1} f_i^{(m-1)}(x+\lambda_i y) = \sum_{k=1}^{n} p_k(x) q_k^{(m-1)}(y).$$

By setting here $y=0$ or $x=0$ it follows that

$$f_i^{(m-1)}(x) = \sum_{k=1}^{n} \sum_{s=0}^{m-1} \alpha_{iks} p_k^{(s)}(x)$$

and

$$f_i^{(m-1)}(\lambda_i y) = \sum_{k=1}^{n} \sum_{s=0}^{m-1} \beta_{iks} q_k^{(s)}(y),$$

respectively. □

3. The equation (5)

In this section we consider the functional equation

$$f(x+y)+g(x-y) = \sum_{i=1}^{n} p_i(x) q_i(y) \qquad (x\in]A, B[, \ y\in]C, D[). \tag{5}$$

We can improve in this case Theorem 1 in the following way.

THEOREM 2. *Let $p_1, p_2, \ldots, p_n$ and $q_1, q_2, \ldots, q_n$ be L-independent on $]A, B[$ and $]C, D[$ (where $]C, D[$ contains 0), respectively, f and g Lebesgue integrable on $]A, B[+]C, D[$ and $]A, B[-]C, D[$, respectively, and (5) satisfied. Then f, g, $p_1, p_2, \ldots, p_n, q_1, q_2, \ldots, q_n$ have derivatives of all orders and $p_1, p_2, \ldots, p_n$ and also $q_1, q_2, \ldots, q_n$ satisfy systems of explicit second order homogeneous linear differential equations with constant coefficients without first order terms, while $f+g$ and $f'-g'$ satisfy linear homogeneous differential equations of order $2n$ with all terms of even order.*

PROOF. It is a consequence of Theorem 1 that all functions in (5) are C^∞. Now differentiate (5) twice with respect to x resp. y and get

$$\sum_{i=1}^{n} p_i''(x) q_i(y) = f''(x+y)+g''(x-y) = \sum_{i=1}^{n} p_i(x) q_i''(y). \tag{12}$$

Again, since $q_1, \ldots, q_k$ are linearly independent, there exist $y_1, \ldots, y_n$ such that $\det q_k(y_j) \neq 0$. Putting these y_j into (12) we get

$$\sum_{i=1}^{n} p_i''(x) q_i(y_j) = \sum_{i=1}^{n} p_i(x) q_i''(y_j) \qquad (j = 1, \ldots, n).$$

Thus we have, again, n linear equations for $p_i''(x)$ $(i=1, \ldots, n)$, whose determinant is different from 0. Therefore each $p_i''(x)$ is a linear combination, with constant coefficients, of $p_k(x)$ $(k=1, \ldots, n)$

(13) $$p_i''(x) = \sum_{k=1}^{n} a_{ik} p_k(x) \qquad (i = 1, \ldots, n).$$

Similarly, also the q_k satisfy an explicit system of linear differential equations of second order with constant coefficients and without first order terms, as asserted.

As to f and g, put $y=0$ into (5) in order to get, with

$$\Phi = f + g$$

and with $c_{0,i} = q_i(0)$, the equation

(14_0) $$\Phi(x) = \sum_{i=1}^{n} c_{0,i} p_i(x).$$

Differentiating this equation twice, we have

$$\Phi''(x) = \sum_{i=1}^{n} c_{0,i} p_i''(x)$$

or, using (13) with $c_{1,k} := \sum_{j=1}^{n} c_{0,j} a_{jk}$ $(k=1, \ldots, n)$,

(14_1) $$\Phi''(x) = \sum_{i=1}^{n} c_{1,i} p_i(x).$$

Similarly,

(14_j) $$\Phi^{(2j)}(x) = \sum_{i=1}^{n} c_{j,i} p_i(x) \qquad (j = 0, \ldots, n),$$

in particular,

(14_{n-1}) $$\Phi^{(2n-2)}(x) = \sum_{i=1}^{n} c_{n-1,i} p_i(x),$$

(14_n) $$\Phi^{(2n)}(x) = \sum_{i=1}^{n} c_{n,i} p_i(x).$$

There are two cases. If $\det c_{j,i} = 0$ $(j=0, \ldots, n-1;\ i=1, \ldots, n)$ then, from (14_0), (14_1), ..., (14_{n-1}), there exist $a_0, a_1, \ldots, a_{n-1}$, not all 0, such that

$$\sum_{j=0}^{n} a_j \Phi^{(2j)}(x) = 0.$$

If, on the other hand, $\det c_{j,i} \neq 0$ then $p_1, \dots, p_n$ can be determined from (14_0), (14_1), …, (14_{n-1}) as linear combinations of $\Phi, \Phi'', \dots, \Phi^{(2n-2)}$ so that (14_n) goes over into

$$\Phi^{(2n)}(x) = \sum_{j=0}^{n-1} a_j \Phi^{(2j)}(x). \tag{15}$$

In both cases Φ satisfies a homogeneous linear differential equation of order $2n$ with all terms of even order and with constant coefficients (not all 0).

If we differentiate (5) with respect to y, put $y=0$ and follow the above procedure, we get a similar equation for $f'-g'$. □

As the reader may have noticed, the result in Theorem 2 about the p_k (and also the q_k), that they satisfy explicit systems of homogeneous linear differential equations with constant coefficients, corresponds to the statement in Theorem 1 that they satisfy such m-th order equations. But there is no equivalent in Theorem 1 to the absence of first order terms in those equations as stated in Theorem 2. This returns *if we* step over into the complex field and *have* $\lambda_l = \varepsilon^l$ *where* $\varepsilon = e^{2\pi i/m}$ *is an m-th root of unity: then the derivatives of orders* $2, \dots, m-1$ *are missing in (10) and (11). In this case* we have also an analogue of the last statement of Theorem 2: *the functions* $\Phi_j = \sum_{l=1}^{m} \varepsilon^{jl} f_l^{(j)}$ *satisfy homogeneous linear differential equations of mn-th order with constant coefficients, having only terms of pm-th order* ($p=0, 1, \dots, n$).

Now we are going to give a general form of the integrable solutions of (5). (For other results on (5), see [2, 10—12].)

THEOREM 3. *Suppose that the functions f, g are Lebesgue integrable on* $]A, B[+]C, D[$, *and* $]A, B[-]C, D[$, *respectively, and suppose that the functions* $p_1, \dots, p_n$ *and also* $q_1, \dots, q_n$ *are L-independent on* $]A, B[$ *or* $]C, D[$, *respectively. Then all solutions of the functional equation* (5) *for* $x \in]A, B[$, $y \in]C, D[$ *(with* $0 \in]C, D[$*) can be written in the form*

$$\begin{aligned}
f(x) &= \sum_{l=1}^{2} \sum_{k=1}^{m} \sum_{j=0}^{n_k-1} A_{kj}^{1l} x^j e^{(-1)^{l-1} w_k x} + \sum_{j=0}^{2N} B_j^1 x^j, \\
g(x) &= \sum_{l=1}^{2} \sum_{k=1}^{m} \sum_{j=0}^{n_k-1} A_{kj}^{2l} x^j e^{(-1)^{l-1} w_k x} + \sum_{j=0}^{2N} B_j^2 x^j, \\
p_i(x) &= \sum_{l=1}^{2} \sum_{k=1}^{m} \sum_{j=0}^{n_k-1} A_{ikj}^{ll} x^j e^{(-1)^{l-1} w_k x} + \sum_{j=0}^{2N-1} B_{ij}^1 x^j,
\end{aligned} \tag{16}$$

$$q_i(y) = \sum_{l=1}^{2} \sum_{k=1}^{m} \sum_{j=0}^{n_k-1} A_{ikj}^{3-l,l} y^j e^{(-1)^{l-1} w_k y} + \sum_{j=0}^{2N-1} B_{ij}^2 y^j, \quad \left(i = 1, \dots, n;\ \sum_{0}^{-1} = 0\right),$$

where the n_k, A and m are integers satisfying

(17)
$$\sum_{k=1}^{m} n_k + N = n \qquad \left(\sum_{1}^{0} = 0\right).$$

The w_k are, in general, complex nonzero constants, and the constants A_{kj}^{sl}, A_{ikj}^{sl}, B_j^l and B_{ij}^l satisfy

$$A_{k0}^{1l} = \sum_{i=1}^{n} A_{ik0}^{ll} A_{ik0}^{3-l,l}, \quad A_{k0}^{2l} = \sum_{i=0}^{n} A_{ik0}^{ll} A_{ik0}^{l,3-l} \quad (l = 1, 2, \quad k = 1, \ldots, m),$$

$$\binom{M}{j} A_{kM}^{1l} = \sum_{i=1}^{n} A_{ik,M-j}^{ll} A_{ikj}^{3-l,l}, \quad (-1)^j \binom{M}{j} A_{kM}^{2l} = \sum_{i=1}^{n} A_{ik,M-j}^{ll} A_{ikj}^{l,3-l}$$

$$(l = 1, 2;\ j = 0, 1, \ldots, M;\ M = 1, 2, \ldots, n_k - 1;\ k = 1, \ldots, m),$$

$$\sum_{i=1}^{n} A_{ikp}^{ll} A_{ikq}^{3-l,l} = \sum_{i=1}^{n} A_{ikp}^{ll} A_{ikq}^{l,3-l} = 0$$

$$(l = 1, 2;\ p, q = 1, \ldots, n_k - 1;\ p+q > n_k - 1;\ k = 1, \ldots, m),$$

(18)
$$\sum_{i=1}^{n} A_{ikp}^{ll} A_{isq}^{3-l,l} = \sum_{i=1}^{n} A_{ikp}^{ll} A_{isq}^{l,3-l} = 0$$

$$(l = 1, 2;\ p = 0, 1, \ldots, n_k - 1;\ q = 0, 1, \ldots, n_s - 1;\ k, s = 1, \ldots, m;\ k \neq s);$$

$$\sum_{i=1}^{n} A_{ikj}^{ll} B_{is}^{2} = \sum_{i=1}^{n} A_{ikj}^{3-l,l} B_{is}^{1} = 0$$

$$(l = 1, 2;\ j = 0, 1, \ldots, n_k - 1;\ k = 1, \ldots, m;\ s = 0, 1, \ldots, 2N - 1);$$

$$B_0^1 + B_0^2 = \sum_{i=1}^{n} B_{i0}^1 B_{i0}^2, \quad \binom{M}{j} [B_M^1 + (-1)^j B_M^2] = \sum_{i=1}^{n} B_{i,M-j}^1 B_{ij}^2$$

$$(j = 0, 1, \ldots, M;\ M = 1, 2, \ldots, 2N - 1),$$

$$B_{(2N)}^1 + B_{(2N)}^2 = 0, \quad \sum_{i=1}^{n} B_{i,2N-k}^1 B_{ik}^2 = \begin{cases} 2\binom{2N}{k} B_{(2N)}^1, & k \text{ odd} \\ 0, & k \text{ even} \end{cases} \quad (k = 1, \ldots, 2N - 1),$$

$$\sum_{i=1}^{n} B_{ip}^1 B_{iq}^2 = 0 \quad (p, q = 2, \ldots, 2N - 1;\ p+q > 2N),$$

but are otherwise arbitrary, and conversely, all systems of functions of the form (16), *with* (17) *and* (18) *satisfy* (5).

The proof of Theorem 3 may be obtained from the following two lemmas.

LEMMA 1. *Under the conditions of Theorem 3, all solutions of the functional equa-*

tion (5) *can be written in the form*

$$
(19)\qquad
\begin{aligned}
f(x) &= \sum_{l=1}^{2}\sum_{k=1}^{m} P_k^{1l}(x)e^{(-1)^{l-1}w_k x}+H_{2N}^1(x),\\
g(x) &= \sum_{l=1}^{2}\sum_{k=1}^{m} P_k^{2l}(x)e^{(-1)^{l-1}w_k x}+H_{2N}^2(x),\\
p_i(x) &= \sum_{l=1}^{2}\sum_{k=1}^{m} P_{ik}^{ll}(x)e^{(-1)^{l-1}w_k x}+H_{i,2N-1}^1(x),\\
q_i(y) &= \sum_{l=1}^{2}\sum_{k=1}^{m} P_{ik}^{3-l,l}(y)e^{(-1)^{l-1}w_k y}+H_{i,2N-1}^2(y),\quad (i=1,\dots,n),
\end{aligned}
$$

where the $P_k^{sl}(x)$ *and* $P_{ik}^{sl}(x)$ *are polynomials of degree* (n_k-1), *the* $H_{2N}^l(x)$ *and* $H_{i\ 2N-1}^l(x)$ *are polynomials of degree* $2N$, $2N-1$, *respectively,*

$$
(17)\qquad \sum_{k=1}^{m} n_k+N=n \qquad \left(\sum_{k=1}^{0} n_k=0\right),
$$

and the w_k *are, in general, complex nonzero constants.*

PROOF. By supposition, the functions $p_1, \dots, p_n$ and also $q_1, \dots, q_n$ are linearly independent and, by Theorem 1, the functions f, g, p_i, q_i $(i=1, \dots, n)$ have derivatives of all orders.

If we differentiate (5) with respect to x and y, we obtain

$$
(20)\qquad f'(x+y)+g'(x-y)=\sum_{i=1}^{n} p_i'(x)q_i(y)
$$

and

$$
(21)\qquad f'(x+y)-g'(x-y)=\sum_{i=1}^{n} p_i(x)q_i'(y),
$$

respectively.

On the other hand, by Theorem 2, the p_i satisfy the explicit system (13) of second order homogeneous linear differential equations with constant coefficients, without first order terms.

Suppose that w_k are nonzero eigenvalues of multiplicity n_k of the matrix (a_{ij}) $(k=1, \dots, m)$ and the zero is an N-fold eigenvalue of the matrix (a_{ij}), then, from the theory of systems of homogeneous linear differential equations with constant coefficients, the general solution of (13) can be written in the form

$$
(22)\qquad p_i(x)=\sum_{k=1}^{m}[P_{ik}^{11}(x)e^{w_k x}+P_{ik}^{22}(x)e^{-w_k x}]+H_{i,2N-1}^1(x)\quad (i=1,\dots,n),
$$

where the P_{ik}^{ll} are polynomials of degree (n_k-1), and the $H_{i,2N-1}^1$ are polynomials of degree $(2N-1)$ $(l=1, 2;\ i=1, \dots, n;\ k=1, \dots, m)$, and

$$
(17)\qquad \sum_{k=1}^{m} n_k+N=n \qquad \left(\sum_{1}^{0}=0\right).
$$

When zero is not an eigenvalue of the matrix (a_{ij}), namely $N=0$, we define $H^1_{i,-1}(x)\equiv 0$.

In order to determine f and g, we set $y=0$ in (20) and (21), and apply (22), so that we obtain

$$f'(x) = \frac{1}{2}\sum_{i=1}^{n}[\alpha_i p_i'(x)+\beta_i p_i(x)] = \sum_{k=1}^{m}[\pi_k^{11}(x)e^{w_k x}+\pi_k^{12}(x)e^{-w_k x}]+h^1_{2N-1}(x)$$

and

$$g'(x) = \frac{1}{2}\sum_{i=1}^{n}[\alpha_i p_i'(x)-\beta_i p_i(x)] = \sum_{k=1}^{m}[\pi_k^{21}(x)e^{w_k x}+\pi_k^{22}(x)e^{-w_k x}]+h^2_{2N-1}(x).$$

By integrating these equations with respect to x we obtain

$$f(x) = \sum_{k=1}^{m}[P_k^{11}(x)e^{w_k x}+P_k^{12}(x)e^{-w_k x}]+H^1_{2N}(x) \tag{23}$$

and

$$g(x) = \sum_{k=1}^{m}[P_k^{21}(x)e^{w_k x}+P_k^{22}(x)e^{-w_k x}]+H^2_{2N}(x), \tag{24}$$

where the P_k^{sl} are polynomials of degree (n_k-1), the H^l_{2N} are polynomials of degree $2N$ $(s, l=1, 2;\ k=1, \dots, m)$.

Substituting (23) and (24) into (5) and setting $x=x_j$ $(j=1, \dots, n)$, such that $\det p_k(x_j)\neq 0$ $(k, j=1, \dots, n)$ (such x_j exist because of the linear independence of the p_k), we obtain

$$q_i(y) = \sum_{k=1}^{m}[P_{ik}^{21}(y)e^{w_k y}+P_{ik}^{12}(y)e^{-w_k y}]+H^2_{i,2N}(y) \quad (i = 1, \dots, n), \tag{25}$$

where the P_{ik}^{21} and P_{ik}^{12} are polynomials of degree (n_k-1), the $H^2_{i,2N}$ are polynomials of degree $2N$.

We now show that the $H^2_{i,2N}$ are polynomials of degree $2N-1$. In fact, let

$$\begin{aligned} H^l_{2N}(x) &= H^l_{2N-1}(x)+a^l_{(2N)}x^{2N} \qquad (l = 1, 2) \\ H^2_{i,2N}(y) &= H^2_{i,2N-1}(y)+a^2_{i,2N}y^{2N} \qquad (i = 1, \dots, n), \end{aligned} \tag{26}$$

where the H^l_{2N-1} and $H^2_{i,2N-1}$ are polynomials of degree $(2N-1)$. If we substitute (23)—(25) with (26) into (5), we get for the coefficients $a^1_{(2N)}$ and $a^2_{(2N)}$ of x^{2N} and y^{2N}

$$a^1_{(2N)}+a^2_{(2N)} = 0$$

and

$$a^1_{(2N)}+a^2_{(2N)} = \sum_{i=1}^{n} a^2_{i,2N}\,p_i(x).$$

Since the p_i are linearly independent, $a^2_{i,2N}=0$ $(i=1, \dots, n)$. Thus, it is also possible to write (25) in the form

$$q_i(y) = \sum_{k=1}^{m}[P_{ik}^{21}(y)e^{w_k y}+P_{ik}^{12}(y)e^{-w_k y}]+H^2_{i,2N-1}(y) \quad (i = 1, \dots, n). \quad \square \tag{27}$$

LEMMA 2. *If we introduce, in Lemma 1,*

$$P_k^{sl}(x) = \sum_{j=0}^{n_k-1} A_{kj}^{sl} x^j, \quad P_{ik}^{sl}(x) = \sum_{j=0}^{n_k-1} A_{ikj}^{sl} x^j,$$

(28)

$$H_{2N}^{l}(x) = \sum_{j=0}^{2N} B_j^l x^j, \quad H_{i,2N-1}^{l}(x) = \sum_{j=0}^{2N-1} B_{ij}^l x^j, \quad (s, l = 1, 2),$$

then the functions (19) *with* (28) *satisfy* (5) *if, and only if, the constants* A_{kj}^{sl}, A_{ikj}^{sl}, B_j^l *and* B_{ij}^l *fulfil* (18).

PROOF. We substitute (19) with (28) into (5):

$$\sum_{l=1}^{2}\sum_{k=1}^{m}\sum_{j=0}^{n_k-1} A_{kj}^{1l}(x+y)^j e^{(-1)^{l-1}w_k(x+y)} + \sum_{j=0}^{2N} B_j^1 (x+y)^j +$$

$$+ \sum_{l=1}^{2}\sum_{k=1}^{m}\sum_{j=0}^{n_k-1} A_{kj}^{2l}(x-y)^j e^{(-1)^{l-1}w_k(x-y)} + \sum_{j=0}^{2N} B_j^2 (x-y)^j =$$

$$= \sum_{i=1}^{n}\left[\sum_{l=1}^{2}\sum_{k=1}^{m}\sum_{j_1=0}^{n_k-1} A_{ikj_1}^{ll} x^{j_1} e^{(-1)^{l-1}w_k x} + \sum_{j=0}^{2N-1} B_{ij}^1 x^j\right]\times$$

$$\times\left[\sum_{\pi=1}^{2}\sum_{s=1}^{m}\sum_{j_2=0}^{n_s-1} A_{isj_2}^{3-\pi,\pi} y^{j_2} e^{(-1)^{\pi-1}w_s y} + \sum_{j=0}^{2N-1} B_{ij}^2 y^j\right],$$

that is,

$$\sum_{l=1}^{2}\sum_{k=1}^{m}\sum_{j=0}^{n_k-1} A_{kj}^{1l}(x+y)^j e^{(-1)^{l-1}w_k(x+y)} + \sum_{l=1}^{2}\sum_{k=1}^{m}\sum_{j=0}^{n_k-1} A_{kj}^{2l}(x-y)^j e^{(-1)^{l-1}w_k(x-y)} +$$

$$+ \sum_{j=0}^{2N} B_j^1 (x+y)^j + \sum_{j=0}^{2N} B_j^2 (x-y)^j =$$

$$= \sum_{i=1}^{n}\sum_{l=1}^{2}\sum_{k=1}^{m}\sum_{s=1}^{m}\sum_{j_1=0}^{n_k-1}\sum_{j_2=0}^{n_s-1} A_{ikj_1}^{ll} A_{isj_2}^{3-l,l} x^{j_1} y^{j_2} e^{(-1)^{l-1}(w_k x + w_s y)} +$$

(29)

$$+ \sum_{i=1}^{n}\sum_{l=1}^{2}\sum_{k=1}^{m}\sum_{s=1}^{m}\sum_{j_1=0}^{n_k-1}\sum_{j_2=0}^{n_s-1} A_{ikj_1}^{ll} A_{isj_2}^{l,3-l} x^{j_1} y^{j_2} e^{(-1)^{l-1}(w_k x - w_s y)} +$$

$$+ \sum_{i=1}^{n}\sum_{l=1}^{2}\sum_{k=1}^{m}\sum_{j=0}^{n_k-1}\sum_{s=0}^{2N-1} A_{ikj}^{ll} B_{is}^2 x^j y^s e^{(-1)^{l-1}w_k x} +$$

$$+ \sum_{i=1}^{n}\sum_{l=1}^{2}\sum_{k=1}^{m}\sum_{j=0}^{n_k-1}\sum_{s=0}^{2N-1} A_{ikj}^{3-l,l} B_{is}^1 x^s y^j e^{(-1)^{l-1}w_k y} +$$

$$+ \sum_{i=1}^{n}\sum_{j_1=0}^{2N-1}\sum_{j_2=0}^{2N-1} B_{ij_1}^1 B_{ij_2}^2 x^{j_1} y^{j_2}.$$

All terms of (29) may be divided into the following five types:

1° The terms containing $A^{ll}_{ksj_1j_2}x^{j_1}y^{j_2}e^{(-1)^{l-1}(w_kx+w_sy)}$.

2° The terms containing $B^{ll}_{ksj_1j_2}x^{j_1}y^{j_2}e^{(-1)^{l-1}(w_kx-w_sy)}$.

3° The terms containing $A^{ll}_{kjs}x^jy^se^{(-1)^{l-1}w_kx}$.

4° The terms containing $B^{ll}_{ksj}x^sy^je^{(-1)^{l-1}w_ky}$.

5° The terms containing $B^{l}_{j_1j_2}x^{j_1}y^{j_2}$.

Due to the fact that the functions contained in i° are independent of those in j° (i≠j=1, 2, 3, 4, 5), the necessary and sufficient condition for (29) to hold is that the sum of each of the five types of terms just mentioned should be equal on both sides of (29).

For type 1, we have

$$\sum_{l=1}^{2}\sum_{k=1}^{m}\sum_{j=0}^{n_k-1} A^{1l}_{kj}(x+y)^je^{(-1)^{l-1}w_k(x+y)} =$$

$$= \sum_{i=1}^{n}\sum_{l=1}^{2}\sum_{k=1}^{m}\sum_{s=1}^{m}\sum_{j_2=0}^{n_k-1}\sum_{j_2=0}^{n_s-1} A^{ll}_{ikj_1}A^{3-l,l}_{isj_2}x^{j_1}y^{j_2}e^{(-1)^{l-1}(w_kx+w_sy)},$$

which is possible only if

$$\sum_{k=1}^{m}\sum_{j=0}^{n_k-1} A^{1l}_{kj}(x+y)^je^{(-1)^{l-1}w_k(x+y)} =$$

$$= \sum_{i=1}^{n}\sum_{k=1}^{m}\sum_{s=1}^{m}\sum_{j_1=0}^{n_k-1}\sum_{j_2=0}^{n_s-1} A^{ll}_{ikj_1}A^{3-l,l}_{isj_2}x^{j_1}y^{j_2}e^{(-1)^{l-1}(w_kx+w_sy)} \quad (l = 1, 2).$$

It is easy to see that this implies

$$A^{1l}_{k0} = \sum_{i=1}^{n} A^{ll}_{ik0}A^{3-l,l}_{ik0}, \quad \binom{M}{j}A^{1l}_{kM} = \sum_{i=1}^{n} A^{ll}_{ik,M-j}A^{3-l,l}_{ikj}$$

$$(l = 1, 2;\ j = 0, 1, \ldots, M;\ M = 1, \ldots, n_k-1;\ k = 1, \ldots, m),$$

(30)
$$\sum_{i=1}^{n} A^{ll}_{ikp}A^{3-l,l}_{ikq} = 0$$

$$(l = 1, 2;\ p, q = 1, \ldots, n_k-1;\ p+q > n_k-1;\ k = 1, \ldots, m),$$

$$\sum_{i=1}^{n} A^{ll}_{ikp}A^{3-l,l}_{isq} = 0$$

$$(l = 1, 2;\ p = 0, 1, \ldots, n_k-1;\ q = 0, 1, \ldots, n_s-1;\ k, s = 1, \ldots, m;\ k \neq s).$$

For type 2, we have

$$\sum_{l=1}^{2}\sum_{k=1}^{m}\sum_{j=0}^{n_k-1} A_{kj}^{2l}(x-y)^j e^{(-1)^{l-1}w_k(x-y)} =$$

$$= \sum_{i=1}^{n}\sum_{l=1}^{2}\sum_{k=1}^{m}\sum_{s=1}^{m}\sum_{j_1=0}^{n_k-1}\sum_{j_2=0}^{n_s-1} A_{ikj_1}^{ll} A_{isj_2}^{l,3-l} x^{j_1} y^{j_2} e^{(-1)^{l-1}(w_k x - w_s y)},$$

which is possible only if

$$A_{k0}^{2l} = \sum_{i=1}^{n} A_{ik0}^{ll} A_{ik0}^{l,3-l}, \quad (-1)^j \binom{M}{j} A_{kM}^{2l} = \sum_{i=1}^{n} A_{ik,M-j}^{ll} A_{ikj}^{l,3-l}$$

$$(l = 1, 2;\ j = 0, 1, \ldots, M;\ M = 1, \ldots, n_k-1;\ k = 1, \ldots, m),$$

(31)
$$\sum_{i=1}^{n} A_{ikp}^{ll} A_{ikq}^{l,3-l} = 0$$

$$(l = 1, 2;\ p, q = 1, \ldots, n_k-1;\ p+q > n_k-1;\ k = 1, \ldots, m),$$

$$\sum_{i=1}^{n} A_{ikp}^{ll} A_{isq}^{l,3-l} = 0$$

$$(l = 1, 2;\ p = 0, 1, \ldots, n_k-1;\ q = 0, 1, \ldots, n_s-1;\ k, s = 1, \ldots, m;\ k \neq s).$$

For type 3, we have

$$\sum_{l=1}^{2}\sum_{k=1}^{m}\sum_{j=0}^{n_k-1}\sum_{s=0}^{2N-1}\left[\sum_{i=1}^{n} A_{ikj}^{ll} B_{is}^{2}\right] x^j y^s e^{(-1)^{l-1}w_k x} = 0,$$

which is possible only if

(32)
$$\sum_{i=1}^{n} A_{ikj}^{ll} B_{is}^{2} = 0$$

$$(l = 1, 2;\ j = 0, 1, \ldots, n_k-1;\ k = 1, \ldots, m;\ s = 0, 1, \ldots, 2N-1).$$

For type 4, we have

$$\sum_{l=1}^{2}\sum_{k=1}^{m}\sum_{j=0}^{n_k-1}\sum_{s=0}^{2N-1}\left[\sum_{i=1}^{n} A_{ikj}^{3-l,l} B_{is}^{1}\right] x^s y^j e^{(-1)^{l-1}w_k y} = 0,$$

which is possible only if

(33)
$$\sum_{i=1}^{n} A_{ikj}^{3-l,l} B_{is}^{1} = 0$$

$$(l = 1, 2;\ j = 0, 1, \ldots, n_k-1;\ k = 1, \ldots, m;\ s = 0, 1, \ldots, 2N-1).$$

For type 5, we have

(34)
$$\sum_{j=0}^{2N}[B_j^1(x+y)^j + B_j^2(x-y)^j] = \sum_{i=1}^{n}\sum_{j_1=0}^{2N-1}\sum_{j_2=0}^{2N-1} B_{ij_1}^{1} B_{ij_2}^{2} x^{j_1} y^{j_2},$$

which holds if and only if the coefficients of $x^{j_1}y^{j_2}$ $(j_1, j_2=0, 1, \dots, 2N)$ of both sides of (34) are equal.

For the constants we obtain

$$B_0^1+B_0^2=\sum_{i=1}^{n} B_{i0}^1 B_{i0}^2. \tag{35}$$

For the coefficients of $x^{M-j}y^j$, we have

$$\binom{M}{j}[B_M^1+(-1)^j B_M^2]=\sum_{i=1}^{n} B_{i,M-j}^1 B_{ij}^2 \tag{36}$$

$$(j=0, 1, \dots, M;\ M=1, \dots, 2N-1).$$

For the coefficients $B_{(2N)}^1$ and $B_{(2N)}^2$ of x^{2N} and y^{2N}, we have

$$B_{(2N)}^1+B_{(2N)}^2=0. \tag{37}$$

For the coefficients of $x^{2N-k}y^k$ we have, in view of (37),

$$\sum_{i=1}^{n} B_{i,2N-k}^1 B_{ik}^2=\begin{cases} 2\binom{2N}{k} B_{(2N)}^1, & k \text{ odd}, \\ 0, & k \text{ even} \end{cases} \quad (k=1, \dots, 2N-1). \tag{38}$$

For the coefficients of $x^p y^q$ with $p+q>2N$, we have

$$\sum_{i=1}^{n} B_{ip}^1 B_{iq}^2=0 \quad (p, q=2, 3, \dots, 2N-1;\ p+q>2N). \tag{39}$$

This completes the proof of Lemma 2, by (30)—(33) and (35)—(39). Thus Theorem 3 is also proved.

4. The case $n=2$

For example, we consider the functional equation

$$f(x+y)+g(x-y)=p_1(x)q_1(y)+p_2(x)q_2(y), \tag{40}$$

where the functions p_1, p_2 and q_1, q_2 are L-independent on the open intervals $]A, B[$ and $]C, D[$, respectively. Because of $\sum_{k=1}^{m} n_k+N=2$, cf. (17), we have only four cases:

$$\text{I. } n_1=n_2=1,\ N=0; \quad \text{II. } n_1=2, \quad N=0;$$
$$\text{III. } n_1=N=1; \quad \text{IV. } N=2.$$

By Theorem 3, the general solutions of (40), where f, g are Lebesgue integrable on $]A, B[\pm]C, D[$ (and (p_1, p_2), (q_1, q_2) L-independent), are

I.
$$f(x)=A_1^{11}e^{w_1x}+A_1^{12}e^{-w_1x}+A_2^{11}e^{w_2x}+A_2^{12}e^{-w_2x}+B,$$
$$g(x)=A_1^{21}e^{w_1x}+A_1^{22}e^{-w_1x}+A_2^{21}e^{w_2x}+A_2^{22}e^{-w_2x}-B,$$
$$p_i(x)=A_{i1}^{11}e^{w_1x}+A_{i1}^{22}e^{-w_1x}+A_{i2}^{11}e^{w_2x}+A_{i2}^{22}e^{-w_2x},$$
$$q_i(y)=A_{i1}^{21}e^{w_1y}+A_{i1}^{12}e^{-w_1y}+A_{i2}^{21}e^{w_2y}+A_{i2}^{12}e^{-w_2y}, \quad (i=1, 2),$$

where

$$A_{1k}^{ll} A_{1s}^{3-l,l} + A_{2k}^{ll} A_{2s}^{3-l,l} = \begin{cases} A_k^{1l}, & s = k, \\ 0, & s \neq k, \end{cases}$$

$$A_{1k}^{ll} A_{1s}^{l,3-l} + A_{2k}^{ll} A_{2s}^{l,3-l} = \begin{cases} A_k^{2l}, & s = k, \\ 0, & s \neq k \end{cases}$$

$$(l = 1, 2;\ s, k = 1, 2);$$

II.

$$f(x) = (A_0^{11} + A_1^{11} x) e^{wx} + (A_0^{12} + A_1^{12} x) e^{-wx} + B,$$

$$g(x) = (A_0^{21} + A_1^{21} x) e^{wx} + (A_0^{22} + A_1^{22} x) e^{-wx} - B,$$

$$p_i(x) = (A_{i0}^{22} + A_{i1}^{11} x) e^{wx} + (A_{i0}^{22} + A_{i1}^{22} x) e^{-wx},$$

$$q_i(y) = (A_{i0}^{21} + A_{i1}^{21} y) e^{wy} + (A_{i0}^{12} + A_{i1}^{12} y) e^{-wy}, \quad (i = 1, 2),$$

where

$$A_{10}^{ll} A_{10}^{3-l,l} + A_{20}^{ll} A_{20}^{3-l,l} = A_0^{1l},$$

$$A_{10}^{ll} A_{10}^{l,3-l} + A_{20}^{ll} A_{20}^{l,3-l} = A_0^{2l},$$

$$A_{1,1-j}^{ll} A_{1j}^{3-l,l} + A_{2,1-j}^{ll} A_{2j}^{3-l,l} = A_1^{1l},$$

$$(-1)^j [A_{1,1-j}^{ll} A_{1j}^{l,3-l} + A_{2,1-j}^{ll} A_{2j}^{l,3-l}] = A_1^{2l},$$

$$A_{11}^{ll} A_{11}^{3-l,l} + A_{21}^{ll} A_{21}^{3-l,l} = A_{11}^{ll} A_{11}^{l,3-l} + A_{21}^{ll} A_{21}^{l,3-l} = 0,$$

$$(l = 1, 2;\ j = 0, 1);$$

III.

$$f(x) = A^{11} e^{wx} + A^{12} e^{-wx} + B_0^1 + B_1^1 x + B_2^1 x^2,$$

$$g(x) = A^{21} e^{wx} + A^{22} e^{-wx} + B_0^2 + B_1^2 x - B_2^1 x^2,$$

$$p_i(x) = A_i^{11} e^{wx} + A_i^{22} e^{-wx} + B_{i0}^1 + B_{i1}^1 x,$$

$$q_i(y) = A_i^{21} e^{wy} + A_i^{12} e^{-wy} + B_{i0}^2 + B_{i1}^2 y, \quad (i = 1, 2),$$

where

$$A_1^{ll} A_1^{3-l,l} + A_2^{ll} A_2^{3-l,l} = A^{1l}, \quad A_1^{ll} A_1^{l,3-l} + A_2^{ll} A_2^{l,3-l} = A^{2l},$$

$$A_1^{ll} B_{ij}^2 + A_2^{ll} B_{2j}^2 = A_1^{3-l,l} B_{1j}^1 + A_2^{3-l,l} B_{2j}^1 = 0,$$

$$B_{10}^1 B_{10}^2 + B_{20}^1 B_{20}^2 = B_0^1 + B_0^2,$$

$$B_{1,1-j}^1 B_{ij}^2 + B_{2,1-j}^1 B_{2j}^2 = B_1^1 + (-1)^j B_1^2,$$

$$B_{11}^1 B_{11}^2 + B_{21}^1 B_{21}^2 = 4 B_2^1,$$

$$(l = 1, 2,;\ j = 0, 1);$$

and

IV.

$$f(x) = B_0^1 + B_1^1 x + B_2^1 x^2 + B_3^1 x^3 + B_4^1 x^4,$$

$$g(x) = B_0^2 + B_1^2 x + B_2^2 x^2 + B_3^2 x^3 - B_4^1 x^4,$$

$$p_i(x) = B_{i0}^1 + B_{i1}^1 x + B_{i2}^1 x^2 + B_{i3}^1 x^3,$$

$$q_i(y) = B_{i0}^2 + B_{i1}^2 y + B_{i2}^2 y^2 + B_{i3}^2 y^3, \quad (i = 1, 2),$$

where

$$B_0^1+B_0^2 = B_{10}^1 B_{10}^2 + B_{20}^1 B_{20}^2,$$

$$\binom{M}{j}[B_M^1+(-1)^j B_M^2] = B_{1,M-j}^1 B_{1j}^2 + B_{2,M-j}^1 B_{2j}^2$$

$$(j = 0, 1, ..., M;\ M = 1, 2, 3),$$

$$8B_4^1 = B_{13}^1 B_{11}^2 + B_{23}^1 B_{21}^2 = B_{11}^1 B_{13}^2 + B_{21}^1 B_{23}^2,$$

$$B_{12}^1 B_{12}^2 + B_{22}^1 B_{22}^2 = B_{12}^1 B_{13}^1 + B_{22}^1 B_{23}^2 =$$

$$= B_{13}^1 B_{12}^2 + B_{23}^1 B_{22}^2 = B_{13}^1 B_{18}^2 + B_{23}^1 B_{23}^2 = 0.$$

REFERENCES

[1] ACZÉL, J., *Lectures on functional equations and their applications*, Mathematics in Science and Engineering, Vol. 19, Academic Press, New York—London, 1966. *MR* **34**#8020.

[2] ACZÉL, J. and DHOMBRES, J., Functional equations containing several variables, *Encyclopedia of Mathematics and its Applications* (to appear).

[3] HAUPT, O., Über einen Eindeutigkeitssatz für gewisse Funktionalgleichungen, *J. Reine Angew. Math.* **186** (1944), 58—64. *MR* **6**—275.

[4] JÁRAI, A., On measurable solutions of functional equations, *Publ. Math. Debrecen* **26** (1979), 17—35. *MR* **80f**: 39008.

[5] SATÔ, R., A study of functional equations, *Proc. Phys. Math. Soc. Japan* (3) **10** (1928), 212—222.

[6] LEVI-CIVITÀ, T., Sulle funzioni che ammettono una formula d'addizione del tipo $f(x+y) = \sum_{i=1}^{n} X_i(x)Y_i(y)$, *Atti Acad. Naz. Lincei Rend.* (5) **22** (1913), Part 2, 181—183.

[7] STÄCKEL, P., Sulla equazione funzionale $f(x+y) = \sum_{i=1}^{n} X_i(x)Y_i(y)$, *Atti Acad. Naz. Lincei Rend.* (5) **22** (1913), Part 2, 392—396.

[8] FENYŐ, I., Über eine Lösungsmethode gewisser Funktionalgleichungen, *Acta Math. Acad. Sci. Hungar.* 7 (1956), 383—396. *MR* **19**—152.

[9] VINCZE, E., Eine allgemeinere Methode in der Theorie der Funktionalgleichungen, I, *Publ. Math. Debrecen* **9** (1962), 149—163; II, *ibid.* **9** (1962), 314—323; III, *ibid.* **10** (1963), 191—202; IV, *ibid.* **10** (1963), 283—318. *MR* **26**#503 (**26**—1544), *MR* **27**#2749, *MR* **29**#3781.

[10] PENNEY, R. C. and RUKHIN, A. L., D'Alembert's functional equation on groups, *Proc. Amer. Math. Soc.* **77** (1979), 73—80. *MR* **80h**: 39013.

[11] RUKHIN, A. L., The solution of the functional equation of D'Alembert's type for commutative groups, *Internat. J. Math. Sci.* **5** (1982), 315—335.

[12] SZÉKELYHIDI, L., Functional equations on abelian groups, *Acta Math. Acad. Sci. Hungar.* **37** (1981), 235—243. *MR* **82f**: 39010.

(Received January 13, 1982)

DEPARTMENT OF PURE MATHEMATICS AND
CENTRE FOR INFORMATION THEORY
UNIVERSITY OF WATERLOO
WATERLOO, ONTARIO
N2L 3G1
CANADA

DEPARTMENT OF MATHEMATICS
SOUTH CHINA INSTITUTE OF TECHNOLOGY
CANTON
PEOPLE'S REPUBLIC OF CHINA

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 69—75

NON-HEREDITARY SEMISIMPLE CLASSES OF NEAR-RINGS

G. BETSCH and R. WIEGANDT

Dedicated to the memory of Professor L. Rédei

1

Much effort has been done to prove the hereditariness of certain concrete near-ring radicals. In this paper, however, we shall show that in many cases the semisimple classes of Kurosh—Amitsur radicals of near-rings are not hereditary. In particular, the semisimple class of a subidempotent radical **R** is never hereditary, provided that **R** contains a 0-symmetric near-ring $N \neq 0$.

The recent results of [7], [8] and [9] gave several characterizations of radical and semisimple classes and of radicals with hereditary semisimple classes for Ω-groups, in particular for 0-symmetric near-rings. The purpose of the present paper is to study the special features of the radicals of near-rings.

One of the classical results of the radical theory is a consequence of the Anderson—Divinsky—Suliński Theorem [1] stating that every semisimple class of associative or alternative rings is hereditary. In [2] Gardner has proved that in the variety of not necessarily associative rings a semisimple class is hereditary if and only if the radical property depends only on the additive group structure. Gardner's result means, in fact, that the usual radicals have never hereditary semisimple classes in that variety. For near-rings Kaarli [5] has proved that the radical $\mathcal{J}_2$ is a Kurosh—Amitsur radical with a hereditary semisimple class. Moreover, by Holcombe and Walker [4] and Holcombe [3] $\mathcal{J}_3$ is a hereditary Kurosh—Amitsur radical for 0-symmetric near-rings with hereditary semisimple class. A recent attempt for studying Kurosh—Amitsur radicals for 0-symmetric near-rings is to be found in [11], claiming that the semisimple class of a radical class need not be hereditary.

Our aim in this note is to prove the non-hereditariness of semisimple classes in many cases. At first we shall construct a near-ring $\Gamma(N)$ for each near-ring N and using this construction we shall prove that certain semisimple classes containing zero-near-rings and satisfying some additional requirements, are not hereditary. Thus, for instance, if **R** is a hereditary radical of abelian near-rings containing a non-constant near-ring N such that every homomorphic image of the zero-near-ring N^0 is **R**-semisimple, then the semisimple class is not hereditary. Further, the semisimple class of a subidempotent radical class containing a 0-symmetric near-ring ($\neq 0$), is never hereditary. Also it turns out that radicals of near-rings with non-hereditary semisimple class are abound.

In what follows we shall work in the variety of all near-rings and we shall adopt the notions and notations of Pilz's book [9]; a near-ring, therefore will always mean a

---

1980 *Mathematics Subject Classification.* Primary 16A76; Secondary 17A65.
Key words and phrases. Kurosh—Amitsur radical and semisimple class, subidempotent radical, lower radical.

right near-ring. For a near-ring N its additive group will be denoted by N^+. Further, N^0 will stand for the *zero-near-ring* built on N^+ by the multiplication $ab=0$ for all $a, b \in N$. The set

$$N_c = \{a \in N\colon\ a0 = a\}$$

is the *constant part* of N. If $N = N_c$, then N is called a *constant near-ring*. The set of all additive commutators of $x \in X \subseteq N$ and $y \in Y \subseteq N$ will be denoted by $[X, Y]$. Further, $I \lhd N$ means that I is an ideal of N.

A class $\mathbf{R}$ is called a *radical class* (or briefly a *radical*) in the sense of Kurosh and Amitsur, if

i) $\mathbf{R}$ is homomorphically closed,

ii) $\mathbf{R}(N) \stackrel{\text{def}}{=} \Sigma(I \lhd N\colon\ I \in \mathbf{R}) \in \mathbf{R}$ for every near-ring N,

iii) $\mathbf{R}$ is closed under extensions, that is, $I \lhd N$, $I \in \mathbf{R}$ and $N/I \in \mathbf{R}$ imply $N \in \mathbf{R}$.

As usual,

$$\mathscr{S}\mathbf{R} = \{N\colon\ \mathbf{R}(N) = 0\}$$

will stay for the *semisimple class* of $\mathbf{R}$. A class $\mathbf{C}$ of near-rings is said to be *hereditary*, if $I \lhd N \in \mathbf{C}$ implies $I \in \mathbf{C}$. As is well-known, a radical $\mathbf{R}$ is hereditary if and only if $\mathbf{R}(I) \supseteq \mathbf{R}(N) \cap I$ for every ideal I of every near-ring N. Moreover, the semisimple class $\mathscr{S}\mathbf{R}$ of a radical $\mathbf{R}$ is hereditary if and only if $\mathbf{R}(I) \subseteq \mathbf{R}(N) \cap I$ (cf. [7] Theorem 1). Pilz [10] calls a radical $\mathbf{R}$ hereditary whenever $\mathbf{R}(I) = \mathbf{R}(N) \cap I$ holds for every ideal I of every near-ring N. Thus Pilz's notion of hereditariness means the hereditariness of both classes $\mathbf{R}$ and $\mathscr{S}\mathbf{R}$. In the sequel we shall use our notion of hereditariness which is in accordance with the usual radical theoretical terminology. For more details concerning the fundamentals of radical theory we refer to [13].

In order to prove the main results of this paper (the Theorems and their Corollaries), we need only the 0-symmetric versions of the construction $\Gamma(N)$, and of the Propositions.

2

For any near-ring N let us consider the cartesian product $\Gamma(N) = N \times N \times N$ and let us define an addition on $\Gamma(N)$ componentwise and a multiplication by the rule

$$(a_,, b_1, c_1)(a_2, b_2, c_2) = (a_1 a_2,\ c_1 c_2 a_2 - c_1 0,\ 0).$$

Obviously, $\Gamma(N)^+$ is a group and a straightforward calculation shows that the multiplication is associative. Looking at the right-distributivity we have

$$((a_1, b_1, c_1) + (a_2, b_2, c_2))(a_3, b_3, c_3) =$$
$$= (a_1 + a_2, b_1 + b_2, c_1 + c_2)(a_3, b_3, c_3) =$$
$$= ((a_1 + a_2)a_3, (c_1 + c_2)c_3 a_3 - (c_1 + c_2)0, 0) =$$
$$= (a_1 a_3 + a_2 a_3, c_1 c_3 a_3 + c_2 c_3 a_3 - c_2 0 - c_1 0, 0),$$

and

$$(a_1, b_1, c_1)(a_3, b_3, c_3) + (a_2, b_2, c_2)(a_3, b_3, c_3) =$$
$$= (a_1 a_3, c_1 c_3 a_3 - c_1 0, 0) + (a_2, a_3\, c_2 c_3 a_3 - c_2 0, 0) =$$
$$= (a_1 a_3 + a_2 a_3, c_1 c_3 a_3 - c_1 0 + c_2 c_3 a_3 - c_2 0, 0).$$

If $[N_c, N^3]=0$, then $-c_1 0+c_2c_3a_3=c_2c_3a_3-c_1 0$, and if $[N_c, N_c]=0$, then $c_2 0+c_1 0=c_1 0+c_2 0$. Consequently, if $[N_c, N_c]=[N_c, N^3]=0$, $\Gamma(N)$ has become a right near-ring. Let us remark that $[N_c, N^3]=[N_c, N_c]=0$ *is satisfied whenever* N *is* 0-*symmetric or abelian.* Furthermore, if N is 0-symmetric or abelian, then so is $\Gamma(N)$, but $\Gamma(N)$ need not be a ring even if N is a ring. Thus *our results are valid in the variety of all near-rings and also in those of* 0-*symmetric near-rings and abelian near-rings.*

PROPOSITION 1. *If N is a near-ring such that* $[N_c, N^3]=[N_c, N_c]=0$, *then* $\Gamma(N)$ *is a near-ring with the following properties:*

(i) $N\cong N_1 \overset{\text{def}}{=} \{(a, 0, 0)\colon a\in N\}$,

(ii) $N\oplus N^0\cong K \overset{\text{def}}{=} \{(a, b, 0)\colon a, b\in N\}$, *and* $N_1\lhd K$,

(iii) $K\lhd\Gamma(N)$,

(iv) $\Gamma(N)/K\cong N^0$.

PROOF. We already proved that $\Gamma(N)$ is a near-ring, meanwhile statements (i)—(iv) are straightforward.

PROPOSITION 2. *Let* $\mathbf{R}$ *be a radical with hereditary semisimple class and let* N *be a near-ring satisfying* $[N_c, N^3]=[N_c, N_c]=0$. *If* $N\in\mathbf{R}$ *and* $N^0\in\mathcal{S}\mathbf{R}$, *then* $\mathbf{R}(\Gamma(N))=N_1$, $xyz=xy0$ *for all* $x, y, z\in N$, *and* $N^3=N^2 0=N_c$. *In particular, if* N *is a* 0-*symmetric near-ring, then* $N\in\mathbf{R}$ *and* $N^0\in\mathcal{S}\mathbf{R}$ *imply* $N^3=0$.

PROOF. Since $\mathcal{S}\mathbf{R}$ is hereditary, we have $\mathbf{R}(K)\subseteq\mathbf{R}(\Gamma(N))$. Applying (iv) of Proposition 1 we get $\Gamma(N)/K\cong N^0\in\mathcal{S}\mathbf{R}$. Hence the intersection representation of the radical yields $\mathbf{R}(\Gamma(N))\subseteq K$, whence $\mathbf{R}(\Gamma(N))\subseteq\mathbf{R}(K)$, implying $\mathbf{R}(\Gamma(N))= =\mathbf{R}(K)$. Taking into account (i) and (ii), it follows

$$\mathbf{R}(\Gamma(N)) = \mathbf{R}(K) = N_1.$$

Thus N_1 is an ideal of $\Gamma(N)$ and hence

$$u(v+i)-uv\in N_1$$

holds for every $u, v\in\Gamma(N)$ and $i\in N_1$. Choosing $u=(0, 0, x)$, $v=(0, 0, y)$ and $i=(z, 0, 0)$ we have

$$(0, 0, x)((0, 0, y)+(z, 0, 0))-(0, 0, x)(0, 0, y) =$$

$$= (0, 0, x)(z, 0, y)-(0, 0, x)(0, 0, y) =$$

$$= (0, xyz-x0, 0)-(0, xy0-x0, 0) =$$

$$= (0, xyz-x0-(xy0-x0), 0) = (0, xyz-xy0, 0).$$

Since $N_1=\mathbf{R}(\Gamma(N))\lhd\Gamma(N)$, it follows $xyz=xy0$ for all $x, y, z\in N$. Thus $N^3= =N^2 0\subseteq N_c$ holds.

The inclusion $N_c\subseteq N^2 0$ is trivial.

A straightforward application of Proposition 2 yields

PROPOSITION 3. *Let* $\mathbf{R}$ *be a radical with hereditary semisimple class* $\mathcal{S}\mathbf{R}$ *and let* N *be a near-ring such that* $[N_c, N^3]=[N_c, N_c]=0$. *If* $\mathbf{R}(N)^0\in\mathcal{S}\mathbf{R}$, *then* $\mathbf{R}(N)^3=$

$=\mathbf{R}(N)^2 0=\mathbf{R}(N)_c$. *In particular, if N is a 0-symmetric near-ring and* $\mathbf{R}(N)^0\in\mathcal{S}\mathbf{R}$, *then* $\mathbf{R}(N)^3=0$.

Let

$$(0:N) = \{a\in N\colon\ aN = 0\}$$

denote the annihilator of N in N which is an ideal of N in view of [10] 1.43 Corollary.

PROPOSITION 4. *Let* $\mathbf{R}$ *be a radical with hereditary semisimple class* $\mathcal{S}\mathbf{R}$ *and* N *a near-ring such that* $[N_c, N^3]=[N_c, N_c]=0$ *and* N_c^+ *is a normal subgroup of* N^+. *If* $N\in\mathbf{R}$, $(0:N)\in\mathbf{R}$, $N^0/N_c^0\in\mathcal{S}\mathbf{R}$, *then* $N=N_c$.

PROOF. By Proposition 2 we have $xy-xy0\in(0:N)$ for all $x, y\in N$. Moreover, $(0:N)^2=0$ implies that $(0:N)$ is an ideal of N^0 satisfying $(0:N)\cap N_c=0$, whence $(0:N)\lhd N^0/N_c^0\in\mathcal{S}\mathbf{R}$. The hereditariness of $\mathcal{S}\mathbf{R}$ now implies $(0:N)\in\mathcal{S}\mathbf{R}$. Hence it follows $(0:N)=0$ yielding $xy=xy0$ for every $x, y\in N$, that is $N^2=N_c$. According to [10] Proposition 1.32(b), $N^2=N_c$ is an invariant subnear-ring of N. Since N_c^+ is normal in N^+, and $a(b+i)-ab\in N^2=N_c$ for all $a, b, i\in N$, we conclude that N_c is an ideal of N. Thus by $N^2=N_c$, $N\in\mathbf{R}$ and by the assumption on N^0 it follows

$$N/N_c = N/N^2 = (N/N^2)^0 = N^0/(N^2)^0\in\mathbf{R}\cap\mathcal{S}\mathbf{R} = 0$$

which implies $N=N_c$.

An immediate application of Proposition 4 yields the following

COROLLARY 1. *Let* $\mathbf{R}$ *be a hereditary radical class. If* $\mathbf{R}$ *contains a non-constant abelian near-ring* N *such that every homomorphic image of* N^0 *is in* $\mathcal{S}\mathbf{R}$, *then* $\mathcal{S}\mathbf{R}$ *is not hereditary.*

According to the previous notation, $(0\colon \mathbf{R}(N))$ will stand for the annihilator of $\mathbf{R}(N)$ in $\mathbf{R}(N)$.

PROPOSITION 5. *Let* $\mathbf{R}$ *be a radical with hereditary semisimple class and* N *a 0-symmetric near-ring with the property* $(0\colon \mathbf{R}(N))\in\mathbf{R}$. *If* $\mathbf{R}(N)^0\in\mathcal{S}\mathbf{R}$, *then* $\mathbf{R}(N)=0$. *In particular,* $N^0\in\mathcal{S}\mathbf{R}$ *implies* $N\in\mathcal{S}\mathbf{R}$.

PROOF. The assertion is, in fact, a special case of that of Proposition 4: the near-ring $R(N)$ satisfies all the assumptions of Proposition 4, so necessarily $R(N)=$ $=R(N)_c=0$, as N is 0-symmetric. Hence $N\in\mathcal{S}\mathbf{R}$.

A reformulation of Proposition 5 is

THEOREM 1. *Let* $\mathbf{R}$ *be a radical. If there exists a 0-symmetric near-ring* N *such that* $\mathbf{R}(N)\neq 0$, $(0\colon \mathbf{R}(N))\in\mathbf{R}$ *and* $N^0\in\mathcal{S}\mathbf{R}$, *then the semisimple class is not hereditary.*

COROLLARY 2. *Let* $\mathbf{R}$ *be a hereditary radical. If there exists a 0-symmetric near-ring* N *such that* $\mathbf{R}(N)\neq 0$ *and* $N^0\in\mathcal{S}\mathbf{R}$, *then the semisimple class* $\mathcal{S}\mathbf{R}$ *is not hereditary.*

PROPOSITION 6. *Let* $\mathbf{R}$ *be a radical with hereditary semisimple class* $\mathcal{S}\mathbf{R}$. *If* N *is a 0-symmetric near-ring with* $N^0\in\mathcal{S}\mathbf{R}$, *then the annihilator* $(0\colon \mathbf{R}(N))$ *of* $\mathbf{R}(N)$ *in* $\mathbf{R}(N)$ *is an ideal of* N^0. *If in addition,* $N^0/(0\colon \mathbf{R}(N))\in\mathcal{S}\mathbf{R}$, *then* $N\in\mathcal{S}\mathbf{R}$.

PROOF. $(0\colon \mathbf{R}(N))$ is an ideal of $\mathbf{R}(N)$, further, $(0\colon \mathbf{R}(N))$ is a zero-near-ring whose additive group is a normal subgroup of N^+ because for any $a\in(0\colon R(N))$, $n\in N$ and $r\in R(N)$ we have $(n+a-n)r=nr+ar-nr=0$ and hence $n+a-n\in(0\colon R(N))$. Therefore $(0\colon \mathbf{R}(N))\lhd N^0\in\mathcal{S}\mathbf{R}$ holds. Hence the hereditariness of $\mathcal{S}\mathbf{R}$ yields $(0\colon \mathbf{R}(N))\in\mathcal{S}\mathbf{R}$. Since N is 0-symmetric, by Proposition 3 $\mathbf{R}(N)$ satisfies $\mathbf{R}(N)^3=0$ and hence $\mathbf{R}(N)^2\subseteq(0\colon \mathbf{R}(N))$ is valid. Consequently, in view of $N^0\in\mathcal{S}\mathbf{R}$ we have

$$\mathbf{R}(N)/(0\colon \mathbf{R}(N)) \cong \mathbf{R}(N)^0/(0\colon \mathbf{R}(N))\lhd N^0/(0\colon \mathbf{R}(N))\in\mathcal{S}\mathbf{R}.$$

Thus we obtained

$$\mathbf{R}(N)/(0\colon \mathbf{R}(N))\in\mathbf{R}\cap\mathcal{S}\mathbf{R} = 0,$$

that is, $\mathbf{R}(N)=(0\colon \mathbf{R}(N))\lhd N^0\in\mathcal{S}\mathbf{R}$ holds, implying $\mathbf{R}(N)=0$.

THEOREM 2. *Let* $\mathbf{R}$ *be a radical which does not contain zero-near-rings* $(\neq 0)$. *If* $\mathbf{R}$ *contains a 0-symmetric near-ring* $N\neq 0$, *then the semisimple class* $\mathcal{S}\mathbf{R}$ *is not hereditary.*

PROOF. Obviously, every zero-near-ring is in $\mathcal{S}\mathbf{R}$. Since $0\neq N\in\mathbf{R}$ and $N^0\in\mathcal{S}\mathbf{R}$, by Proposition 6 $\mathcal{S}\mathbf{R}$ is not hereditary.

Let $\mathbf{Z}$ denote the class of all zero-near-rings, and let us consider the upper radical class

$$\mathcal{U}\mathbf{Z}=\{N\colon N \text{ has no nonzero homomorphic image in } \mathbf{Z}\}.$$

Since $\mathbf{Z}$ is hereditary, $\mathcal{U}\mathbf{Z}$ is a Kurosh—Amitsur radical class, moreover $\mathcal{U}\mathbf{Z}$ is the largest radical such that $\mathbf{Z}\cap\mathcal{U}\mathbf{Z}=0$. In what follows (N^2) will denote the ideal of N generated by N^2.

PROPOSITION 7. $N\in\mathcal{U}\mathbf{Z}$ *if and only if* $N=(N^2)$.

PROOF. If $N\notin\mathcal{U}\mathbf{Z}$, then there is an ideal I of N such that $0\neq N/I\in\mathbf{Z}$, that is, $(N^2)\subseteq I$. Hence $(N^2)\subseteq I\neq N$ holds.

If $N\in\mathcal{U}\mathbf{Z}$, then $N/(N^2)\in\mathbf{Z}\cap\mathcal{U}\mathbf{Z}=0$. Hence $N=(N^2)$ holds.

In view of Proposition 7 a radical class $\mathbf{R}$ will be called *subidempotent,* if $\mathbf{R}\subseteq\mathcal{U}\mathbf{Z}$. Theorem 2 and Proposition 7 immediately yield

COROLLARY 3. *If a subidempotent radical* $\mathbf{R}$ *contains a 0-symmetric near-ring* $N\neq 0$, *then the semisimple class* $\mathcal{S}\mathbf{R}$ *is not hereditary.*

As we have seen in proving the non-hereditariness of certain semisimple classes, the construction $\Gamma(N)$ played a decisive role. In fact, using $\Gamma(N)$ we could prove statements of the following type if $\mathcal{S}\mathbf{R}$ is a hereditary semisimple class, then the assumption $N^0\in\mathcal{S}\mathbf{R}$ attracts also other near-rings into the class $\mathcal{S}\mathbf{R}$. Thus in all of our non-hereditary results the semisimple class $\mathcal{S}\mathbf{R}$ contained a zero-near-ring $N^0\neq 0$. As far as radical classes $\mathbf{R}$ containing the class $\mathbf{Z}$ are concerned, they may have hereditary semisimple classes. Kaarli [5] has proved that $\mathcal{J}_2$ is a hereditary Kurosh—Amitsur radical with hereditary semisimple class, and also $\mathbf{Z}\subseteq\mathcal{J}_2$ holds. Moreover, for 0-symmetric near-rings Holcombe [3] has introduced the radical $\mathcal{J}_3$ containing $\mathcal{J}_2$, and Holcombe and Walker [4] have shown that $\mathcal{J}_3$ is a Kurosh—Amitsur radical.

According to Holcombe [3] $\mathcal{J}_3$ is hereditary and has a hereditary semisimple class. (For the definitions of $\mathcal{J}_2$ and $\mathcal{J}_3$ we refer to [10].) It is worth mentioning that both $\mathcal{J}_2$ and $\mathcal{J}_3$ satisfy the following condition: If I is an invariant subnear-ring of N and $I\in\mathcal{J}_i$, then $I\subseteq\mathcal{J}_i(N)$ holds for $i=2, 3$. (In Kaarli's terminology an invariant subnear-ring is called a quasi-ideal.) In a 0-symmetric near-ring every ideal is an invariant subnear-ring, see [10] Proposition 1.34 (b).

A radical containing all zero-near-rings, may have a non hereditary semisimple class. For instance, the class $\mathcal{N}$ of all nil near-rings is a hereditary radical class containing all zero-near-rings. But as Kaarli has kindly informed us, it follows from his Example 5.4 of [6] that the semisimple class of $\mathcal{N}$ is not hereditary.

3

Concerning radicals with non-hereditary semisimple classes one may ask whether there are sufficiently many such classes. Making use of the lower radical construction given by Tangeman and Kreiling [12] we give an affirmative answer to this question. Though in [12] not necessarily associative rings were considered, the results are valid also for near-rings. Let $\mathbf{C}$ be any class of near-rings, and define $\mathbf{C}_1$ as its homomorphic closure. Furthermore,

$$\mathbf{C}_\beta = \{N : I,\ N/I\in\mathbf{C}_{\beta-1} \text{ for some } I\lhd N\}$$

if $\beta-1$ exists, and

$$\mathbf{C}_\beta=\{N:\ N \text{ contains a chain } \{I_\gamma\} \text{ of ideals such that } I_\gamma\in\bigcup_{\alpha<\beta}\mathbf{C}_\alpha \text{ and } N=\bigcup I_\gamma\}$$

if $\beta-1$ does not exist. Now $\mathcal{L}\mathbf{C}=\bigcup\mathbf{C}_\beta$ is the lower radical determined by the class $\mathbf{C}$ ([12] Theorem 2). Moreover, if $\mathbf{C}$ is also hereditary, then so is $\mathcal{L}\mathbf{C}$ ([12] Theorem 3). Obviously, one can easily construct plenty of radicals having non-hereditary semisimple classes. For instance, let $\mathbf{D}$ be any homomorphically closed (and hereditary) class of near-rings such that $\mathbf{D}$ does not contain zero-near-rings ($\neq 0$), but does contain a 0-symmetric near-ring $N\neq 0$. Then $\mathcal{L}\mathbf{D}$ is a (hereditary) radical class such that the semisimple class $\mathcal{S}\mathcal{L}\mathbf{D}$ is *not* hereditary, as stated in Theorem 2.

We summarize the hereditary properties of some radicals in the variety of 0-symmetric near-rings.

radical	radical class	semisimple class
$\mathcal{J}_2$	hereditary	hereditary
$\mathcal{J}_3$	hereditary	hereditary
$\mathcal{N}$	hereditary	non-hereditary
subidempotent radicals ($\neq 0$)	hereditary or not	non-hereditary

Acknowledgement. The authors are indebted to Professor R. Mlitz for his valuable advices.

REFERENCES

[1] Anderson, T., Divinsky, N. and Suliński A., Hereditary radicals in associative and alternative rings, *Canad. J. Math.* **17** (1965), 594—603. *MR* **31** # 215.

[2] Gardner, B. J., Some degeneracy and pathology in non-associative radical theory, *Ann. Univ. Sci. Budapest Eötvös Sect. Math.* **22/23** (1979/80), 65—74. *MR* **81j**: 17005.

[3] Holcombe, M., A hereditary radical for near-rings, *Studia Sci. Math. Hungar.* **17** (1982), 453—456.

[4] Holcombe M. and Walker, R., Radicals in categories, *Proc. Edinburgh Math. Soc.* (2) **21** (1978/79), 111—128. *MR* **80b**: 18009.

[5] Kaarli, K., Радикалы в почти-кольцах, (Radicals in near-rings, in Russian), *Tartu Riikl. Ül. Toimetised* **390** (1976), 134—171. *MR* **57** # 9760.

[6] Kaarli K., Классификация неприводимых модулей над полупримарным почти-кольцом, *Tartu Riikl. Ül. Toimetised* No. **556** *Trudy Mat. Meh.* (1981), 47—63. *MR* **82k**: 16047.

[7] Leavitt, W. G. and Wiegandt, R., Torsion theory for not necessarily associative rings, *Rocky Mountain J. Math.* **9** (1979), 259—271. *MR* **80d**: 17005.

[8] Leeuwen, L. C. A. van and Wiegandt, R., Radicals, semisimple classes and torsion theories, *Acta Math. Acad. Sci. Hungar.* **36** (1980), 37—37. *MR* **82j**: 17004.

[9] Mlitz, R., Radicals and semisimple classes of Ω-groups, *Proc. Edinburgh Math. Soc.* (2) **23** (1980), 37—41. *MR* **82e**: 17005.

[10] Pilz, G., *Near-rings.* The theory and its applications, North-Holland Mathematics Studies, No. 23, North-Holland Publishing Co., Amsterdam—New York—Oxford, 1977. *MR* **57** # 9761.

[11] Saxena, P. K. and Bhandari, M. C., General radical theory for near-rings, *Tamkang J. Math.* **12** (1981), 91—97.

[12] Tangeman, R. and Kreiling, D., Lower radicals in nonassociative rings, *J. Austral. Math. Soc.* **14** (1972), 419—423. *MR* **47** # 3470.

[13] 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 January 27, 1982)

MATHEMATISCHES INSTITUT
UNIVERSITÄT TÜBINGEN
AUF DER MORGENSTELLE 10
D—7400 TÜBINGEN
FEDERAL REPUBLIC OF GERMANY

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

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 77—90

SIMPLE EIGENVALUES OF TRANSITIVE GRAPHS

HORST SACHS and MICHAEL STIEBITZ

Dedicated to the memory of László Rédei

Abstract

For the classes of directed graphs, undirected graphs, multigraphs, and schlicht graphs, all finite and vertex transitive, using representation theoretical and combinatorial means upper bounds for the number of simple eigenvalues are found.

Introduction

A transitive graph **G** is a graph whose automorphism group Γ acts transitively on its vertex set **V**. The eigenvalues of **G** are the eigenvalues of its adjacency matrix **A**. Clearly, a permutation γ acting on **V** is an automorphism of **G** if and only if the corresponding permutation matrix $\mathbf{P}_\gamma$ satisfies $\mathbf{P}_\gamma^{-1}\mathbf{A}\mathbf{P}_\gamma=\mathbf{A}$ or, equivalently, $\mathbf{A}\mathbf{P}_\gamma=\mathbf{P}_\gamma\mathbf{A}$. Thus, if $\gamma\in\Gamma$ and **x** is an eigenvector of **A** (i.e., $\mathbf{A}\mathbf{x}=\lambda\mathbf{x}$), then so is $\mathbf{P}_\gamma\mathbf{x}$ belonging to the same eigenvalue λ, and if λ is a simple eigenvalue then **x** and $\mathbf{P}_\gamma\mathbf{x}$ are linearly dependent for every $\gamma\in\Gamma$. Therefore, it is reasonable to ask under what conditions a transitive graph can have (non-trivial) simple eigenvalues at all, and how many it can have.

This problem was already investigated by the authors in a previous paper [5], however, since that paper was written, many extensions and improvements have been found. Therefore, in this paper the authors present a condensed theory systematically using representation theoretical means. In order to make the paper self-contained, they have included in it some of the proofs (improved versions) already given in [5].

The authors wish to thank Professor Gerhard Pazderski at the Wilhelm-Pieck-Universität Rostock for many helpful discussions.

1. Preliminaries

In what follows $\mathbf{G}=(\mathbf{V}, \mathbf{E})$ will always denote a finite, directed or undirected, graph with vertex set **V** and edge set **E**, loops and multiple edges being allowed. **G** is called *schlicht* if it has neither loops nor multiple edges.

Let $|\mathbf{V}|=n$: then the vertices of **G** can be labelled by $1, 2, \ldots, n$ and identified with their labels.

Let a_{ij} denote the number of edges going from vertex i to vertex j, where, for an undirected graph **G**, a_{ii} is defined to be twice the number of loops attached to the vertex i. The square matrix $\mathbf{A}=(a_{ij})$ of order n is called the *adjacency matrix* of **G**. The graph **G** is called *regular of degree r* iff all row sums and all column sums of **A** are equal to r. The eigenvectors (eigenvalues) of $\mathbf{G}$ are the eigenvectors (eigenvalues) of its adjacency matrix $\mathbf{A}$.

For $\mathbf{V}_1, \mathbf{V}_2 \subseteq \mathbf{V}$, let $W^+_{\mathbf{G}}(\mathbf{V}_1, \mathbf{V}_2) = W^-_{\mathbf{G}}(\mathbf{V}_2, \mathbf{V}_1)$ denote the number of edges issuing from $\mathbf{V}_1$ and terminating in $\mathbf{V}_2$, i.e.,

$$W^+_{\mathbf{G}}(\mathbf{V}_1, \mathbf{V}_2) = W^-_{\mathbf{G}}(\mathbf{V}_2, \mathbf{V}_1) = \sum_{\substack{i \in \mathbf{V}_1 \\ j \in \mathbf{V}_2}} a_{ij}.$$

If $\mathbf{V}_1 = \{a\}$ $(a \in \mathbf{V})$, we shall briefly write $W^+_{\mathbf{G}}(a, \mathbf{V}_2)$ instead of $W^+_{\mathbf{G}}(\{a\}, \mathbf{V}_2)$, etc.

Let $\mathbf{S}_n$ denote the symmetric group of degree n. The group of all adjacency preserving (1,1)-mappings of $\mathbf{V} = \{1, \dots, n\}$ onto itself is called the *automorphism group* $\mathbf{Aut}(\mathbf{G})$ of $\mathbf{G}$:

$$\gamma \in \mathbf{Aut}(\mathbf{G}) \quad \text{iff} \quad \gamma \in \mathbf{S}_n \quad \text{and} \quad a_{\gamma(i)\gamma(j)} = a_{ij}, \quad i, j = 1, \dots, n.$$

Let $\mathbf{P}_\gamma$ be the permutation matrix assigned to the permutation $\gamma \in \mathbf{S}_n$: then $\gamma \in \mathbf{Aut}(\mathbf{G})$ iff $\mathbf{A}\mathbf{P}_\gamma = \mathbf{P}_\gamma \mathbf{A}$.

$\mathbf{G}$ is called *transitive* iff $\mathbf{Aut}(\mathbf{G})$ acts transitively on $\mathbf{V}$.

Let $\mathscr{C}$ denote the complex number field and $\mathbf{H}$ be an abstract group; recall that a homomorphism of $\mathbf{H}$ into the multiplicative group of $\mathscr{C}$ (or the set of images under this homomorphism) is called a *representation of* $\mathbf{H}$ *of degree* 1.

Let $\mathbf{x} = (x_1, \dots, x_n)^{\mathrm{T}}$ and $\mathbf{y} = (y_1, \dots, y_n)^{\mathrm{T}}$ be vectors over $\mathscr{C}$. The Hermitian scalar product $x_1\bar{y}_1 + \dots + x_n\bar{y}_n$ of $\mathbf{x}$ and $\mathbf{y}$ is denoted by $\langle \mathbf{x}, \mathbf{y} \rangle$; $\mathbf{x}, \mathbf{y}$ are called *orthogonal* iff $\langle \mathbf{x}, \mathbf{y} \rangle = 0$. The vector $\mathbf{u} = (1, \dots, 1)^{\mathrm{T}}$ is called the *identity vector*.

For positive integers l and n, put

$$\mathbf{U}_l(n) = \{\mathbf{x} = (x_1, \dots, x_n)^{\mathrm{T}} \mid x_i^l = 1 \text{ for } i = 1, \dots, n\}.$$

Under the composition $\otimes$ defined by

$$\mathbf{x} \otimes \mathbf{y} = (x_1 y_1, \dots, x_n y_n)^{\mathrm{T}},$$

$\mathbf{U}_l(n)$ becomes an abelian group, namely, the direct product of n cyclic groups of order l.

Let $\mathbf{x}$ be an eigenvector of the graph $\mathbf{G}$ with corresponding eigenvalue λ, i.e., $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$. Then, for all $\gamma \in \mathbf{Aut}(\mathbf{G})$,

$$\mathbf{A}\mathbf{P}_\gamma \mathbf{x} = \mathbf{P}_\gamma \mathbf{A}\mathbf{x} = \mathbf{P}_\gamma \lambda \mathbf{x} = \lambda \mathbf{P}_\gamma \mathbf{x}.$$

This means that, together with $\mathbf{x}$, also $\mathbf{P}_\gamma\mathbf{x}$ is an eigenvector of $\mathbf{G}$ belonging to the same eigenvalue λ. If λ is a simple eigenvalue of $\mathbf{A}$, then $\mathbf{x}$ and $\mathbf{P}_\gamma\mathbf{x}$ must be linearly dependent. This implies that, for each $\gamma \in \mathbf{Aut}(\mathbf{G})$, there is a complex number $a = a_\gamma(\mathbf{x})$ satisfying $\mathbf{P}_\gamma\mathbf{x} = a_\gamma(\mathbf{x}) \cdot \mathbf{x}$; thus $\mathbf{x}$ is a common eigenvector of all $\mathbf{P}_\gamma$, too.

The number of simple eigenvalues of $\mathbf{G}$ will be denoted by $\sigma(G)$.

Let $\Gamma \subseteq \mathbf{S}_n$ be a transitive permutation group of degree n and order g and let $\mathbf{x} = (x_1, \dots, x_n)^{\mathrm{T}} \neq 0$ satisfy

$$\mathbf{P}_\gamma \mathbf{x} = a_\gamma(\mathbf{x}) \cdot \mathbf{x} \quad \text{for all} \quad \gamma \in \Gamma$$

where the $a_\gamma(\mathbf{x})$ are complex numbers. Then $\{a_\gamma(\mathbf{x}) \mid \gamma \in \Gamma\}$ is a representation of Γ of degree 1, the numbers $a_\gamma(\mathbf{x})$ are g-th roots of unity, and the transitivity of Γ immediately implies $|x_1| = \dots = |x_n| \neq 0$.

Put
$\mathbf{Z}(\Gamma)=\{\mathbf{x}|x_1=1$ *and for each* $\gamma\in\Gamma$ *there is a complex number* $a_\gamma(\mathbf{x})$ *such that* $\mathbf{P}_\gamma\mathbf{x}=a_\gamma(\mathbf{x})\cdot\mathbf{x}\}$.

$\mathbf{Z}(\Gamma)$ has the following properties (see Section 2.2, **(D)** and **(F)**):

(i) *For* $l=\exp\Gamma$, $\mathbf{Z}(\Gamma)$ *is a subgroup of* $\mathbf{U}_l(n)$,

(ii) $\langle\mathbf{x},\mathbf{y}\rangle=0$ *for* $\mathbf{x},\mathbf{y}\in\mathbf{Z}(\Gamma)$, $\mathbf{x}\neq\mathbf{y}$.

2. Eigenvectors of transitive graphs

2.1. *Feasible vector sets*

Definition 1. A subgroup $\mathbf{E}$ of $\mathbf{U}_l(n)$ is called a *feasible vector set of degree l and dimension n* iff any two distinct elements of $\mathbf{E}$ are orthogonal.

Denote the set of all feasible vector sets of degree l and dimension n by $\mathscr{E}_l(n)$ and put

$$s_l(n)=\max_{\mathbf{E}\in\mathscr{E}_l(n)}|\mathbf{E}|.$$

Lemma 1 (see, e.g., [3]). *Let* $\mathbf{H}=\{h_1=\mathrm{id}, h_2, \ldots, h_m\}$ *be an abelian group of order m and let* $\chi_1, \ldots, \chi_m$ *be the characters of the m irreducible representations of* $\mathbf{H}$ *(in what follows these will be briefly called "the simple characters" of* $\mathbf{H}$*;* $\mathbf{H}$ *being abelian,* $\chi_1, \ldots, \chi_m$ *correspond to the m distinct representations of degree 1 of* $\mathbf{H}$*). Put* $\mathbf{M}=(\chi_i(h_j))$, $i, j=1, \ldots, m$. *Then*

(a) *the row vectors (column vectors) of* $\mathbf{M}$ *are pairwise orthogonal,*

(b) $\det(\mathbf{M})\neq0$ *and* $\mathbf{M}^{-1}=\frac{1}{m}\overline{\mathbf{M}}^{\mathrm{T}}$. ∎

With respect to the usual composition, the characters of an abelian group $\mathbf{H}$ form a group isomorphic to $\mathbf{H}$. Thus the row vectors as well as the column vectors of $\mathbf{M}$ form feasible vector sets of degree $l=\exp\mathbf{H}$ and dimension m. In particular, putting

$$\chi(\mathbf{H})=\{\mathbf{x}_i=(\chi_1(h_i), \ldots, \chi_m(h_i))^{\mathrm{T}}\,|\,i=1, \ldots, m\},$$

we obtain

$\chi(\mathbf{H})\in\mathscr{E}_l(m)$ where $l=\exp\mathbf{H}$.

Further, put

$$\chi_k(\mathbf{H})=\{\mathbf{x}_i^{(k)}=(\underbrace{\chi_1(h_i), \ldots, \chi_1(h_i)}_{k\ times}; \ldots; \chi_m(h_i), \ldots, \chi_m(h_i))^{\mathrm{T}}\,|\,i=1, \ldots, m\}.$$

Then, clearly, $\chi_k(\mathbf{H})\in\mathscr{E}_l(k\cdot m)$ and $\chi_k(\mathbf{H})\cong\chi_1(\mathbf{H})=\chi(\mathbf{H})\cong\mathbf{H}$.

Lemma 2. *Let* $\mathbf{E}\in\mathscr{E}_l(n)$ *be a feasible vector set of order m. Then*

(a) $m|n$,

(b) *the components of the vectors of* $\mathbf{E}$ *can be so labelled that* $\mathbf{E}=\chi_k(\mathbf{E})$, *where* $k=\frac{n}{m}$.

PROOF. The mappings ψ_i defined by $\psi_i(\mathbf{x})=x_i$ for $\mathbf{x}=(x_1, \ldots, x_n)^{\mathrm{T}}\in\mathbf{E}$ are simple characters of the abelian group $\mathbf{E}$; further, $\mathbf{x}=(\psi_1(\mathbf{x}), \ldots, \psi_n(\mathbf{x}))^{\mathrm{T}}$.

Let $\chi_1, \ldots, \chi_m$ be the m simple characters of $\mathbf{E}$. Then, for $\mathbf{x}\in\mathbf{E}$,

$$\langle \mathbf{x}, \mathbf{u}\rangle = \sum_{i=1}^{n} \psi_i(\mathbf{x}) = \sum_{j=1}^{m} k_j\chi_j(\mathbf{x}) = \begin{cases} 0 & \text{for} \quad \mathbf{x} \neq \mathbf{u} \\ n & \text{for} \quad \mathbf{x} = \mathbf{u}. \end{cases}$$

Lemma 1 now immediately implies $k_1=\ldots=k_m=k=\dfrac{n}{m}$ yielding the assertions. ∎

The following proposition is a simple consequence of Lemma 2.

PROPOSITION 1. *Let* $\mathbf{E}, \mathbf{E}'\in\mathscr{E}_l(n)$, $\mathbf{E}\cong\mathbf{E}'$, *and let* φ *be an isomorphism mapping* $\mathbf{E}$ *onto* $\mathbf{E}'$; *then there is a permutation* $\gamma\in\mathbf{S}_n$ *satisfying* $\varphi(\mathbf{x})=\mathbf{P}_\gamma\mathbf{x}$ *for all* $\mathbf{x}\in\mathbf{E}$.

LEMMA 3. *Let* $\mathbf{x}$ *be an element of order* d *of* $\mathbf{E}\in\mathscr{E}_l(n)$. *Then* $d|l$ *and* $d|n$.

PROOF. $\mathbf{x}^l=\mathbf{u}$ implies $d|l$ and, because of $d|m$, by Lemma 2 also $d|n$. ∎

THEOREM 1. *Let* $l=p_1^{\alpha_1}\ldots p_r^{\alpha_r}$ *and* $n=p_1^{k_1}\ldots p_r^{k_r}k$ *where* $p_1, \ldots, p_r$ *are distinct primes,* $r, \alpha_1, \ldots, \alpha_r$, *and* k *are positive integers,* $k_1, \ldots, k_r$ *are non-negative integers, and* $(k, l)=1$. *Then*

$$s_l(n) = p_1^{k_1}\ldots p_r^{k_r}.$$

PROOF. As an immediate consequence of Lemma 3, $s_l(n)|p_1^{k_1}\ldots p_r^{k_r}$.

Let $\mathbf{H}=\mathbf{H}_1\times\ldots\times\mathbf{H}_r$ where $\mathbf{H}_i=\mathbf{H}_i^{(1)}\times\ldots\times\mathbf{H}_i^{(k_i)}$ $(i=1, \ldots, r)$ is the direct product of k_i cyclic groups each of order p_i. Then $\mathbf{H}$ is abelian of order $|\mathbf{H}|=p_1^{k_1}\ldots p_r^{k_r}$. Put $l'=\exp\mathbf{H}$: clearly, l' is the product of all those p_i which have $k_i>0$, thus $l'|l$. Now, $\mathbf{E}=\chi_k(\mathbf{H})\in\mathscr{E}_{l'}(n)\subseteq\mathscr{E}_l(n)$ is a feasible vector set of order $|\mathbf{E}|=|\mathbf{H}|=p_1^{k_1}\ldots p_r^{k_r}$ from which the assertion immediately follows. ∎

With respect to $\mathbf{E}\in\mathscr{E}_l(n)$, define a relation $i\sim j|\mathbf{E}$ $(i, j\in\mathbf{V}=\{1, \ldots, n\})$ as follows:

$$i \sim j|\mathbf{E} \quad \text{iff} \quad x_i = x_j \quad \text{for all} \quad \mathbf{x}\in\mathbf{E}.$$

Clearly, this relation is an equivalence relation (which we shall call an $\mathbf{E}$-*equivalence*) partitioning $\mathbf{V}=\{1, \ldots, n\}$ into equivalence classes $\mathbf{X}_1(\mathbf{E}), \ldots, \mathbf{X}_a(\mathbf{E})$, say. From the proof of Lemma 2, $a=m=|\mathbf{E}|$ and $|\mathbf{X}_i(\mathbf{E})|=\dfrac{n}{m}$ $(i=1, \ldots, m)$ immediately follows.

Next, put $\bar{\mathbf{E}}=\chi_1(\mathbf{E})$; clearly, $\bar{\mathbf{E}}\in\mathscr{E}_l(m)$ and $\bar{\mathbf{E}}$ is isomorphic to $\mathbf{E}$. To each $\mathbf{E}\in\mathscr{E}_l(n)$ assign a permutation group $\Gamma(\mathbf{E})$ where $\gamma\in\Gamma(\mathbf{E})$ iff $\gamma\in\mathbf{S}_n$ and to each $\mathbf{x}\in\mathbf{E}$ there is a complex number $a_\gamma(\mathbf{x})$ such that $P_\gamma\mathbf{x}=a_\gamma(\mathbf{x})\cdot\mathbf{x}$.

LEMMA 4. *Let* $\mathbf{E}\in\mathscr{E}_l(n)$, $\bar{\mathbf{E}}=\chi_1(\mathbf{E})$. *Then* $\Gamma(\bar{\mathbf{E}})$ *is a transitive abelian permutation group isomorphic to* $\mathbf{E}$.

PROOF. Let $\chi_1, \ldots, \chi_m$ $(m=|\mathbf{E}|)$ be the simple characters of $\mathbf{E}$. Then, for all $\mathbf{x}\in\mathbf{E}$, $\bar{\mathbf{x}}=(\chi_1(\mathbf{x}), \ldots, \chi_m(\mathbf{x}))^{\mathrm{T}}\in\bar{\mathbf{E}}$ and for $\gamma\in\Gamma(\bar{\mathbf{E}})$ we obtain $\mathbf{P}_\gamma\bar{\mathbf{x}}=a_\gamma(\mathbf{x})\cdot\bar{\mathbf{x}}$, i.e., $\chi_{\gamma(i)}(\mathbf{x})=a_\gamma(\mathbf{x})\chi_i(\mathbf{x})$, for all $\mathbf{x}\in\mathbf{E}$ and $i\in\{1, \ldots, m\}$: this means that the permutations of $\Gamma(\bar{\mathbf{E}})$ are in (1,1)-correspondence with the simple characters of $\mathbf{E}$, thus

$\Gamma(\bar{\mathbf{E}})$ is a transitive abelian permutation group of order m being isomorphic to $\mathbf{E}$. ∎

THEOREM 2. *Let* $\mathbf{E}\in\mathscr{E}_l(n)$, $|\mathbf{E}|=m$. *By Lemma 2,* $m|n$; *put* $n=km$. *Then* $\Gamma(\mathbf{E})$ *is a transitive permutation group of order* $m(k!)^m$ *isomorphic to the wreath product of* $\Gamma(\bar{\mathbf{E}})$ *and* $\mathbf{S}_k$.

PROOF. By $\Gamma(\mathbf{E})$ the equivalence classes $\mathbf{X}_1(\mathbf{E}), \ldots, \mathbf{X}_m(\mathbf{E})$ are permuted according to $\Gamma(\bar{\mathbf{E}})$; in particular, with respect to $\Gamma(\mathbf{E})$ they are domains of imprimitivity. For each $i\in\{1, \ldots, m\}$, the restriction of $\Gamma(\mathbf{E})$ to $X_i(\mathbf{E})$ is isomorphic to $\mathbf{S}_k$, establishing the assertion. ∎

2.2. *Partition vectors of transitive permutation groups*

In this section, let $\Gamma\subseteq\mathbf{S}_n$ be a transitive permutation group, let $\mathbf{Z}(\Gamma)$ denote the vector set introduced in Section 1, and let $l=\exp\Gamma$. By reasons soon to become visible, the elements $\mathbf{x}\in\mathbf{Z}(\Gamma)$ are called the *partition vectors* of Γ.

First, let us list some simple facts.

(A) *Let* $\mathbf{x}, \mathbf{y}\in\mathbf{Z}(\Gamma)$. *Then also* $\mathbf{x}\otimes\mathbf{y}\in\mathbf{Z}(\Gamma)$ *and, for all* $\gamma\in\Gamma$, $a_\gamma(\mathbf{x}\otimes\mathbf{y})=a_\gamma(\mathbf{x})\cdot a_\gamma(\mathbf{y})$, *i.e.* $\{a_\gamma(\mathbf{x})|\mathbf{x}\in\mathbf{Z}(\Gamma)\}$ *is a representation of degree* 1 *of* $\mathbf{Z}(\Gamma)$.
(B) *Let* $\mathbf{x}, \mathbf{y}\in\mathbf{Z}(\Gamma)$. *Then* $\mathbf{x}=\mathbf{y}$ *iff* $a_\gamma(\mathbf{x})=a_\gamma(\mathbf{y})$ *for all* $\gamma\in\Gamma$.
(C) *Let* $\gamma, \gamma'\in\Gamma$. *Then, for all* $\mathbf{x}\in\mathbf{Z}(\Gamma)$, $a_{\gamma\gamma'}(\mathbf{x})=a_\gamma(\mathbf{x})\cdot a_{\gamma'}(\mathbf{x})$, *i.e.,* $\{a_\gamma(\mathbf{x})|\gamma\in\Gamma\}$ *is a representation of degree* 1 *of* Γ *which is contained in the matrix representation* $\{\mathbf{P}_\gamma|\gamma\in\Gamma\}$ *of* Γ *as an irreducible component.*
(D) *For all* $\gamma\in\Gamma$ *and all* $\mathbf{x}\in\mathbf{Z}(\Gamma)$, $(a_\gamma(x))^l=1$, *i.e.,* $\mathbf{Z}(\Gamma)$ *is a subgroup of* $\mathbf{U}_n(l)$.
(E) *If, for* $\gamma, \gamma'\in\Gamma$ *and some* $i\in\{1, \ldots, n\}$, $\gamma(i)=\gamma'(i)$ *then* $a_\gamma(\mathbf{x})=a_{\gamma'}(\mathbf{x})$ *for all* $\mathbf{x}\in\mathbf{Z}(\Gamma)$.
(F) *Let* $\mathbf{x}, \mathbf{y}\in\mathbf{Z}(\Gamma)$, $\mathbf{x}\neq\mathbf{y}$. *According to* (B), *there is a* $\gamma\in\Gamma$ *with* $a_\gamma(\mathbf{x})\neq a_\gamma(\mathbf{y})$: *therefore,* $\langle\mathbf{x}, \mathbf{y}\rangle=\langle\mathbf{P}_\gamma\mathbf{x}, \mathbf{P}_\gamma\mathbf{y}\rangle=a_\gamma(\mathbf{x})\cdot\overline{a_\gamma(\mathbf{y})}\cdot\langle\mathbf{x}, \mathbf{y}\rangle$ *with* $a_\gamma(\mathbf{x})\cdot\overline{a_\gamma(\mathbf{y})}\neq1$, *implying* $\langle\mathbf{x}, \mathbf{y}\rangle=0$.
(G) *Let* $\gamma_i\in\Gamma$ *with* $\gamma_i(1)=i$ $(i=1, \ldots, n)$. *Then, for all* $\mathbf{x}\in\mathbf{Z}(\Gamma)$, $\mathbf{x}=(a_{\gamma_1}(\mathbf{x}), \ldots, a_{\gamma_n}(\mathbf{x}))^{\mathsf{T}}$.

As an immediate consequence of these propositions, we obtain

THEOREM 3. $\mathbf{Z}(\Gamma)\in\mathscr{E}_l(n)$.

Put $\mathbf{Z}_d(\Gamma) = \mathbf{Z}(\Gamma)\cap\mathbf{U}_d(n)$, $z_d(\Gamma) = |\mathbf{Z}_d(\Gamma)|$, $z(\Gamma) = |\mathbf{Z}(\Gamma)|$. ∎

As a corollary of Theorems 1 and 3, we have

LEMMA 5. $\mathbf{Z}_d(\Gamma)\in\mathscr{E}_d(n)$, $z_d(\Gamma)\leqq s_d(n)$. ∎

THEOREM 4. *Let* Γ' *be the commutator subgroup of* Γ, *let* Γ^* *be the subgroup generated by all permutations of* Γ *keeping an element fixed, and put* $\Gamma^d=\langle\{\gamma|\gamma\in\Gamma, \gamma^d=\mathrm{id}\}\rangle$. *Then*

$$\mathbf{Z}(\Gamma)\cong\Gamma/(\Gamma^*\cdot\Gamma'),\quad \mathbf{Z}_d(\Gamma)\cong\Gamma/(\Gamma^*\Gamma'\Gamma^d).$$

For a proof see [5]. ∎

LEMMA 6. *Let* $\mathbf{E}\in\mathscr{E}_l(n)$. *Then*

(a) $\mathbf{Z}(\Gamma(\mathbf{E}))=\mathbf{E}$,

(b) *if* $\Gamma\subseteqq\mathbf{S}_n$ *is a transitive permutation group with* $\mathbf{Z}(\Gamma)=\mathbf{E}$ *then* $\Gamma\subseteqq\Gamma(\mathbf{E})$.

PROOF. Put $\mathbf{E}'=\mathbf{Z}(\Gamma(\mathbf{E}))$. Then $\mathbf{E}\subseteqq\mathbf{E}'$ and, therefore, $\Gamma(\mathbf{E}')\subseteqq\Gamma(\mathbf{E})\subseteqq\Gamma(\mathbf{Z}(\Gamma(\mathbf{E})))=\Gamma(\mathbf{E}')$, thus $\Gamma(\mathbf{E})=\Gamma(\mathbf{E}')$. Now, $i\sim j|\mathbf{E}'$ implies $i\sim j|\mathbf{E}$. Conversely, if $i\sim j|\mathbf{E}$ then the permutation $\gamma\in\mathbf{S}_n$ with $\gamma(i)=j$, $\gamma(j)=i$, and $\gamma(k)=k$ for $k\neq i,j$ is an element of $\Gamma(\mathbf{E})$ and, consequently, also of $\Gamma(\mathbf{E}')$. This means that γ fixes the equivalence classes of $\mathbf{E}'$, i. e., $i\sim\gamma(i)|\mathbf{E}'$ and thus $i\sim j|\mathbf{E}'$. This implies $|\mathbf{E}|=|\mathbf{E}'|$ and, therefore, $\mathbf{E}=\mathbf{E}'$. ∎

Let $\mathbf{E}\in\mathscr{E}_l(n)$ and put

$\mathscr{G}(\mathbf{E})=\{\Gamma|\Gamma\subseteqq\mathbf{S}_n,\ \Gamma \text{ is transitive},\ \mathbf{Z}(\Gamma)=\mathbf{E}\}$.

Lemma 6 says that $\Gamma(\mathbf{E})\in\mathscr{G}(\mathbf{E})$ and that $\Gamma'\in\mathscr{G}(\mathbf{E})$ implies $\Gamma'\subseteqq\Gamma(\mathbf{E})$. If, in particular, $|\mathbf{E}|=n$ then, according to Theorem 2, $\Gamma(\mathbf{E})$ has order n; therefore, if $|\mathbf{E}|=n$ and $\Gamma'\in\mathscr{G}(\mathbf{E})$ then, necessarily, $\Gamma'=\Gamma(\mathbf{E})$ and from Lemma 4 (with $\bar{\mathbf{E}}=\mathbf{E}$) we conclude that Γ' is isomorphic to $\mathbf{E}=\mathbf{Z}(\Gamma')$ implying, in particular, that Γ' is abelian.

Conversely, if $\tilde{\Gamma}\subseteqq\mathbf{S}_n$ is any transitive abelian permutation group then, by Theorem 4, $z(\tilde{\Gamma})=n$.

Summing up (using Theorem 3 and Lemma 2) we obtain

PROPOSITION 2. *For a transitive permutation group* $\Gamma\subseteqq\mathbf{S}_n$, *the following four statements are equivalent.*

(i) $z(\Gamma)>\frac{n}{2}$,

(ii) $z(\Gamma)=n$,

(iii) Γ *is abelian,*

(iv) Γ *is isomorphic to* $\mathbf{Z}(\Gamma)$.

The term "partition vectors" for the elements of $\mathbf{Z}(\Gamma)$ can now be given a simple interpretation: if $\mathbf{x}\in\mathbf{Z}(\Gamma)$ has order d, then $\mathbf{E}=\{\mathbf{x}^k|k=1,\ldots,d\}\subseteqq\mathbf{Z}(\Gamma)$ is a cyclic group of order d and the partitioning $\mathbf{X}_1(\mathbf{E}),\ldots,\mathbf{X}_d(\mathbf{E})$ of $\{1,\ldots,n\}$ induced by the $\mathbf{E}$-equivalence has the following properties: by Γ the classes $\mathbf{X}_i(\mathbf{E})$ are cyclically permuted (according to $\Gamma(\bar{\mathbf{E}})$) thus forming domains of imprimitivity of Γ. For $d=1$ and $d=n$, the trivial partitionings are obtained.

Conversely, for any system of cyclically permuted imprimitivity domains of Γ there exists a partition vector generating this particular partitioning.

As a simple consequence, we obtain

LEMMA 7. *Let* $\Gamma\subseteqq\mathbf{S}_n$ *be transitive and primitive.*

(a) *If n is not a prime, then* $z(\Gamma)=1$;

(b) *if n is a prime, then* $z(\Gamma)=1$ *or* $z(\Gamma)=n$ *where the latter case occurs iff* Γ *is cyclic of order n.* ∎

The following theorem was proved in [5].

THEOREM 4. *Let p be a prime and $v_p(\Gamma)$ the number of intransitive normal subgroups of Γ having index p. Then*

$$z_p(\Gamma)-1=(p-1)v_p(\Gamma).$$ ∎

3. Simple eigenvalues of transitive graphs

3.1. *The maximum number of simple eigenvalues of a transitive graph*

Let **G** be a transitive graph with n vertices and **x** an eigenvector of **G** belonging to a simple eigenvalue. Then **x** is a common eigenvector of all permutation matrices $\mathbf{P}_\gamma$ with $\gamma \in \mathbf{Aut}(\mathbf{G})$ and from the transitivity of **Aut (G)** it follows that

$$|x_1| = |x_2| = \ldots = |x_n| \neq 0.$$

Therefore, $\frac{1}{x_1}\mathbf{x} \in \mathbf{Z}(\mathbf{Aut}(\mathbf{G}))$, implying

THEOREM 5. *If the graph* **G** *is transitive then* $\sigma(\mathbf{G}) \leq z(\mathbf{Aut}(\mathbf{G}))$. *(Recall that* $\sigma(\mathbf{G})$ *is the number of simple eigenvalues of* **G**.*)* ∎

Let **G** be an undirected, transitive graph with n vertices. The adjacency matrix of **G** being symmetric, the eigenvalues of **G** are real and its eigenvectors can also be chosen real. In particular, to each simple eigenvalue of **G** there belongs exactly one eigenvector contained in $\mathbf{Z}_2(\mathbf{Aut}(\mathbf{G})) \in \mathscr{E}_2(n)$.
This immediately implies

THEOREM 6. *Let* **G** *be an undirected transitive graph with* n *vertices. Then*

$$\sigma(\mathbf{G}) \leq z_2(\mathbf{Aut}(\mathbf{G})) \leq s_2(n).$$ ∎

Thus, for undirected transitive graphs, we have obtained an upper bound for the number of simple eigenvalues depending on the number of vertices only. The next theorem says that, in fact, this bound is sharp.

THEOREM 7. *Let* n *be an arbitrary positive integer and put* $n=2^q k$ *where* k *is odd. Then there exist connected graphs with* n *vertices which are transitive, undirected, loopless, and have* $s_2(n)=2^q$ *simple eigenvalues.*

First we prove

LEMMA 8. *Let* **A** *be the adjacency matrix of a transitive undirected graph* **G** *with* n *vertices which has exactly* s *simple eigenvalues. Then there is an infinite set* $\mathbf{T}=\mathbf{T}(\mathbf{G})$ *of positive integers such that, for every* $t \in \mathbf{T}$, *the graph* $\mathbf{G}'=\mathbf{G}_t(\mathbf{G})$ *with adjacency matrix*

$$\mathbf{A}' = \begin{pmatrix} t\mathbf{A} & \mathbf{I} \\ \mathbf{I} & t\mathbf{A} \end{pmatrix}$$

has exactly $2s$ *simple eigenvalues;* **G′** *is undirected, transitive, and has* $2n$ *vertices.*

PROOF. Let $\lambda_1, \lambda_2, \ldots, \lambda_m$ be the distinct eigenvalues of **G** and let t be so chosen that $t(\lambda_i-\lambda_j)\neq 2$ for $i, j=1, 2, \ldots, m$. Then **G′** has the $2m$ distinct eigenvalues $t\lambda_1 \pm 1, \ldots, t\lambda_m \pm 1$; thus, in particular, $\sigma(\mathbf{G}')=2s$. Clearly, **G′** is undirected and has $2n$ vertices; further, **G** being transitive and **G′** admitting the automorphism $(1, n+1)(2, n+2)\ldots(n, 2n)$, **G′** is transitive, too.

PROOF of THEOREM 7. Let $\mathbf{G}_1^0$ denote the graph consisting of exactly one isolated vertex and, for odd $k \geq 3$, let $\mathbf{G}_k^0$ denote the circuit of length k. The graphs $\mathbf{G}_k^0$

($k=1, 3, 5, \dots$) are connected, transitive, undirected, loopless, and have exactly 2^0k vertices and one simple eigenvalue. By virtue of Lemma 8, starting from $\mathbf{G}_k^0$ we can now easily construct sequences of graphs $\mathbf{G}_k^q$ ($q=0, 1, 2, \dots$) (where $\mathbf{G}_k^{q+1}=$ $=\mathbf{G}_t(\mathbf{G}_k^q)$, $t\in\mathbf{T}(\mathbf{G}_k^q)$) which are connected, transitive, undirected, loopless, and have exactly 2^qk vertices and 2^q simple eigenvalues. ∎

REMARK. If $k=1$ then we may choose $t=2$ in each step of our construction. The graphs so obtained (see Fig. 1) have the following properties:

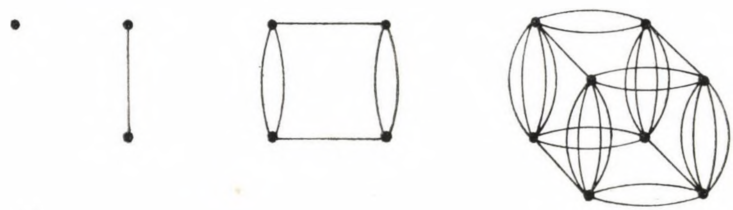
(i) $\mathbf{G}_1^q$ *has* 2^q *vertices,*

(ii) $\mathbf{G}_1^q$ *is a bipartite graph which is regular of degree* $r_q=2^q-1$,

(iii) *replacing multiple edges by single ones transforms* $\mathbf{G}_1^q$ *into the graph of the q-dimensional cube,*

(iv) *all non-trivial automorphisms of* $\mathbf{G}_1^q$ *are involutions, i.e.,* $\exp \mathbf{Aut}(\mathbf{G})=2$ *for* $q>0$,

(v) *all eigenvalues of* $\mathbf{G}_1^q$ *are simple and form the equidistant spectrum* $[-r_q, -r_q+2, \dots, r_q-2, r_q]$.



G_1^0 G_1^1 G_1^2 G_1^3

Fig. 1

The problem of determining all integers n for which there exists a *schlicht*, undirected, connected, transitive graph with exactly n vertices and $s_2(n)$ simple eigenvalues, however, seems to be difficult. We can only present

CONJECTURE 1. *Let q and k be positive integers, k being odd. Then there is a schlicht, undirected, transitive, connected graph with* $n=2^qk$ *vertices and* $s_2(n)=2^q$ *simple eigenvalues iff* $k\geqq 2^{q-1}$.

Using well-known constructions (forming products of graphs), in [6] the following theorems could be proved.

THEOREM 8. *Let q be a positive integer and let* $k_i\geqq 3$ ($i=1, 2, \dots, q$) *be odd integers (not necessarily distinct). Put* $n=2^qk_1k_2\dots k_q$. *Then there are schlicht, undirected, transitive, connected graphs having exactly n vertices and* $s_2(n)=2^q$ *simple eigenvalues.* ∎

THEOREM 9. *For any odd integer* $k\geqq 3$ *there is a schlicht, undirected, transitive, connected graph having exactly 4k vertices and 4 simple eigenvalues. For* $k=1$, *there is no such graph.* ∎

3.2. *Simple eigenvalues and divisors of transitive graphs*

DEFINITION 2. Let $\mathbf{G}$ be a graph (not necessarily connected) and let $\mathbf{G}'$ be a graph with vertex set $\mathbf{V}(\mathbf{G}')=\{1, 2, \ldots, m\}$ and adjacency matrix $A'=(a'_{ij})$.

$\mathbf{G}'$ is called a *divisor of* $\mathbf{G}$ (in symbols: $\mathbf{G}'|\mathbf{G}$) iff the vertex set $\mathbf{V}(\mathbf{G})$ can be partitioned into m non-void classes $\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_m$ in such a way that, for all $a \in \mathbf{V}_i$ $(i=1, 2, \ldots, m)$ and $j=1, 2, \ldots, m$, $W^+_{\mathbf{G}}(a, \mathbf{V}_j)=a'_{ij}$ and $W^-_{\mathbf{G}}(a, \mathbf{V}_j)=a'_{ji}$.

If, in addition, all subgraphs $\mathbf{G}[\mathbf{V}_j]$ of $\mathbf{G}$ spanned by the sets $\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_m$ are isomorphic then $\mathbf{G}'$ is called an *i-divisor of* $\mathbf{G}$ (in symbols: $\mathbf{G}'|_i\mathbf{G}$).

If $\mathbf{G}$ is undirected then so is every divisor of $\mathbf{G}$ (neglecting directed loops that may possibly occur). (For a detailed treatment of the divisor concept, the reader is referred to [1, Chapter 4].)

Now consider a transitive graph $\mathbf{G}$ having n vertices and let Γ denote a transitive subgroup of $\mathbf{Aut}(\mathbf{G})$ and $\mathbf{E} \subseteq \mathbf{Z}(\Gamma)$ a feasible vector set of order m. Then $\Gamma \subseteq \Gamma(\mathbf{E})$ and the partitioning $(\mathbf{X}_1(\mathbf{E}), \mathbf{X}_2(\mathbf{E}), \ldots, \mathbf{X}_m(\mathbf{E}))$ of $\mathbf{V}(\mathbf{G})=\{1, 2, \ldots, n\}$ has the following properties:

(1) $|\mathbf{X}_i(\mathbf{E})|=\dfrac{n}{m}$, $i=1, 2, \ldots, m$ (see 2.1), *and the equivalence classes* $\mathbf{X}_i(\mathbf{E})$ *are transitively permuted by* Γ *according to* $\Gamma(\tilde{\mathbf{E}})$ (see Theorem 2), $\Gamma(\tilde{\mathbf{E}})$ *being an abelian permutation group of order* m *which is isomorphic to* $\mathbf{E}$.

(2) *Let* $i \in \{1, 2, \ldots, m\}$, $a \in \mathbf{X}_i(\mathbf{E})$, $\gamma \in \Gamma$ *and let* $\tilde{\gamma}$ *denote the permutation of* $\Gamma(\tilde{\mathbf{E}})$ *determined by* γ *(i. e., satisfying* $\tilde{\gamma}(i)=j$ *iff* $\gamma(\mathbf{X}_i(\mathbf{E}))=\mathbf{X}_j(\mathbf{E})$). *Then*

$$W^+_{\mathbf{G}}(a, \mathbf{X}_j(\mathbf{E})) = W^+_{\mathbf{G}}(\gamma(a), \mathbf{X}_k(\mathbf{E})) \qquad (j = 1, 2, \ldots, m;\ k = \tilde{\gamma}(j)),$$

where $\gamma(a) \in \mathbf{X}_l(\mathbf{E})$ *with* $l=\tilde{\gamma}(i)$.

(3) *Because of the transitivity of* Γ, *for any pair of vertices* $a, b \in \mathbf{X}_i(\mathbf{E})$, *both the following equalities hold:*

$$W^+_{\mathbf{G}}(a, \mathbf{X}_j(\mathbf{E})) = W^+_{\mathbf{G}}(b, \mathbf{X}_j(\mathbf{E})),$$

$$W^-_{\mathbf{G}}(a, \mathbf{X}_j(\mathbf{E})) = W^-_{\mathbf{G}}(b, \mathbf{X}_j(\mathbf{E})), \quad j = 1, 2, \ldots, m.$$

(4) *All the* $\mathbf{X}_i(\mathbf{E})$ *having the same cardinality* (see (1)), *for any pair of vertices* $a \in \mathbf{X}_i(\mathbf{E})$, $c \in \mathbf{X}_j(\mathbf{E})$ *we obtain*

$$W^-_{\mathbf{G}}(a, \mathbf{X}_j(\mathbf{E})) = W^+_{\mathbf{G}}(c, \mathbf{X}_i(\mathbf{E})) \qquad (i, j = 1, 2, \ldots, m).$$

Now let $a \in \mathbf{X}_i(\mathbf{E})$ and put $W^+_{\mathbf{G}}(a, \mathbf{X}_j(\mathbf{E}))=\tilde{a}_{ij}$ $(i, j=1, 2, \ldots, m)$; by (3) and (4), $\tilde{a}_{ij}$ is well defined. Let $\tilde{\mathbf{G}}=\mathbf{G}(\Gamma, \mathbf{E})$ denote the graph with vertex set $\mathbf{V}(\tilde{\mathbf{G}})= =\{1, 2, \ldots, m\}$ and adjacency matrix $\tilde{\mathbf{A}}=(\tilde{a}_{ij})$: then, clearly, $\tilde{\mathbf{G}}$ is a divisor of $\mathbf{G}$. Further, it is easy to see that all the spanned subgraphs $\mathbf{G}[\mathbf{X}_i(\mathbf{E})]$ $(i=1, 2, \ldots, m)$ are isomorphic, thus $\tilde{\mathbf{G}}|_i\mathbf{G}$.

From (2) it immediately follows that $\Gamma(\tilde{\mathbf{E}})$ is a subgroup of $\mathbf{Aut}(\tilde{\mathbf{G}})$, where, by Lemma 6, $\mathbf{Z}(\Gamma(\tilde{\mathbf{E}}))=\tilde{\mathbf{E}}$.

We shall now consider the case that $\mathbf{E}$ is a cyclic group generated by some element $\mathbf{x} \in \mathbf{Z}(\Gamma)$ of order d.

Let ε be a primitive d-th root of unity. Then the sets $\mathbf{X}_k=\{i|x_i=\varepsilon^{k-1}\}$ $(k=1, 2, \dots$ $\dots, d)$ are the $\mathbf{E}$-equivalence classes. $\Gamma(\bar{\mathbf{E}})$ is a cyclic permutation group of order d and, therefore, the adjacency matrix $\tilde{\mathbf{A}}=(\tilde{a}_{ij})$ of $\bar{\mathbf{G}}=\mathbf{G}(\Gamma, \mathbf{E})$ is a circulant matrix. Recall that $\mathbf{A}=(a_{ij})$ is the adjacency matrix of $\mathbf{G}$. Putting $r_k=\tilde{a}_{1k}$ we obtain for all $i\in\mathbf{X}_k$

$$\sum_{j=1}^{n} a_{ij}x_j = \sum_{l=1}^{d} \tilde{a}_{kl}\varepsilon^{l-1} = \varepsilon^{k-1}\sum_{l=1}^{d} r_{l-k+1}\varepsilon^{l-k} = x_i\sum_{j=1}^{d} r_j\varepsilon^{j-1}$$

(subscripts of the letter r to be reduced modulo d).

Thus we have proved

THEOREM 10. *Let* $\mathbf{G}$ *be a transitive graph being regular of degree r, let* $\mathbf{x}\in\mathbf{Z}(\Gamma)$ *be a partition vector of order d with respect to a transitive subgroup Γ of* **Aut** ($\mathbf{G}$), *and let ε be a primitive d-th root of unity. Then* $\mathbf{x}$ *is an eigenvector of* $\mathbf{G}$, *the corresponding eigenvalue λ having the form*

$$\lambda = r_1+r_2\varepsilon+\dots+r_d\varepsilon^{d-1},$$

where $r_1, r_2, \dots, r_d$ *are non-negative integers satisfying* $r_1+r_2+\dots+r_d=r.$ ∎

In what follows the eigenvalue of $\mathbf{G}$ corresponding to the partition vector $\mathbf{x}\in\mathbf{Z}(\Gamma)$ will be briefly denoted by $\lambda_{\mathbf{G}}(\mathbf{x})$.

Combining the last results with the statements **(1)—(4)**, we obtain

THEOREM 11. *Let* $\mathbf{G}$ *be a transitive graph and Γ a transitive subgroup of* **Aut** ($\mathbf{G}$). *Let* $\mathbf{E}=\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}\subseteq\mathbf{Z}(\Gamma)$ *be a feasible vector set of order m and put* $\bar{\mathbf{G}}=\mathbf{G}(\Gamma, \mathbf{E})$. *Then the following assertions hold.*

(I) $\bar{\mathbf{G}}$ *is an i-divisor of* $\mathbf{G}$ *having m vertices.*

(II) *If* $\mathbf{G}$ *is undirected then so is* $\bar{\mathbf{G}}$ *(neglecting directed loops).*

(III) $\bar{\mathbf{G}}$ *is transitive and* $\Gamma(\bar{\mathbf{E}})$ *is a transitive abelian subgroup of* **Aut** ($\bar{\mathbf{G}}$) *which is isomorphic to* $\mathbf{E}$.

(IV) *Every vector from* $\bar{\mathbf{E}}$ *is an eigenvector of* $\bar{\mathbf{G}}$ *and, for all* $\mathbf{x}\in\mathbf{E}$, $\lambda_{\bar{\mathbf{G}}}(\bar{\mathbf{x}})=\lambda_{\mathbf{G}}(\mathbf{x})$, *implying that* $\bar{\mathbf{G}}$ *has the m (not necessarily distinct) eigenvalues* $\lambda_{\mathbf{G}}(\mathbf{x}_1)$, $\lambda_{\mathbf{G}}(\mathbf{x}_2), \dots, \lambda_{\mathbf{G}}(\mathbf{x}_m)$.

(V) *If* $\mathbf{E}=\mathbf{Z}(\Gamma)$ *then* $\sigma(\mathbf{G})\leq\sigma(\mathbf{G}(\Gamma, \mathbf{Z}(\Gamma)))\leq z(\Gamma)$. ∎

THEOREM 12. *Let* $\mathbf{G}$ *be a transitive graph with n vertices. If* **Aut** ($\mathbf{G}$) *contains a transitive abelian subgroup then, for every prime factor p of n, there is an i-divisor of* $\mathbf{G}$ *having exactly p vertices.*

PROOF. If $\Gamma\subseteq$ **Aut** ($\mathbf{G}$) is transitive and abelian then $\mathbf{Z}(\Gamma)$ is a feasible vector set of order n and to each prime factor p of n there exists $\mathbf{x}\in\mathbf{Z}(\Gamma)$ of order p. Let $\mathbf{E}$ be the subgroup of $\mathbf{Z}(\Gamma)$ generated by $\mathbf{x}$: then $\mathbf{G}(\Gamma, \mathbf{E})$ is an i-divisor of $\mathbf{G}$ with exactly p vertices. ∎

For (undirected) transitive graphs, we can draw some more conclusions from Theorem 10.

PROPOSITION 3. *If* **G** *is a transitive graph being regular of degree r and* Γ *is a transitive subgroup of* **Aut (G)** *then, for all* $\mathbf{x}\in \mathbf{Z}_2(\Gamma)$,

$$\lambda_{\mathbf{G}}(\mathbf{x}) = \varrho-(r-\varrho) = 2\varrho-r \quad \textit{where} \quad \varrho\in\{0, 1, \dots, r\}.$$

PROPOSITION 4. *If* **G** *is an undirected transitive graph being regular of degree r then* $\sigma(\mathbf{G})\leqq r+1$.

If, in addition, **G** is schlicht and has more than 2 vertices then $\sigma(\mathbf{G})<n$.

Thus every simple eigenvalue of an undirected transitive graph **G** which is regular of degree r has the form $\lambda=2\varrho-r$ where $\varrho\in\{0, 1, \dots, r\}$. If **G** has an eigenvalue $\lambda\neq r$ then its vertex set is covered by a pair of isomorphic spanned subgraphs which are transitive and regular of degree $\varrho=(\lambda+r)/2$. If **G** has n vertices then each of these subgraphs has $n/2$ vertices and, therefore, $(\varrho n)/4$ edges implying $\varrho n\equiv 0$, mod 4. Combining these results with Theorem 6, we obtain

THEOREM 13. *Let* **G** *be an undirected, transitive, connected graph with* $n=2^q k$ *(k being odd) vertices being regular of degree r. Then the following assertions hold.*

(i) *If* $q=0$ *then* $\lambda=r$ *is the only simple eigenvalue of G.*

(ii) *If* $q=1$ *then* **G** *has at most one simple eigenvalue* λ *different from r and, if it exists, it is of the form* $\lambda=4\varrho-r$ *where* $\varrho\in\{0, 1, \dots, \frac{1}{2}(r-1)\}$.

(iii) *If* $q\geqq 2$ *then* **G** *has at most* 2^q *simple eigenvalues including* $\lambda=r$; *all of them have the form* $\lambda=2\varrho-r$ *where* $\varrho\in\{0, 1, \dots, r\}$. ∎

3.3. *Simple eigenvalues and automorphism group of transitive graphs*

In this section, we shall investigate the relations between the automorphism group of a transitive graph and the number of its simple eigenvalues.

From Proposition 2 and Theorem 5 we deduce

THEOREM 14. *If a transitive graph* **G** *with n vertices has more than* $\frac{n}{2}$ *simple eigenvalues then its automorphism group is abelian.* ∎

If, conversely, **G** is a transitive graph with n vertices whose automorphism group has a transitive abelian subgroup Γ, then, in the general case, we have no statements on $\sigma(\mathbf{G})$. But from Theorem 10 and Proposition 2 we know that **G** has n linearly independent eigenvectors which, with respect to the operation $\otimes$, constitute a group isomorphic to Γ. In particular it follows from this observation that the adjacency matrix of **G** is non-derogatory iff all eigenvalues of **G** are simple. In 1971, A. Mowshowitz (see, e.g., [1, p. 144]) showed that the automorphism group of an arbitrary graph which has a non-derogatory adjacency matrix is abelian, thus we have proved

THEOREM 15. *The adjacency matrix of a transitive graph G is non-derogatory iff all eigenvalues of G are simple.* ∎

LEMMA 10. *Let* **G** *and* **G′** *be transitive graphs each having n vertices and assume that each of* **Aut (G)** *and* **Aut (G′)** *contains a transitive abelian subgroup,* $\Gamma\subseteqq$ **Aut (G)** *and* $\Gamma'\subseteqq$ **Aut (G′)**, *say; further suppose that there is an isomorphism* φ *mapping* $\mathbf{Z}(\Gamma)$ *onto* $\mathbf{Z}(\Gamma')$ *and satisfying* $\lambda_{\mathbf{G}}(\mathbf{x})=\lambda_{\mathbf{G}'}(\varphi(\mathbf{x}))$ *for all* $\mathbf{x}\in\mathbf{Z}(\Gamma)$. *Then* **G** *and* **G′** *are isomorphic.*

PROOF. Let $\mathbf{A}$ and $\mathbf{A}'$ be the adjacency matrices of $\mathbf{G}$ and $\mathbf{G}'$, respectively, let $\mathbf{X}=(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ be the matrix whose column vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$ are the elements of $\mathbf{Z}(\Gamma)$, and let Λ denote the diagonal matrix having the eigenvalues $\lambda_{\mathbf{G}}(\mathbf{x}_1), \lambda_{\mathbf{G}}(\mathbf{x}_2), \dots, \lambda_{\mathbf{G}}(\mathbf{x}_n)$ in its main diagonal. Then, by Theorem 10, $\mathbf{AX}=\mathbf{X}\Lambda$, implying $\mathbf{A}'\mathbf{X}'=\mathbf{X}'\Lambda$ with $\mathbf{X}'=(\varphi(\mathbf{x}_1), \varphi(\mathbf{x}_2), \dots, \varphi(\mathbf{x}_n))$. By Proposition 1 there is a $\gamma\in\mathbf{S}_n$ satisfying $\varphi(\mathbf{x})=\mathbf{P}_\gamma\mathbf{x}$ for all $\mathbf{x}\in\mathbf{Z}(\Gamma)$, i.e., $\mathbf{X}'=\mathbf{P}_\gamma\mathbf{X}$, immediately implying $\mathbf{A}'=\mathbf{P}_\gamma\mathbf{X}\Lambda(\mathbf{P}_\gamma\mathbf{X})^{-1}=\mathbf{P}_\gamma\mathbf{X}\Lambda\mathbf{X}^{-1}\mathbf{P}_\gamma^{-1}=\mathbf{P}_\gamma\mathbf{A}\mathbf{P}_\gamma^{-1}$. ▮

THEOREM 16. *Let* $\mathbf{G}$ *be a connected transitive graph with n vertices whose automorphism group is primitive. Then the following assertions hold.*

(1) *If n is not a prime then* $\sigma(\mathbf{G})=1$.

(2) *If* $n=p$ *is a prime then* $\sigma(\mathbf{G})=1$ *or* $\sigma(\mathbf{G})=p$, *where* $\sigma(\mathbf{G})=p$ *iff* $\mathbf{Aut}(\mathbf{G})$ *is a cyclic group of order p.*

PROOF. Assertion (1) follows immediately from Lemma 7 and Theorem 5.

If $n=p$ is a prime then there is a transitive cyclic permutation group $\Gamma\subseteq\mathbf{Aut}(\mathbf{G})$ having order p. Then $\mathbf{E}=\mathbf{Z}(\Gamma)=\mathbf{Z}_p(\Gamma)\in\mathscr{E}_p(p)$ and, denoting the first row of the adjacency matrix of $\mathbf{G}$ by $\mathbf{a}^T$, we obtain $\lambda_{\mathbf{G}}(\mathbf{x})=\langle\mathbf{x}, \mathbf{a}\rangle$ for all $\mathbf{x}\in\mathbf{E}$.

If $\sigma(\mathbf{G})=p$ then, by Theorem 14, $\Gamma=\mathbf{Aut}(\mathbf{G})$. Assume $\sigma(\mathbf{G})<p$. Then there is at least one pair of distinct vectors $\mathbf{x}_1, \mathbf{x}_2\in\mathbf{E}$ satisfying $\lambda_{\mathbf{G}}(\mathbf{x}_1)=\lambda_{\mathbf{G}}(\mathbf{x}_2)$ where $\mathbf{x}_1\neq\mathbf{u}\neq\mathbf{x}_2$ since $\lambda_{\mathbf{G}}(\mathbf{u})$ is a simple eigenvalue of $\mathbf{G}$ (because $\mathbf{G}$ is connected). Then $\mathbf{x}_2=\mathbf{x}_1^k$ for some $k\in\{2, 3, \dots, p-1\}$ and the mapping $\varphi(\mathbf{x})=\mathbf{x}^k$ $(\mathbf{x}\in\mathbf{E})$ is an automorphism of $\mathbf{E}$; using Proposition 1 we conclude that there is a $\gamma\in\mathbf{S}_n$ with $\mathbf{P}_\gamma\mathbf{x}=\mathbf{x}^k$ for all $\mathbf{x}\in\mathbf{E}$.

As is well known (see [2]), $\lambda_{\mathbf{G}}(\mathbf{x}_1)=\langle\mathbf{x}_1, \mathbf{a}\rangle=\langle\mathbf{P}_\gamma\mathbf{x}_1, \mathbf{P}_\gamma\mathbf{a}\rangle=\langle\mathbf{x}_2, \mathbf{P}_\gamma\mathbf{a}\rangle=\langle\mathbf{x}_2, \mathbf{a}\rangle=\lambda_{\mathbf{G}}(\mathbf{x}_2)$ implies $\mathbf{P}_\gamma\mathbf{a}=\mathbf{a}$. Thus $\lambda_{\mathbf{G}}(\mathbf{x})=\langle\mathbf{x}, \mathbf{a}\rangle=\langle\mathbf{P}_\gamma\mathbf{x}, \mathbf{P}_\gamma\mathbf{a}\rangle=\langle\mathbf{x}^k, \mathbf{a}\rangle=\lambda_{\mathbf{G}}(\varphi(\mathbf{x}))$ for all $\mathbf{x}\in\mathbf{E}$ implying $\sigma(\mathbf{G})=1$.

Further, according to Lemma 10, γ is a non-trivial automorphism of $\mathbf{G}$ with $\gamma(1)=1$, i.e., $\gamma\notin\Gamma$, thus $\Gamma\neq\mathbf{Aut}(\mathbf{G})$. ▮

For undirected transitive graphs we obtain from Theorems 6 and 4

THEOREM 17. *Let* $\mathbf{G}$ *be a connected, undirected, transitive graph. Then* $\sigma(\mathbf{G})\leqq v_2(\mathbf{Aut}(\mathbf{G}))+1$. ▮

COROLLARY. *Let* $\mathbf{G}$ *be an undirected transitive graph, let* Γ *be a transitive subgroup of* $\mathbf{Aut}(\mathbf{G})$, *and let t denote the number of elements of order two contained in* Γ. *Then* $\sigma(\mathbf{G})\leqq t+1$.

As is well known (see [4]), every connected, undirected, transitive graph whose automorphism group contains a transitive abelian subgroup has a Hamiltonian circuit. For such graphs, we shall now establish a somewhat stronger theorem.

We need some more notation.

The *product* $\mathbf{G}_1\times\mathbf{G}_2$ of the undirected schlicht graphs $\mathbf{G}_1, \mathbf{G}_2$ is defined to be the graph $\mathbf{G}$ with vertex set $\mathbf{V}(\mathbf{G})=\mathbf{V}(\mathbf{G}_1)\times\mathbf{V}(\mathbf{G}_2)$ in which two vertices (i_1, j_1) and (i_2, j_2) are adjacent iff either $i_1=i_2$ and j_1 and j_2 are adjacent in G_2, or i_1 and i_2 are adjacent in G_1 and $j_1=j_2$. Let $\mathbf{C}_2$ denote the complete graph on two vertices and let, for $l>2$, $\mathbf{C}_l$ be the circuit of length l; put $\mathbf{C}_{d_1}\times\mathbf{C}_{d_2}\times\dots\times\mathbf{C}_{d_r}=\mathbf{C}_{d_1,d_2,\dots,d_r}$. If $d_1=d_2=\dots=d_r=d$ then we shall briefly write $\mathbf{C}_d^r$ instead of $\mathbf{C}_{d_1,d_2,\dots,d_r}$. Every graph $\mathbf{C}_{d_1, d_2, \ldots, d_r}$ has a Hamiltonian circuit (see [8]); $\mathbf{C}_2^r$ is the r-dimensional cube graph.

For a group $\mathbf{H}$ and a subset Δ of $\mathbf{H}$, let $\mathbf{G}=\mathbf{G}(\mathbf{H}, \Delta)$ denote the Cayley graph of $\mathbf{H}$ with respect to Δ (i.e., $V(\mathbf{G})=\mathbf{H}$ and two vertices $h_1, h_2 \in \mathbf{H}$ are adjacent in $\mathbf{G}$ iff $h_1 h_2^{-1} \in \Delta \cup \Delta^{-1}$).

LEMMA 11. *Let* $\mathbf{H}$ *be an abelian group and* $\Delta=\{h_1, h_2, \ldots, h_r\} \subseteq \mathbf{H}$ *a basis of* $\mathbf{H}$ *where* h_i *has order* d_i $(i=1, 2, \ldots, r)$ ($\mathbf{H}$ is said to be of type $(d_1, d_2, \ldots, d_r)$, see [3]). *Then*

$$\mathbf{G}(\mathbf{H}, \Delta) \cong \mathbf{G}(\mathbf{H}_1, \{h_1\}) \times \ldots \times \mathbf{G}(\mathbf{H}_r, \{h_r\}) \cong \mathbf{C}_{d_1, d_2, \ldots, d_r}$$

where $\mathbf{H}_i$ *is the subgroup of* $\mathbf{H}$ *generated by* h_i.

PROOF. $\mathbf{H}$ is the direct product of subgroups $\mathbf{H}_i$ $(i=1, 2, \ldots, r)$. Let $h, h' \in \mathbf{H}$ where $h=h_1^{\alpha_1} h_2^{\alpha_2} \ldots h_r^{\alpha_r}$ and $h'=h_1^{\alpha_1'} h_2^{\alpha_2'} \ldots h_r^{\alpha_r'}$. If $h \cdot h'^{-1}=h_j$ or h_j^{-1} for some $j \in \{1, 2, \ldots, r\}$, then $\alpha_i=\alpha_i'$ for $i \neq j$ and $\alpha_j=\alpha_j' \pm 1$, implying immediately the first part of the assertion.

It is easy to see that $\mathbf{G}(\mathbf{H}_i, \{h_i\})$ is isomorphic to $\mathbf{C}_{d_i}$ $(i=1, 2, \ldots, r)$, thus $\mathbf{G}(\mathbf{H}, \Delta) \cong \mathbf{C}_{d_1, d_2, \ldots, d_r}$. ∎

THEOREM 18. *Let* $\mathbf{G}$ *be a connected, undirected, transitive graph with* $n>1$ *vertices whose automorphism group contains a transitive abelian group* Γ *of type* $(d_1, d_2, \ldots, d_r)$ *as a subgroup. Then* $\mathbf{G}$ *contains a subgraph isomorphic to* $\mathbf{C}_{d_1, d_2, \ldots, d_r}$.

PROOF. According to a theorem of G. Sabidussi [7, Theorem 2], Γ contains a generating set Δ^* such that $\mathbf{G}(\Gamma, \Delta^*)$ is isomorphic to a subgraph of $\mathbf{G}$. Then there is a basis $\Delta \subseteq \Delta^*$: clearly, $\mathbf{G}(\Gamma, \Delta)$ is a subgraph of $\mathbf{G}(\Gamma, \Delta^*)$ and, by Lemma 11, $\mathbf{G}(\Gamma, \Delta) \cong \mathbf{C}_{d_1, d_2, \ldots, d_r}$. ∎

Summing up, from Theorem 1, Theorem 6, Lemma 5, Proposition 2, Proposition 4, and Theorem 18 we obtain

THEOREM 19. *Let* $\mathbf{G}$ *be an undirected transitive graph with* n *vertices all of whose eigenvalues are simple. Then the following assertions hold.*

(I) $\mathbf{G}$ *is connected.*
(II) $n=2^q$ *for some non-negative integer* q.
(III) *The automorphism group of* $\mathbf{G}$ *is transitive and abelian and, if* $n>1$, *has exponent* 2.
(IV) $\mathbf{G}$ *is regular of degree* $r \geqq 2^q-1$.
(V) $\mathbf{G}$ *contains a q-dimensional cube graph* $\mathbf{C}_2^q$ *as a spanning subgraph.*
(VI) *If* $-r$ *is an eigenvalue of* $\mathbf{G}$, *then, by replacing multiple edges by single ones,* $\mathbf{G}$ *is transformed into the cube graph* $\mathbf{C}_2^q$.
(VII) *If* $n>2$, *then* $\mathbf{G}$ *has multiple edges, i.e.,* $\mathbf{G}$ *is not a schlicht graph.*

Let $\mathscr{G}(q)$ denote the set of all undirected transitive graphs having 2^q vertices and 2^q simple eigenvalues. On the one hand, $\mathscr{G}(q)$ is not empty (see Section 3.1) but, on the other hand, for $q \geqq 2$ does not contain any schlicht graphs. Thus the question arises how far a graph from $\mathscr{G}(q)$ can "deviate" from the property of being schlicht. As a measure of this deviation the maximum number of pairwise parallel edges occurring in $\mathbf{G}$ (called the *edge multiplicity* em$(\mathbf{G})$) may be used. For the special graphs

$\mathbf{G}_1^q \in \mathcal{G}(q)$ constructed in Section 3.1, em $(\mathbf{G}_1^q)=2^{q-1}$ $(q \geqq 1)$, and we conjecture that, in fact, these graphs have the minimum edge multiplicity among all graphs of $\mathcal{G}(q)$:

CONJECTURE 2. *For* $q \geqq 1$ *and all* $\mathbf{G} \in \mathcal{G}(q)$, em $(\mathbf{G}) \geqq$ em $(\mathbf{G}_1^q)=2^{q-1}$.

Conjecture 2 is closely related to Conjecture 1, as the following theorem shows.

THEOREM 20. *If for a positive integer* q *and a positive odd integer* k *there exists an undirected, schlicht, transitive graph* $\mathbf{G}$ *with exactly* $2^q k$ *vertices and* 2^q *simple eigenvalues, then there is also an undirected graph* $\mathbf{G}' \in \mathcal{G}(q)$ *with* em $(\mathbf{G}') \leqq k$.

PROOF. Put $\mathbf{E}=\mathbf{Z}_2(\mathbf{Aut}(\mathbf{G}))$. Then, by Theorem 11, $\mathbf{G}'=\mathbf{G}(\mathbf{Aut}(\mathbf{G}), \mathbf{E})$ is an undirected transitive graph having 2^q vertices and 2^q simple eigenvalues. Loops — if they occur — may be deleted since they have no influence on the number of simple eigenvalues.

Further, $|X_i(E)|=k$ $(i=1, 2, \ldots, 2^q)$ and, since $\mathbf{G}$ is schlicht, we obtain as an immediate consequence that em $(\mathbf{G}') \leqq k$. ▮

REFERENCES

[1] CVETKOVIĆ, D. M., DOOB, M. and SACHS, H., *Spectra of graphs. Theory and application,* VEB Deutscher Verlag der Wissenschaften, Berlin, 1980; Pure and Applied Mathematics, 87, Academic Press, Inc., New York—San Francisco—London, 1980. 2nd edition, 1982. *MR* **81i**: 05054.

[2] ELSPAS, B. and TURNER, J., Graphs with circulant adjacency matrices, *J. Combinatorial Theory* **9** (1970), 297—307. *MR* **42** # 7540.

[3] HUPPERT, B., *Endliche Gruppen I,* Die Grundlehren der mathematischen Wissenschaften, Band 134, Springer-Verlag, Berlin—Heidelberg—New York, 1967. *MR* **37** # 302.

[4] LOVÁSZ, L., *Combinatorial problems and exercises,* North-Holland Publishing Co., Amsterdam—New York, 1979 (§ 12, no. 17 on p. 76). *MR* **80m**: 05001.

[5] SACHS, H. and STIEBITZ, M., Automorphism group and spectrum of a graph, *Studies in Pure Mathematics* (to the memory of Paul Turán), Akadémiai Kiadó, Budapest, 1983, 587—604.

[6] SACHS, H., and STIEBITZ, M., Konstruktion schlichter transitiver Graphen mit maximaler Anzahl einfacher Eigenwerte, *Math. Nachr.* **100** (1981), 145—150.

[7] SABIDUSSI, G., Vertex-transitive graphs, *Monatsh. Math.* **68** (1964), 426—438. *MR* **31** # 91.

[8] VIZING, V. G., The cartesian product of graphs, *Vyčisl. Sistemy,* No. 9 (1963), 30—43. *MR* **35** # 81.

(Received February 15, 1982)

TECHNISCHE HOCHSCHULE ILMENAU
POSTFACH 327
DDR—6300 ILMENAU
GERMAN DEMOCRATIC REPUBLIC

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 91—99

ON S-SYMMETRIC MATRICES OVER A FIELD F OF CHARACTERISTIC DIFFERENT FROM TWO

A. LEE

To the memory of Professor L. Rédei

Abstract

Let **F** be a field, char $\mathbf{F}\neq 2$, and let $M_n(\mathbf{F})$ denote the set of all $n\times n$ matrices over **F**. A matrix $A\in M_n(\mathbf{F})$ is called s-symmetric (secondary symmetric) if it is symmetric concerning the secondary diagonal, consisting of the j, k^{th} entries with $j+k=n+1$. The object of this paper is to prove the following

THEOREM. *There exists an s-symmetric matrix over* **F** *with arbitrary prescribed elementary divisors (over* **F***).*

The significance of this theorem is enlightened by the facts that on the one hand

(i) the ordinary symmetry may condition very strongly the elementary divisors of a matrix over certain fields, for example over the reals, and on the other hand

(ii) whenever $-\varepsilon$ as well as 2ε (where ε denotes the unit of **F**) is a square in **F**, then there is an inner automorphism φ of $M_n(\mathbf{F})$, $\varphi: A\to U^{-1}AU$, which is a one-to-one mapping of the s-symmetric matrices onto the symmetric matrices.

Statement (ii) is also proved in this paper.

1. Introduction

Let **F** be a field and let $M_n(\mathbf{F})$ denote the set of all $n\times n$ matrices over **F**. A matrix $A\in M_n(\mathbf{F})$ is s-symmetric* — i.e. symmetric concerning the secondary diagonal — iff it satisfies the equality

$$VA^TV = A, \tag{1.1}$$

where $V=[\delta_{j,n+1-k}]$ and δ_{pq} denotes the Kronecker delta function with values in **F**, i.e. V is the permutation matrix containing the ε's (the unit of **F**) in the secondary diagonal.

The main theme of this paper is to clarify the question of what elementary divisors the s-symmetric matrices over **F** can have. This question has been answered in our previous paper [3] if $\mathbf{F}=\mathbf{C}$ (the complexes). The extended answer is (Theorem 4) the main result of this paper:

Let **F** be a field and char $\mathbf{F}\neq 2$. There exists an s-symmetric matrix over **F** with arbitrary prescribed elementary divisors (over **F**).

Hence the s-symmetry does not influence the elementary divisors of a matrix over an arbitrary field **F**, char $\mathbf{F}\neq 2$, which is not true in the case of the ordinary

---

1980 *Mathematics Subject Classification*. Primary 15A57; Secondary 15A18, 15A21.

Key words and phrases. S-symmetric matrices, symmetric matrices, elementary divisors, classical canonical matrix.

* In the literature the s-symmetric matrices are called also as persymmetric matrices.

symmetry, as it is shown by the spectral property of real symmetric matrices. But as a consequence of Theorem 4 and 5 of this paper the following statement is valid.

There is a symmetric matrix in $M_{2k}(\mathbf{F})$ (resp. in $M_{2k+1}(\mathbf{F})$) with arbitrary prescribed elementary divisors over $\mathbf{F}$, if $-\varepsilon$ is a square (resp. $-\varepsilon$ and 2ε are squares) in $\mathbf{F}$, where $\mathbf{F}$ is a field char $\mathbf{F}\neq 2$.

The considerations of the main theme need some preparations, namely decompositions of some Hankel matrices. These are proposed in Section 2. The main results are contained in Section 3.

2. Preliminaries and decompositions of some Hankel matrices

Let $\mathbf{F}$ be an arbitrary field, char $\mathbf{F}\neq 2$. The unit and zero element of $\mathbf{F}$ will be denoted by ε and 0. The set of $n\times n$ matrices over $\mathbf{F}$ will be denoted by $M_n(\mathbf{F})$, while $\mathbf{F}^n$ denotes the set of $n+1$ matrices. I or I_n is the unit matrix in $M_n(\mathbf{F})$ and $N\in M_n(\mathbf{F})$ or N_n denote the nilpotent matrix with ε's in the positions $j, j+1$ $(j=1, \dots, n-1)$ and 0's anywhere else.

Let $a_1, a_2, \dots, a_{2n-1}$ be a sequence of elements $a_j\in\mathbf{F}$. The symmetric matrix $S=[s_{jk}]$ with entries $s_{jk}=a_{j+k-1}$ $(j, k=1, \dots, n)$ is called a *Hankel matrix* (cf. [1, pp. 538]). If $a_{n+1}=\dots=a_{2n+1}=0$, then $S=[a_{j+k-1}]$ is an *upper triangular Hankel matrix,* shortly an UH-matrix, which will be denoted by $\bar{S}(a_1, \dots, a_n)$. Similarly, $\underline{S}(a_n, \dots, a_1)$ will denote an LH-matrix (lower triangular Hankel matrix):

$$\bar{S}(a_1, \dots, a_n) = \begin{bmatrix} a_1 & a_2 & \dots & a_n \\ a_2 & & \cdot^{\cdot^{\cdot}} & 0 \\ \vdots & \cdot^{\cdot^{\cdot}} & & \vdots \\ a_n & 0 & \dots & 0 \end{bmatrix}, \quad \underline{S}(a_n, \dots, a_1) = \begin{bmatrix} 0 & \dots & 0 & a_n \\ \vdots & & \cdot^{\cdot^{\cdot}} & \vdots \\ 0 & \cdot^{\cdot^{\cdot}} & & a_2 \\ a_n & \dots & a_2 & a_1 \end{bmatrix}.$$

Obviously, $V\bar{S}V=\underline{S}$ and vice versa. The UH-matrix $\bar{S}$ can be written with help of V and N in the form

$$\bar{S} = \bar{S}(a_1, \dots, a_n) = \left(\sum_{j=0}^{n-1} a_{n-j}N^j\right)V = V\left(\sum_{j=0}^{n-1} a_{n-j}(N^T)^j\right) \tag{2.1}$$

and it is nonsingular iff $a_n\neq 0$.

The triangular matrices $U\in M_n(\mathbf{F})$ of the form $U=\sum_{j=0}^{n-1} b_jN^j$, $b_j\in\mathbf{F}$ are called UT-matrices (upper triangular Toeplitz). Obviously, $VUV=U^T=T$ is an LT-matrix (lower triangular Toeplitz). Let us consider the following LT-matrices $T_j(x)\in \in M_n(\mathbf{F})$

$$T_j(x) = I-x(N^T)^j \tag{2.2}$$

for $j=1, \dots, n-1$ with arbitrary $x\in\mathbf{F}$.

LEMMA 1. *Let* $\bar{S}=\bar{S}(a_1, \dots, a_{n-j}, 0, \dots, 0, a_n)$ *be a nonsingular UH-matrix over a field* $\mathbf{F}$, *char* $\mathbf{F}\neq 0$ *and let* $T_j=T_j((2a_n)^{-1}a_{n-j})$ *be the LH-matrix* (2.2) *with* $x=(2a_n)^{-1}a_{n-j}$. *If* $\bar{S}'$ *denotes the matrix*

$$\bar{S}' = T^T\bar{S}T, \tag{2.3}$$

then $\bar{S}' = \bar{S}(a_1' a_2', \ldots, a_{n-j-1}', 0, \ldots, 0, a_n)$, *i.e.*

(2.4) $$a_{n-j}' = 0, \quad a_{n-v}' = a_{n-v} = 0, \qquad (v = 1, \ldots, j-1), \quad a_n' = a_n.$$

PROOF. Taking into account (2.1) and (2.2) the direct computation of (2.3) results for the parameters a_{n-v}'

$$a_{n-v}' = \begin{cases} a_{n-v} & \text{if } v = 0, \ldots, j-1, \\ 0 & \text{if } v = j, \\ a_{n-v} & \text{if } v = j+1, \ldots, 2j-1, \\ a_{n-2j} - 2(2a_n)^{-1} a_{n-j}^2 + (2a_n)^{-2} a_n^2 a_{n-j}^2 & \text{if } v = 2j, \\ a_{n-v} - 2(2a_n)^{-1} a_{n-j} a_{n-j-v} & \text{if } v = 2j+1, \ldots, 3j-1, \\ a_{n-v} - 2(2a_n)^{-1} a_{n-j} a_{n-j-v} + (2a_n)^{-2} a_n^2 a_{n-2j-v} & \text{if } v = 3j, \ldots, n-1, \end{cases}$$

and this gives for $v=0, \ldots, j$ the stated relations (2.4). □

Let us observe that Lemma 1 makes use of assumption char $\mathbf{F} \neq 2$. The next lemma is an immediate consequence of Lemma 1.

LEMMA 2. *Let $\bar{S}=\bar{S}(a_1, \ldots, a_n)$ be a nonsingular UH-matrix (resp. $\underline{S}=\underline{S}(a_n, \ldots$ $\ldots, a_1)$ a nonsingular LH-matrix) over a field* $\mathbf{F}$, *char* $\mathbf{F} \neq 2$. *Then there is an LT-matrix T (resp. a UT-matrix U) with ε's in the main diagonal such that*

(2.5) $$T^T \bar{S} T = a_n V, \quad \textit{resp.} \quad U^T S U = a_n V.$$

PROOF. It suffices to prove the statement for the UH-matrix $\bar{S}$. Let us define inductively the matrices

$$T_j = T_j((2a_n)^{-1} a_{n-j}^{(j-1)})$$
$$\bar{S}^{(j)} = \bar{S}(a_1^{(j)}, \ldots, a_n^{(j)}) = T_j^T \bar{S}^{(j-1)} T_j$$
$$j = 1, \ldots, n-1; \quad \bar{S}^{(0)} \overset{\text{def}}{=} \bar{S}(a_1, \ldots, a_n).$$

Then by Lemma 1 it holds

(2.6) $$\bar{S}^{(j)} = \bar{S}(a_1^{(j)}, \ldots, a_{n-j-1}^{(j)}, 0, \ldots, 0, a_n) \qquad j = 1, 2, \ldots, n-1.$$

The matrix $T = T_1 T_2 \ldots T_{n-1}$ is again an LH-matrix with ε's in the main diagonal, and according to (2.6) we have

$$\bar{S}^{(n-1)} = T^T \bar{S} T = a_n V,$$

which was to be proved. □

REMARK. Observe that the triangular matrices T and U involved in Lemma 2 have the inverses

(2.7) $$T^{-1} = a_n^{-1} V T^T \bar{S}, \quad U^{-1} = a_n^{-1} V U^T \underline{S}.$$

LEMMA 3. *Let $H \in M_{2r}(\mathbf{F})$ be the Hankel matrix having the partitioned form*

$$H = \operatorname{diag}(V_{2k+1}, -V_{2l+1}) = \begin{bmatrix} V_{2k+1} & 0 \\ 0 & -V_{2l+1} \end{bmatrix},$$

where $k+l+1=r$ $(k, l \geqq 0)$. *If* $\operatorname{char} \mathbf{F} \neq 2$, *then there is a nonsingular matrix* Q *such that*

$$Q^T H Q = 2V.$$

PROOF. Let us consider the permutation matrix P and the blockdiagonal matrix D, that have the partitioned form

$$P = \begin{array}{c} \left[\begin{array}{c|c|c|c|c|c} I_k & & & & & \\ \hline & & \varepsilon & & & \\ \hline & & & & & I_k \\ \hline & I_l & & & & \\ \hline & & & \varepsilon & & \\ \hline & & & & I_l & \end{array}\right] \begin{array}{l} \}k \\ \\ \}k \\ \}l \\ \\ \}l \end{array} \\ \begin{array}{ccccc} k & l & 2 & l & k \end{array} \end{array} \qquad D = \begin{array}{c} \left[\begin{array}{c|c|c|c|c} I_k & & & & \\ \hline & I_l & & & \\ \hline & & \begin{matrix} \varepsilon & \varepsilon \\ -\varepsilon & \varepsilon \end{matrix} & & \\ \hline & & & -2I_l & \\ \hline & & & & 2I_k \end{array}\right] \begin{array}{l} \}k \\ \}l \\ \}2 \\ \}l \\ \}k \end{array} \\ \begin{array}{ccccc} k & l & 2 & l & k \end{array} \end{array}$$

where on the empty spaces stand everywhere 0's.

A direct verification shows the validity of

$$H^{(1)} = P^T H P = \begin{array}{c} \left[\begin{array}{c|c|c|c|c} & & & & V_k \\ \hline & & & -V_l & \\ \hline & & \begin{matrix} \varepsilon & 0 \\ 0 & -\varepsilon \end{matrix} & & \\ \hline & -V_l & & & \\ \hline V_k & & & & \end{array}\right] \begin{array}{l} \}k \\ \}l \\ \}2 \\ \}l \\ \}k \end{array} \\ \begin{array}{ccccc} k & l & 2 & l & k \end{array} \end{array}$$

and

$$D^T H^{(1)} D = 2V,$$

which prove the lemma. □

3. S-symmetric matrices

From relation (1.1) it is obvious that an s-symmetric matrix of size $2k\times 2k$ and $(2k+1)\times(2k+1)$ can be written in the form

$$A_{2k}=\begin{bmatrix} B & CV \\ VD & VB^TV \end{bmatrix},\quad A_{2k+1}=\begin{bmatrix} B & b & CV \\ d^T & c & b^TV \\ VD & Vd & VB^TV \end{bmatrix}, \tag{3.1}$$

where $C^T=C$; $D^T=D$; $B, C, D\in M_k(\mathbf{F})$; $b, d\in\mathbf{F}^k$; $c\in\mathbf{F}$ are arbitrary.

Taking into account form (3.1) of an arbitrary s-symmetric matrix, the following proposition can be directly verified.

PROPOSITION 1. *Let A_{2k} and A_n be s-symmetric matrices and let us consider the $(2k+n)\times(2k+n)$ permutation matrix P having the partitioned form*

$$P=\begin{bmatrix} I_k & 0 & 0 \\ 0 & 0 & I_k \\ 0 & I_n & 0 \end{bmatrix},$$

then the matrix M, where

$$M = P^T \operatorname{diag}(A_{2k}, A_n) P$$

is s-symmetric.

Further on it is always assumed that the field $\mathbf{F}$ has char $\neq 2$.

The next three theorems take the preliminary steps in proving the main result.

THEOREM 1. *Let $\mu(\lambda)\in\mathbf{F}[\lambda]$ be given, where $\mu(\lambda)=\lambda^n-a_1\lambda^{n-1}-\ldots-a_n$ is irreducible over $\mathbf{F}$ and $n\geq 1$. There is an s-symmetric matrix $A\in M_n(\mathbf{F})$ with elementary divisor $\mu(\lambda)$.*

PROOF. If $n=1$, then $A=a_1$ is s-symmetric. If $n>1$ then the irreducibility of $\mu(\lambda)$ implies $a_n\neq 0$. The companion matrix of $\mu(\lambda)$ is

$$F=\begin{bmatrix} 0 & \varepsilon & 0 & \ldots & 0 \\ 0 & 0 & \varepsilon & & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & & & & \varepsilon \\ a_n & a_{n-1} & & \ldots & a_1 \end{bmatrix}. \tag{3.2}$$

Consider the nonsingular LH-matrix $\underline{S}=\underline{S}(\varepsilon, a_n^{-1}a_{n-1}, \ldots, a_n^{-1}a_1)$. According to Lemma 2 there is a UT-matrix U such that

$$U^{-1}=VU^T\underline{S}. \tag{3.3}$$

A direct verification shows the validity of

$$\underline{S}F=F^T\underline{S}. \tag{3.4}$$

We state that $A=U^{-1}FU$ is s-symmetric. Taking into account (3.3) and (3.4), it really holds

$$VA^TV = V(VU^TSFU)^TV = VU^TF^TSU = VU^TSFU = A,$$

which prove the theorem. □

THEOREM 2. *Let* $\mu(\lambda)^q \in \mathbf{F}[\lambda]$ *be given, where* $\mu(\lambda)^q=(\lambda^n-a_1\lambda^{n-1}-\ldots-a_n)^q$, *and* $\mu(\lambda)$ *is irreducible over* $\mathbf{F}$, $q>1$, $n\geqq 1$. *There is an s-symmetric matrix* $A\in M_{nq}(\mathbf{F})$ *with elementary divisor* $\mu(\lambda)^q$.

PROOF. If $n=1$, then to $\mu(\lambda)^q=(\lambda-a_1)^q$ belongs the Jordan matrix

$$J=\begin{bmatrix} a_1 & & \\ \varepsilon & \ddots & \\ & \ddots & \ddots \\ & & \varepsilon \quad a_1 \end{bmatrix}$$

as the classical canonical matrix with elementary divisor $(\lambda-a_1)^q$, and J is s-symmetric. Suppose $n>1$, then $a_n\neq 0$. The classical canonical matrix with the elementary divisor $\mu(\lambda)^q$ is the matrix (cf. [2, p. 72])

$$B=\begin{bmatrix} F & & \\ E_{n1}F & & \\ & \ddots & \ddots \\ & & E_{n1} \quad F \end{bmatrix}, \tag{3.5}$$

where F is the companion matrix of $\mu(\lambda)$ and $E_{n1}\in M_n(\mathbf{F})$ denotes the n, 1^{th} matrix unit, with ε in the position n, 1 and 0's everywhere else. Let us consider now the UH-matrix $\bar{S}=c\left(\sum_{j=0}^{q-1} N^{jn}\right)V$, $c\neq 0$, which has the form

$$\bar{S}=c\begin{bmatrix} V_n & V_n & \ldots & V_n \\ V_n & & \cdot^{\cdot^{\cdot}} & \\ \vdots & \cdot^{\cdot^{\cdot}} & & \\ V_n & & & \end{bmatrix},$$

and let T be the LT-matrix with inverse

$$T^{-1}=c^{-1}VT^T\bar{S}.$$

A direct verification shows that

$$\bar{S}B=B^T\bar{S}.$$

In the same way as before we can see that the matrix

$$A=T^{-1}BT$$

is s-symmetric.

THEOREM 3. *Let* $e_i(\lambda)=\mu_i(\lambda)^{q_i}\in F[\lambda]$ $(i=1,2)$ *be given, where* $\mu_i(\lambda)=\lambda^{n_i}-a_{i1}\lambda^{n_i-1}-\ldots-a_{in_i}$ *is irreducible over* $\mathbf{F}$, *and* $n_iq_i=2k_i+1$ $(k_i\geqq 0)$. *There is an s-symmetric matrix* $A\in M_{2r}(\mathbf{F})$, *(where* $r=k_1+k_2+1$*), with elementary divisors* $e_1(\lambda)$, $e_2(\lambda)$.

PROOF. Let C_i be the classical canonical matrix with elementary divisor $e_i(\lambda)$. C_i may be any of the form (3.2) or (3.5) including their special cases of $F=a\in \mathbf{F}$ or $B=J$, too. Then $C\in M_{2r}(\mathbf{F})$, where

$$C = \operatorname{diag}(C_1, C_2)$$

is the classical canonical matrix with elementary divisors $e_i(\lambda)$ $(i=1, 2)$. On the basis of Lemma 2 there are triangular Hankel matrices S_i and triangular Toeplitz matrices T_i, such that

$$S_iC_i = C_i^TS_i; \quad T_i^TS_iT_i = V_{2k_i+1} \qquad i = 1, 2.$$

Then the matrices

$$S = \operatorname{diag}(S_1, -S_2) \qquad T = \operatorname{diag}(T_1, T_2)$$

will satisfy the relations

$$SC = C^TS; \quad T^TST = \operatorname{diag}(V_{2k_1+1}-V_{2k_2+1}). \tag{3.6}$$

According to Lemma 3 there is a nonsingular matrix Q such that

$$Q^T \operatorname{diag}(V_{2k_1+1}, -V_{2k_2+1})Q = (2\varepsilon)V_{2r} \qquad (r = k_1+k_2+1). \tag{3.7}$$

Combining (3.6) and (3.7) the matrix $U=TQ$ has the inverse

$$U^{-1} = (2\varepsilon)^{-1}VU^TS$$

and the matrix

$$A = U^{-1}CU$$

is s-symmetric. □

Now we are in the position to prove the main theorem of this paper.

THEOREM 4. *There exists an s-symmetric matrix over the field* $\mathbf{F}$, *char* $\mathbf{F}\neq 2$, *with arbitrary prescribed elementary divisors (over* $\mathbf{F}$*).*

PROOF. Let the system of the given elementary divisors

$$e_i(\lambda) = \mu_i(\lambda)^{q_i}: \ \mu_i(\lambda) = \lambda^{n_i}-a_{i1}\lambda^{n_i-1}-\ldots-a_{in_i};$$

$$n_iq_i = 2k_i; \quad k_i \geqq 1; \quad i=1, \ldots, s,$$

$$f_j(\lambda) = \nu_j(\lambda)^{p_j}: \ \nu_j(\lambda) = \lambda^{m_j}-b_{j1}\lambda^{m_j-1}-\ldots-b_{jm_j};$$

$$m_jp_j = 2l_j+1; \quad l_j \geqq 0; \quad j = 1, \ldots, t.$$

According to the preceding theorems we can construct the following s-symmetric matrices

(1) for $i=1, \ldots, s$:

$A_{(i)}M_{2k_i}(\mathbf{F})$ with elementary divisor $e_i(\lambda)$,

(2) for $j=2, 4, \ldots, 2\left[\frac{t}{2}\right]$:

$B_{(j)} \in M_{2r_j}(\mathbf{F})$ with elementary divisors $f_{j-1}(\lambda)$, $f_j(\lambda)$, where $r_j = l_{j-1} + l_j + 1$, and

(3) if $\left[\frac{t}{2}\right] < \frac{t}{2}$:

$C_t M_{2tl+1}(\mathbf{F})$ with elementary divisor $f_t(\lambda)$.

Then the repeated application of Proposition 1 results a matrix $A \in M_d(\mathbf{F})$, where $d = \sum_1^s n_i q_i + \sum_1^t m_j p_j$, with the required properties. □

The next theorem refers to a connection existing between the set of s-symmetric matrices and the set of (ordinary) symmetric matrices in $M_n(\mathbf{F})$.

THEOREM 5. *Let* $\mathbf{F}$ *be a field* char $\mathbf{F} \neq 2$. *If* $-\varepsilon$ *is a square (resp.* $-\varepsilon$ *and* 2ε *are squares) in* $\mathbf{F}$, *then there is a nonsingular matrix* $U \in M_{2k}(\mathbf{F})$ *(resp.* $U \in M_{2k+1}(\mathbf{F})$) *such that the inner automorphism* φ_U: $A \to U^{-1}AU$ *is a one-to-one mapping of the s-symmetric matrices onto the ordinary symmetric matrices in* $M_{2k}(\mathbf{F})$ *(resp. in* $M_{2k+1}(\mathbf{F})$).

PROOF. Let us consider the matrix $Q \in M_n(\mathbf{F})$, where

$$Q = \begin{bmatrix} I_k & -V_k \\ V_k & I_k \end{bmatrix} \quad \text{if} \quad n = 2k, \quad Q = \begin{bmatrix} I_k & 0 & -V_k \\ 0 & \varepsilon & 0 \\ V_k & 0 & I_k \end{bmatrix} \quad \text{if} \quad n = 2k+1,$$

and the diagonal matrix $D \in M_n(\mathbf{F})$, where

$$D = \begin{cases} \operatorname{diag}(I_k, -I_k) & \text{if} \quad n = 2k \\ \operatorname{diag}(I_k, 2\varepsilon, -I_k) & \text{if} \quad n = 2k+1. \end{cases}$$

The following decomposition of the permutation matrix V is valid

(2.8) $$V = (2\varepsilon)^{-1} QDQ^T.$$

The assumption that $-\varepsilon$ is a square (resp. $-\varepsilon$ and 2ε are squares) in $\mathbf{F}$ implies that the diagonal matrix D has a diagonal square root $\Delta \in M_{2k}(\mathbf{F})$, (resp. $\Delta \in M_{2k+1}(\mathbf{F})$). Using the notation $U = Q\Delta$, we have from (2.8) $V = (2\varepsilon)^{-1} UU^T$, hence the nonsingular matrix U has the inverse

$$U^{-1} = (2\varepsilon)^{-1} U^T V.$$

We show that the inner automorphism φ_U with this U possesses the property stated in the theorem. Really, let A be s-symmetric, then $B = U^{-1}AU$ is symmetric, since

$$B^T = (2\varepsilon)^{-1}(U^T VAU)^T = (2\varepsilon)^{-1} U^T VAVVU = (2\varepsilon) U^T VAU = B.$$

The converse, that each symmetric matrix B is an φ_U-image of an s-symmetric matrix, can be proved similarly.

From the preceding two theorems we have the

COROLLARY. *Let* $\mathbf{F}$ *be a field* char $\mathbf{F} \neq 2$. *There is a symmetric matrix in* $M_{2k}(\mathbf{F})$ *(resp. in* $M_{2k+1}(\mathbf{F})$) *with arbitrary prescribed elementary divisors (over* $\mathbf{F}$), *if* $-\varepsilon$ *is a square (resp.* $-\varepsilon$ *and* 2ε *are squares) in* $\mathbf{F}$.

REFERENCES

[1] Gantmacher, F. R., *The theory of matrices,* Vol. I, Chelsea Publishing Co., New York, 1959. *MR* **21** # 6372c.

[2] Jacobson, N., *Lectures in abstract algebra,* Vol. II, Linear Algebra, D. van Nostrand, Co., Inc., Toronto—New York—London, 1953. *MR* **14**—837.

[3] Lee, A., Secondary symmetric, skewsymmetric and orthogonal matrices, *Period. Math. Hungar.* **7** (1976), 63—70. *MR* **55** # 8070a.

(Received February 18, 1982)

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

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 101—111

CONSTRUCTIONS OF VARIETIES THAT SATISFY THE AMALGAMATION PROPERTY OR THE CONGRUENCE EXTENSION PROPERTY

F. J. PASTIJN

To the memory of Professor L. Rédei

Given a suitable variety K, we shall investigate two techniques for constructing a variety that satisfies the amalgamation property or the congruence extension property. The first technique consists in constructing the inflations of the universal algebras that belong to K, the second one consists in performing the Płonka sums of the semilattice ordered systems of universal algebras that belong to K.

Besides the concepts which will be defined below, we shall essentially use the notations and the terminology of [4].

We shall only deal with algebras without nullary fundamental operations. Let K be a variety (=equational class) of algebras. Then $E(K)$ will denote the set of all identities which are satisfied in all the algebras that belong to K. An identity $\mathbf{p}=\mathbf{q}$ of type τ will be called *regular* if the same variables occur in both $\mathbf{p}$ and $\mathbf{q}$. We denote by $R(K)$ the set of all regular identities which are satisfied in all the algebras that belong to K. Let us consider the identity

$$\mathbf{x}_l = \mathbf{p}(\mathbf{x}_1, \ldots, \mathbf{x}_n), \qquad 1 \leqq l \leqq n \tag{1}$$

of type τ. If in the right-hand side of (1) $\mathbf{x}_l$ and at least one more variable occurs, then we call (1) a *strongly non-regular identity*. Groups, lattices, rings, ..., each satisfy some strongly non-regular identity. If E is a set of identities of some fixed type τ, then K_E denotes the variety that consists of the algebras that satisfy all the identities in the set E.

Let K be a variety. Let $\mathfrak{A}=\langle A; F\rangle$, $\mathfrak{B}=\langle B; F\rangle$ and $\mathfrak{C}=\langle C; F\rangle$ be algebras in K, and let $\beta: A\to B$ and $\gamma: A\to C$ be monomorphisms. Consider the amalgam $\langle\mathfrak{A}, \beta, \mathfrak{B}, \gamma, \mathfrak{C}\rangle$ in K. This amalgam is *weakly embeddable* in K if there exists an algebra $\mathfrak{D}=\langle D; F\rangle$ in K, and monomorphisms $\beta': B\to D$, $\gamma': C\to D$, such that $\beta\beta'=\gamma\gamma'$. The amalgam will be called *strongly embeddable* in K, if there exist $\mathfrak{D}$, β' and γ' satisfying the above, and in addition $A\beta\beta'=A\gamma\gamma'=B\beta'\cap C\gamma'$. The variety K satisfies WAP (the *weak amalgamation property*) if every amalgam in K is weakly embeddable in K, and the variety K satisfies SAP (the *strong amalgamation property*) if every amalgam in K is strongly embeddable in K. We refer to [2], [4], [6], [7] for examples of varieties that satisfy WAP or SAP.

A variety K satisfies CEP (the *congruence extension property*) if every congruence relation on a subalgebra of an algebra $\mathfrak{A}$ in K can be extended to a congruence relation on $\mathfrak{A}$. We refer to [1], [4], [7] for more information and for examples.

1980 *Mathematics Subject Classification.* Primary 08B05, 08B25; Secondary 20M07.
Key words and phrases. Amalgamation property, congruence extension property, Płonka sums.

Let $\mathfrak{A}=\langle A; F\rangle$ be an algebra of type τ, and let $\{X_a|a\in A\}$ be a set of mutually disjoint sets, such that for $a\in A$, we have $X_a\cap A=\{a\}$. Let $B=\bigcup_{a\in A} X_a$. For any $f_\gamma\in F$, with $\gamma<o(\tau)$ and $n_\gamma=\tau(\gamma)$, and any $b_i\in X_{a_i}$, $i=1, \ldots, n_\gamma$, we define

$$f_\gamma(b_1, \ldots, b_{n_\gamma})=f_\gamma(a_1, \ldots, a_{n_\gamma}).$$

Then $\mathfrak{B}=\langle B; F\rangle$ becomes an algebra of type τ which will be called an *inflation* of the algebra $\mathfrak{A}$.

Let $Z=\{0, \infty\}$, and let τ be a type. Let $F=\langle f_\gamma, \gamma<o(\tau)\rangle$, and consider $\mathfrak{Z}=\langle Z; F\rangle$, where $f_\gamma(a_1, \ldots, a_{n_\gamma})=0$ for all $\gamma<o(\tau)$ and all $a_1, \ldots, a_{n_\gamma}\in Z$. Remark that $\mathfrak{Z}$ satisfies all the identities of type τ, except the identities which are of the form (1) where $\mathbf{p}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ is a polynomial symbol which is different from $\mathbf{x}_l$. We call $\mathfrak{Z}$ the two-element *zero algebra* of type τ.

We refer to [2] for special cases of Theorems 1 and 2 below. For Theorem 1 we refer to [5] and [14]. We include a proof for the sake of completeness.

THEOREM 1. *Let K be a variety consisting of algebras of type τ. Then the following are equivalent.*

(i) *$\bar{K}$ is the join of K and the variety* HSP($\mathfrak{Z}$), *where $\mathfrak{Z}$ is the two-element zero algebra of type τ.*

(ii) *$\bar{K}$ is the variety that consists of the algebras of type τ which satisfy all the identities in $E(K)$, except perhaps the identities* (1) *where $\mathbf{p}(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ is a polynomial symbol which is different from $\mathbf{x}_l$.*

(iii) *$\bar{K}$ consists of the algebras which are inflations of the algebras in K.*

(iv) *An algebra belongs to $\bar{K}$ if and only if it is a subdirect product of an algebra in K and an algebra in* HSP($\mathfrak{Z}$).

PROOF. The equivalence of (i) and (ii) follows from the observation made before the statement of the theorem. If $E(K)$ does not contain any non-trivial identity of the form (1), then $K=\bar{K}$ in the four statements (i), (ii), (iii), (iv), and the equivalence prevails. Therefore we shall henceforth suppose that $E(K)$ contains an identity (1), where $\mathbf{p}(\mathbf{x}_1, \ldots, \mathbf{x}_k)$ is a polynomial symbol which is different from $\mathbf{x}_k$ in $\mathbf{P}(\tau)$, for $k=1, \ldots, n$.

Let $\mathfrak{B}=\langle B; F\rangle$ be an inflation of $\mathfrak{A}=\langle A; F\rangle\in K$. Then $B=\bigcup_{a\in A} X_a$, where

(2) $$X_a\cap X_{a'}=\Box \quad \text{if} \quad a\neq a'$$

and

(3) $$X_a\cap A=\{a\}$$

for all $a, a'\in A$, and where

(4) $$f_\gamma(a_1, \ldots, a_{n_\gamma})=f_\gamma(b_1, \ldots, b_{n_\gamma})$$

for every $f_\gamma\in F$ and $b_i\in X_{a_i}$, $i=1, \ldots, n_\gamma$. Let

$$\varrho_1=\bigcup_{a\in A} X_a\times X_a,$$

$$\varrho_2=(A\times A)\cup\{(x, x)|x\in B\}.$$

Then ϱ_1 and ϱ_2 are congruence relations on $\mathfrak{B}$. Since $\varrho_1\cap\varrho_2=\iota_B$, we see that $\mathfrak{B}$ is a subdirect product of $\mathfrak{B}/\varrho_1$ and $\mathfrak{B}/\varrho_2$. Further, $\mathfrak{B}/\varrho_1\cong\mathfrak{A}\in K$ and $\mathfrak{B}/\varrho_2\in$ HSP($\mathfrak{Z}$).

Let us consider an algebra $\mathfrak{B}=\langle B; F\rangle$ which is a subdirect product of an algebra $\mathfrak{A}_1=\langle A; F\rangle\in K$ and an algebra $\mathfrak{A}_2=\langle A_2; F\rangle\in \mathrm{HSP}(\mathfrak{Z})$. Then there exists an element o in A_2 such that $o=f_\gamma(b_1, \ldots, b_{n_\gamma})$ for every $\gamma<o(\tau)$ and every $b_1, \ldots, b_{n_\gamma}\in$ $\in A_2$. Let a be any element in A_1. Then there exists an element $b\in A_2$ such that $(a, b)\in B$. Let p be the polynomial which is induced by the polynomial symbol $\mathbf{p}$ which appears in (1). Then $p((a, b), \ldots, (a, b))=(a, o)\in B$, and we may conclude that $\mathfrak{A}=\langle A; F\rangle$, with $A=\{(a, o)|a\in A_1\}$, is a subalgebra of $\mathfrak{B}$ which is isomorphic to $\mathfrak{A}_1$. Thus we also have $\mathfrak{A}\in K$. Putting $X_{(a,0)}=\{(a, b)|b\in A_2, (a, b)\in B\}$, we see that $\mathfrak{B}$ is an inflation of $\mathfrak{A}$. We proved the equivalence of (iii) and (iv).

In view of the equivalence of (iii) and (iv), we can now say that in order to show the equivalence of (i) and (iii), it suffices to show that the class $\bar{K}$ which is defined by (iii) constitutes a variety. Let I be a set, and for each $i\in I$, let $\mathfrak{B}_i$ be an inflation of $\mathfrak{A}_i\in K$. Then the direct product $\Pi(\mathfrak{B}_i|i\in I)$ is an inflation of $\Pi(\mathfrak{A}_i|i\in I)\in K$. Thus the class $\bar{K}$ which is defined by (iii) is closed for taking direct products. Let $\mathfrak{B}=$ $=\langle B; F\rangle$ be an inflation of $\mathfrak{A}=\langle A; F\rangle\in K$, where (2), (3) and (4) hold. Let $\mathfrak{C}=$ $=\langle C; F\rangle$ be a subalgebra of $\mathfrak{B}$. Clearly $\mathfrak{A}\cap\mathfrak{C}=\langle A\cap C; F\rangle$ is a subalgebra of $\mathfrak{A}$, and so $\mathfrak{A}\cap\mathfrak{C}\in K$. Let us put $Y_a=X_a\cap C$ for $a\in A\cap C$. Let $q\in C$. Then $q\in X_a$ for some $a\in A$. Since (1) holds in K, we have $a=p(q, \ldots, q)\in A\cap C$. Thus $q\in Y_a$, with $a\in A\cap C$. We conclude that C is the disjoint union of the sets Y_a, $a\in A\cap C$, where for every $a\in A\cap C$, we have $Y_a\cap(A\cap C)=\{a\}$. It is now easy to see that $\mathfrak{C}$ is an inflation of $\mathfrak{A}\cap\mathfrak{C}\in K$. Thus the class $\bar{K}$ defined by (iii) is closed for taking subalgebras. Let us again consider the above algebra $\mathfrak{B}$, and let φ: $B\to E$ be a homomorphism of $\mathfrak{B}$ onto $\mathfrak{E}=\langle E; F\rangle$. The image $\mathfrak{D}=\langle D; F\rangle$ of $\mathfrak{A}$ under φ is a subalgebra of $\mathfrak{E}$ and $\mathfrak{D}\in K$. Let $Z_{a\varphi}=X_a\varphi$ for every $a\in A$. Every element of $Z_{a\varphi}$ is of the form $q\varphi$, with $q\in X_a$. But then $a\varphi=(p(q, \ldots, q))\varphi=p(q\varphi, \ldots, q\varphi)$. This shows that E is the disjoint union of the sets Z_d, $d\in D$, where $Z_d\cap D=\{d\}$ for every $d\in D$. By this partitioning $\mathfrak{E}$ becomes an inflation of $\mathfrak{D}\in K$. Consequently, the class $\bar{K}$ which is defined by (iii) is closed for taking homomorphic images. We may now conclude that this class $\bar{K}$ constitutes a variety.

THEOREM 2. *Let K and $\bar{K}$ be as in Theorem 1. Then K satisfies WAP [SAP, CEP] if and only if $\bar{K}$ satisfies WAP [SAP, CEP].*

PROOF. By the foregoing we know that $K=\bar{K}$ in case $E(K)$ does not contain any non-trivial identity of the form (1). In this case our theorem holds trivially. We suppose henceforth that K satisfies (1), where $\mathbf{p}(\mathbf{x}_1, \ldots, \mathbf{x}_n)\neq\mathbf{x}_k$, $1\leq k\leq n$, in $\mathbf{P}(\tau)$.

For any $\mathfrak{X}'=\langle X; F\rangle$ in $\bar{K}$, we note that $\mathfrak{X}'$ is the inflation of $\mathfrak{X}=\langle X; F\rangle\in K$, which is the image of $\mathfrak{X}'$ under the idempotent endomorphism

$$\theta\colon\ X'\to X,\ x'\to p(x', \ldots, x'),$$

of $\mathfrak{X}'$ onto $\mathfrak{X}$. X' is the disjoint union of the $\theta\theta^{-1}$-classes $X_x=x\theta\theta^{-1}$, $x\in X$.

Let $\langle\mathfrak{A}', \beta', \mathfrak{B}', \gamma', \mathfrak{C}'\rangle$ be an amalgam in $\bar{K}$. Without loss of generality we may suppose that $\mathfrak{A}'=\langle A'; F\rangle$, $\mathfrak{B}'=\langle B'; F\rangle$, $\mathfrak{C}'=\langle C'; F\rangle$, $A'=B'\cap C'$, where β' and γ' are inclusion mappings. Then $\langle\mathfrak{A}, \beta'|A, \mathfrak{B}, \gamma'|A, \mathfrak{C}\rangle$ is an amalgam in K where $\beta'|A$ and $\gamma'|A$ are inclusion mappings, and $A=B\cap C=\{a\in A'=B'\cap C'|a=p(a, \ldots$ $\ldots, a)\}$. If K satisfies WAP, then there exists an algebra $\mathfrak{D}=\langle D; F\rangle$ in K, and monomorphisms β: $B\to D$, γ: $C\to D$ of $\mathfrak{B}$ and $\mathfrak{C}$ into $\mathfrak{D}$ which agree on A. We can easily construct an inflation $\mathfrak{D}'$ of $\mathfrak{D}$ such that β [γ] extends to a monomorphism

β'' [γ''] of $\mathfrak{B}'$ [$\mathfrak{C}'$] into $\mathfrak{D}'$, where β'' and γ'' agree on A'. If the amalgam $\langle \mathfrak{A}, \beta|A, \mathfrak{B}, \gamma|A, \mathfrak{C}\rangle$ is strongly embedded in $\mathfrak{D}$, then we can manage to construct $\mathfrak{D}'$ in such a way that the amalgam $\langle \mathfrak{A}', \beta', \mathfrak{B}', \gamma', \mathfrak{C}'\rangle$ is strongly embedded in $\mathfrak{D}'$ by β'' and γ''. A precise construction is left as an easy exercise to the reader. Thus, if K satisfies WAP [SAP], then $\bar{K}$ satisfies WAP [SAP].

Let us suppose that $\langle \mathfrak{A}, \beta, \mathfrak{B}, \gamma, \mathfrak{C}\rangle$ be an amalgam in K, and let $\mathfrak{D}'=\langle D'; F\rangle$ be an algebra in $\bar{K}$, and $\beta'\colon B\to D'$, $\gamma'\colon C\to D'$ monomorphisms such that $\beta\beta'=\gamma\gamma'$. Let $\theta\colon D'\to D$, $x\to p(x, \dots, x)$ be the idempotent endomorphism considered above: θ maps $\mathfrak{D}'$ homomorphically onto the subalgebra $\mathfrak{D}=\langle D; F\rangle$ of $\mathfrak{D}'$. We note that $\mathfrak{D}\in K$, and that $\beta'\theta$ and $\gamma'\theta$ are monomorphisms such that $\beta(\beta'\theta)=\gamma(\gamma'\theta)$. Hence the above amalgam is weakly embedded in $\mathfrak{D}\in K$ by the monomorphisms $\beta'\theta$ and $\gamma'\theta$. If the amalgam is strongly embedded in $\mathfrak{D}'\in\bar{K}$ by β' and γ', then the amalgam is strongly embedded in $\mathfrak{D}\in K$ by $\beta'\theta$ and $\gamma'\theta$. Thus, if $\bar{K}$ satisfies WAP [SAP], then K satisfies WAP [SAP].

The proof for the corresponding statements involving CEP is straightforward

Let $L=\{0, 1\}$ and let τ be a type. Let $F=\langle f_\gamma, \gamma<o(\tau)\rangle$, and consider $\mathfrak{L}=\langle L; F\rangle$ where

$$f_\gamma(a_1, \dots, a_{n_\gamma}) = a \quad \text{whenever} \quad a_1 = a_2 = \dots = a_{n_\gamma} = a$$

and

$$f_\gamma(a_1, \dots, a_{n_\gamma}) = 0 \quad \text{if} \quad 0\in\{a_1, \dots, a_{n_\gamma}\}.$$

$\mathfrak{L}$ satisfies the identity $\mathbf{p}=\mathbf{q}$ of type τ if and only if $\mathbf{p}=\mathbf{q}$ is a regular identity. We call $\mathfrak{L}$ the two-element *semilattice of type* τ.

Let us consider the variety K consisting of algebras of type τ. Recall that a semilattice ordered system of algebras from K

$$\mathscr{A} = \langle\langle I; \wedge\rangle;\ (\mathfrak{A}_i|i\in I);\ (\varphi_{ii'}|i, i'\in I, i'\leqq i)\rangle \tag{5}$$

is just a direct family of algebras in the sense of Definition 3 of § 21 of [4], where the underlying partially ordered set is a semilattice $\langle I; \wedge\rangle$, and where the carriers A_i of the algebras $\mathfrak{A}_i$, $i\in I$, are mutually disjoint. The *Płonka sum* $S(\mathscr{A})$ of the system $\mathscr{A}$ is the algebra of type τ with carrier $\bigcup_{i\in I} A_i$, and where for every $\gamma<o(\tau)$, and $a_j\in A_{i_j}$, $j=1, \dots, n_\gamma$, we have

$$f_\gamma(a_1, \dots, a_{n_\gamma}) = f(a_1\varphi_{i_1 i_0}, \dots, a_{n_\gamma}\varphi_{i_{n_\gamma} i_0}) \tag{6}$$

with $i_0=\bigwedge_{j=1}^{n_\gamma} i_j$.

If the algebras of the variety K all satisfy some strongly non-regular identity (1), then the class consisting of the algebras that are Płonka sums of semilattice ordered systems of algebras from K constitutes a variety. If this is the case, then the variety under consideration is exactly $K_{R(K)}$, and this variety $K_{R(K)}$ is the join of K and HSP($\mathfrak{L}$) [3], [10], [13].

Special cases for part of the following theorem may be found in [2], [6], [7].

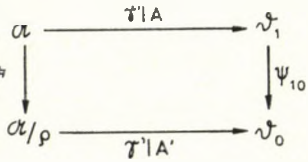
THEOREM 3. *Let K be a variety, where $E(K)$ contains a strongly non-regular identity* (1). *Then the variety $K_{R(K)}$ satisfies WAP [SAP] if and only if K satisfies WAP [SAP] and CEP.*

PROOF. Let us first suppose that $K_{R(K)}$ satisfies WAP. Let $\langle\mathfrak{A}, \beta, \mathfrak{B}, \gamma, \mathfrak{C}\rangle$ be an amalgam in K. This amalgam can be weakly embedded in an algebra $\mathfrak{D}=\langle D; F\rangle$ of $K_{R(K)}$ be the monomorphisms β': $B\to D, \gamma'$: $C\to D$. We know that $\mathfrak{D}$ is the Płonka sum of a semilattice ordered system of algebras $\mathfrak{D}_i=\langle D_i; F\rangle$, $i\in I$, of K. Since the subalgebras $\mathfrak{B}\beta'$ and $\mathfrak{C}\gamma'$ of $\mathfrak{D}$ satisfy the strongly non-regular identity (1) we have $B\beta'\subseteq D_{i_1}$ and $C\gamma'\subseteq D_{i_2}$ for some $i_1, i_2\in I$. Further, $\square\neq B\beta'\cap C\gamma'\subseteq\subseteq D_{i_1}\cap D_{i_2}$ implies $i_1=i_2$. It follows that the above amalgam is weakly embedded in $\mathfrak{D}_i=\langle D_i; F\rangle\in K$ by β' and γ'. We conclude that K satisfies WAP. Obviously, if $K_{R(K)}$ satisfies SAP, then K also satisfies SAP.

Let $\mathfrak{A}=\langle A; F\rangle$ be a subalgebra of the algebra $\mathfrak{B}=\langle B; F\rangle\in K$, and let ϱ be a congruence on $\mathfrak{A}$. The canonical homomorphism of $\mathfrak{A}$ onto $\mathfrak{A}/\varrho=\langle A'; F\rangle$ will be denoted by $\varrho^\natural$. We now consider the semilattice ordered system $\mathscr{A}$ given by (5), where $I=\{1, 0\}$, $\mathfrak{A}_1=\mathfrak{A}$, $\mathfrak{A}_0=\mathfrak{A}/\varrho$, and $\varphi_{10}=\varrho^\natural$. We can and shall suppose that $B\cap A'=\square$. Then $\langle\mathfrak{A}, \beta, \mathfrak{B}, \gamma, S(\mathscr{A})\rangle$ is an amalgam in $K_{R(K)}$ where β and γ are inclusion mappings, and where $B\cap(A\cup A')=A$. This amalgam is weakly embeddable in an algebra $\mathfrak{D}=\langle D; F\rangle\in K_{R(K)}$ by the monomorphisms β' and γ'. We can always suppose that $\mathfrak{D}$ is generated by the elements of $B\beta'\cup A'\gamma'$. Consequently $\mathfrak{D}$ is the sum of a semilattice ordered system

$$(7)\qquad \langle\langle I; \wedge\rangle; (\mathfrak{D}_i|i\in I); (\psi_{ii'}, i, i'\in I, i'\leqq i)\rangle$$

of algebras $\mathfrak{D}_i=\langle D_i; F\rangle\in K$, $i\in I=\{1, 0\}$, where β' maps $\mathfrak{B}$ onto $\mathfrak{D}_1$ and γ' maps $\mathfrak{A}/\varrho$ into $\mathfrak{D}_0$. Further, the embedding of the above amalgam is necessarily a strong embedding. This all follows from the fact that the semilattice ordered system (7) which is associated with $\mathfrak{D}$ is uniquely determined by $\mathfrak{D}$ [10], [13]. Note that ψ_{10}: $D_1\to D_0$ is given by the following. If $d\in D_1$, then $d\psi_{10}=p(d_1, \ldots, d_n)$ with $d_l=d$, and $d_k\in D_0$ for all $1\leqq k\leqq n$, $k\neq l$. Thus



is commutative, and ϱ is the congruence relation on $\mathfrak{A}$ which is induced by $(\gamma'|A)\psi_{10}$. Let ϱ' be the congruence relation on $\mathfrak{B}$ which is induced by $\beta'\psi_{10}$. Since $\beta'|A=\gamma'|A$, the above implies that $\varrho'\cap A\times A=\varrho$. Thus the congruence relation ϱ on $\mathfrak{A}$ can be extended to a congruence ϱ' on $\mathfrak{B}$. We conclude that K satisfies CEP.

We now proceed to show the converse part of the theorem. Accordingly, let us suppose that K satisfies WAP and CEP. Let $\langle\mathfrak{A}, \beta, \mathfrak{B}, \gamma, \mathfrak{C}\rangle$ be an amalgam in $K_{R(K)}$. There is no loss of generality if we suppose that $\mathfrak{A}=\langle A; F\rangle$, $\mathfrak{B}=\langle B; F\rangle$, $\mathfrak{C}=\langle C; F\rangle$, $B\cap C=A$, where β and γ are inclusion mappings. Let $\mathfrak{A}$ be the Płonka sum of the system (5), where $\mathfrak{A}_i\in K$ for all $i\in I$. Let $\mathfrak{B}$ and $\mathfrak{C}$ be the Płonka sums of the semilattice ordered systems of algebras from K

$$(8)\qquad \langle\langle J; \wedge\rangle; (\mathfrak{B}_j|j\in J); (\psi_{jj'}|j, j'\in J, j'\leqq j)\rangle$$

and

$$(9)\qquad \langle\langle M; \wedge\rangle; (\mathfrak{C}_m|m\in M); (\zeta_{mm'}|m, m'\in M, m'\leqq m)\rangle,$$

respectively. The above considered decompositons of $\mathfrak{A}$, $\mathfrak{B}$ and $\mathfrak{C}$ as sums of systems of algebras from K is unique: the $\mathfrak{A}_i$, $i\in I$, are the maximal subalgebras of $\mathfrak{A}$ that belong to K, and a similar statement holds for the $\mathfrak{B}_j$, $j\in J$, and the $\mathfrak{C}_m$, $m\in M$.

The mapping $\bar{\beta}$: $I\to J$ which is given by $A_i\subseteq B_{i\bar{\beta}}$ for all $i\in I$, is well-defined and gives rise to a monomorphism of the semilattice $\mathfrak{I}=\langle I;\wedge\rangle$ into the semilattice $\mathfrak{J}=\langle J;\wedge\rangle$. Analogously, the mapping $\bar{\gamma}$: $I\to M$ which is given by $A_i\subseteq C_{i\bar{\gamma}}$ for all $i\in I$, is a well-defined monomorphism of $\mathfrak{I}$ into the semilattice $\mathfrak{M}=\langle M;\wedge\rangle$. As a result we have the amalgam $\langle\mathfrak{I},\bar{\beta},\mathfrak{J},\bar{\gamma},\mathfrak{M}\rangle$ of semilattices. Without loss of generality, we may suppose that $I=J\cap M$, and that $\bar{\beta}$ and $\bar{\gamma}$ are inclusion mappings. This semilattice amalgam can be embedded strongly into a semilattice $\mathfrak{H}=\langle H;\wedge\rangle$ by the monomorphisms $\bar{\beta}'$: $J\to H$ and $\bar{\gamma}'$: $M\to H$. For the sake of simplicity we shall again suppose that β' and γ' are inclusion mappings, and that $\mathfrak{H}$ is generated by the elements of $J\cup M$ (thus, $H=J\cup M\cup(J\wedge M)$). The embedding of the above amalgam can be done in such a way that for $j\in J$ and $m\in M$ we have $j\leqq m$ [$m\leqq j$] in $\mathfrak{H}$ if and only if there exists $i\in I$ such that $j\leqq i$ in $\mathfrak{J}$ and $i\leqq m$ in $\mathfrak{M}$ [$m\leqq i$ in $\mathfrak{M}$ and $i\leqq j$ in $\mathfrak{J}$].

If $i,i'\in I$ and $i'\leqq i$ in $\mathfrak{I}$, then $i,i'\in J\cap M$, $i'\leqq i$ in $\mathfrak{J}$ and in $\mathfrak{M}$, and $A_i=B_i\cap C_i$, $A_{i'}=B_{i'}\cap C_{i'}$. Further, if $a\in A_i$, then $a\zeta_{ii'}=a\varphi_{ii'}=a\psi_{ii'}=p(a_1,\dots,a_n)$, where $a_l=a$ and $a_k\in A_{i'}$ for $k\neq l$, $k\in\{1,\dots,n\}$. Hence

(10) $$\psi_{ii'}|A_i=\zeta_{ii'}|A_i=\varphi_{ii'},\quad i,i'\in I,\quad i'\leqq i\quad\text{in}\quad\mathfrak{I}.$$

For any $h\in H$, let

$$\theta_h\colon\ (\bigcup_{j\geqq h}B_j)\cup(\bigcup_{m\geqq h}C_m)\to Z_h$$

be a bijection, and let $\mathfrak{V}_h=\langle V_h;F\rangle$ be the algebra of K which is freely generated by the elements of Z_h subject to the identities of $E(K)$, and subject to

(11) $$b_j\theta_h=(b_j\psi_{jj'})\theta_h\quad\text{for all}\quad h\leqq j'\leqq j\quad\text{in}\quad\mathfrak{H},\quad j,j'\in J,\quad\text{and all}\quad b_j\in B_j,$$

(12) $$c_m\theta_h=(c_m\zeta_{mm'})\theta_h\quad\text{for all}\quad h\leqq m'\leqq m\quad\text{in}\quad\mathfrak{H},\quad m,m'\in M,\quad\text{and all}\quad c_m\in C_m,$$

(13) $$f(x_1\theta_h,\dots,x_p\theta_h)=g(y_1\theta_h,\dots,y_q\theta_h)\quad\text{for all}\quad x_1,\dots,x_p,y_1,\dots,y_q\in B_j\quad\text{with}\quad j\in J,\ \text{and}\ j\geqq h\ \text{in}\ \mathfrak{H},\ \text{if}\quad f(x_1,\dots,x_p)=g(y_1,\dots,y_q)\ \text{holds in}\ \mathfrak{B}_j,$$

(14) $$f(x_1\theta_h,\dots,x_p\theta_h)=g(y_1\theta_h,\dots,y_q\theta_h)\quad\text{for all}\quad x_1,\dots,x_p,y_1,\dots,y_q\in C_m\quad\text{with}\quad m\in M,\ \text{and}\ m\geqq h\ \text{in}\ \mathfrak{H},\ \text{if}\quad f(x_1,\dots,x_p)=g(y_1,\dots,y_q)\ \text{holds in}\ \mathfrak{C}_m.$$

We shall suppose that the carriers V_h, $h\in H$, are pairwise disjoint, and we consider the semilattice ordered system

(15) $$\langle\langle H; \wedge\rangle; (\mathfrak{B}_h | \in H); (\eta_{hh'} | h, h'\in H, h' \leqq h)\rangle$$

where for each $h, h'\in H, h'\leqq h$ in $\mathfrak{H}$, $\eta_{hh'}$ is the homomorphism of $\mathfrak{B}_h$ into $\mathfrak{B}_{h'}$ which extends the mapping $Z_h\to Z_{h'}$, $x\theta_h\to x\theta_{h'}$. Let $\mathfrak{B}$ be the Płonka sum of the system (15).

Let $h\in J$. Then $\theta_h|B_h$: $B_h\to V_h$ is a homomorphism of $\mathfrak{B}_h$ into $\mathfrak{B}_h$. If $b_1, b_2\in B_h$, and $b_1\theta_h = b_2\theta_h$ in $\mathfrak{B}_h$, then there exists a finite proof which establishes the equality of $b_1\theta_h$ and $b_2\theta_h$. This proof is based on the rules (11), (12), (13), (14) and the identities of $E(K)$. In other words, there exists a sequence

(16) $$b_1\theta_h = p_1(x_1, \dots, x_t), \dots, p_s(x_1, \dots, x_t) = b_2\theta_h$$

where $p_1, \dots, p_s$ are polynomials of $\mathfrak{B}_h$, $x_1, \dots, x_t\in Z_h$, and where the transition of one of the elements in (16) to the subsequent one is accomplished by a single application of one of the rules (11), (12), (13), (14) or an identity in $E(K)$. Let N be the subset of M which consists of the elements n for which $c_n\theta_h\in\{x_1, \dots, x_t\}$ for some $c_n\in C_n$. If $N=\square$, then $b_1=b_2$ in B_h. We shall henceforth suppose $N\neq\square$. Let $m_0=\wedge N$ in the semilattice $\mathfrak{M}$. Since $m_0\in M$, $h\in J$ and $h\leqq m_0$, there exists an $i\in I$ such that $h\leqq i\leqq m_0$ in $\mathfrak{H}$. For $k=1, \dots, t$, let

$$x_k' \begin{cases} =(b_j\psi_{jh})\theta_h & \text{if } b_j\in B_j,\ b_j\theta_h = x_k \text{ and } i\not\leqq j \text{ in } \mathfrak{J}, \\ =(b_j\psi_{ji})\theta_h & \text{if } b_j\in B_j,\ b_j\theta_h = x_k \text{ and } i\leqq j \text{ in } \mathfrak{J}, \\ =(c_n\zeta_{ni})\theta_h & \text{if } c_n\in C_n,\ c_n\theta_h = x_h. \end{cases}$$

The above elements are well-defined because of (10). Then the sequence

(17) $$b_1\theta_h = p_1(x_1', \dots, x_t'), \dots, p_s(x_1', \dots, x_t') = b_2\theta_h$$

also constitutes a proof for the equality of $b_1\theta_h$ and $b_2\theta_h$ in $\mathfrak{B}_h$. This proof involves elements of $(B_i\cup C_i\cup B_h)\theta_h$ only. Again every transition in (17) is an application of one of the rules (11), (12), (13), (14) or an identity of $E(K)$.

The congruence relation ϱ on $\mathfrak{A}_i$ which is induced by the homomorphism $\psi_{ih}|A_i$ can be extended to a congruence relation $\bar\varrho$ on $\mathfrak{C}_i$ since K satisfies CEP. Let us consider the amalgam $\langle\mathfrak{A}_i/\varrho, \iota, \mathfrak{C}_i/\bar\varrho, \eta, \mathfrak{B}_h\rangle$, where ι is the inclusion mapping, and η the monomorphism of $\mathfrak{A}_i/\varrho$ into $\mathfrak{B}_h$ which is defined by $(a\varrho^\natural)\eta = a\psi_{ih}$ for all $a\in A_i$. Since K satisfies WAP, the above amalgam can be embedded in an algebra $\mathfrak{D}=\langle D; F\rangle\in K$ by the monomorphisms $\bar\iota$ and $\bar\eta$. For $k=1, \dots, t$, let

$$x_k'' \begin{cases} = x_k'\theta_h^{-1}\bar\varrho^\natural\bar\iota & \text{if } x_k'\in C_i\theta_h, \\ = x_k'\theta_h^{-1}\psi_{ih}\bar\eta & \text{if } x_k'\in B_i\theta_h, \\ = x_k'\theta_h^{-1}\bar\eta & \text{if } x_k'\in B_h\theta_h. \end{cases}$$

Then

(18) $$b_1\bar\eta = p_1(x_1'', \dots, x_t'') = \dots = p_s(x_1'', \dots, x_t'') = b_2\bar\eta$$

holds in $\mathfrak{D}$. Thus $b_1=b_2$ in B_h. We conclude that $\theta_h|B_h$ is a monomorphism of $\mathfrak{B}_h$ into $\mathfrak{B}_h$.

In a similar way we can show that for every $h \in M$, $\theta_h | C_h$ is a monomorphism of $\mathfrak{C}_h$ into $\mathfrak{V}_h$. Further, if $j, j' \in J$, with $j' \leq j$, then

$$\begin{array}{ccc} B_j & \xrightarrow{\theta_j | B_j} & V_j \\ \downarrow{\scriptstyle \psi_{jj'}} & & \downarrow{\scriptstyle \eta_{jj'}} \\ B_{j'} & \xrightarrow[\theta_{j'} | B_{j'}]{} & V_{j'} \end{array}$$

is commutative. Therefore $\bigcup_{j \in J} (\theta_j | B_j) = \beta'$ is a monomorphism of $\mathfrak{B}$ into $\mathfrak{V}$. In the same way one shows that $\bigcup_{m \in M} (\theta_m | C_m) = \gamma'$ is a monomorphism of $\mathfrak{C}$ into $\mathfrak{V}$. Since $\beta' | A = \gamma' | A$, we may conclude that the amalgam $\langle \mathfrak{A}, \beta, \mathfrak{B}, \gamma, \mathfrak{C} \rangle$ is embeddable in $\mathfrak{V}$ by β' and γ'. Thus $K_{R(K)}$ satisfies WAP.

Let us again consider the above amalgam $\langle \mathfrak{A}, \beta, \mathfrak{B}, \gamma, \mathfrak{C} \rangle$ which can be embedded in $\mathfrak{V}$ by the monomorphisms β' and γ'. Let $i \in I$. The algebra $\mathfrak{V}_i \in K$ may considered to be freely generated by the elements of $(B_i \cup C_i)\theta_i$, subject to the identities in $E(K)$, and subject to the equalities (13) and (14) that hold in $\mathfrak{B}_i$ and $\mathfrak{C}_i$. If the amalgam $\langle \mathfrak{A}_i, \beta | A_i, \mathfrak{B}_i, \gamma | A_i, \mathfrak{C}_i \rangle$ is strongly embeddable in K, then $B_i \theta_i \cap C_i \theta_i = A_i \theta_i$ in $\mathfrak{V}_i$. Consequently, if K satisfies SAP, then the amalgam $\langle \mathfrak{A}, \beta, \mathfrak{B}, \gamma, \mathfrak{C} \rangle$ is strongly embedded in $\mathfrak{V}$ by the monomorphisms β' and γ'. Thus if K satisfies CEP and SAP, then $K_{R(K)}$ satisfies SAP.

For a special case of the following theorem, see [7] §3.

THEOREM 4. *Let K be a variety, where $E(K)$ contains a strongly non-regular identity* (1). *Then K satisfies CEP if and only if $K_{R(K)}$ satisfies CEP.*

PROOF. Let us suppose that K satisfies CEP. Let $\mathfrak{A}$ be a subalgebra of the algebra $\mathfrak{B} \in K_{R(K)}$. We know that $\mathfrak{B}$ is the sum of a semilattice ordered system (8), and that $\mathfrak{A}$ is the sum of a semilattice ordered system (5), where we suppose that the components $\mathfrak{A}_i$, $i \in I$, and $\mathfrak{B}_j$, $j \in J$, belong to K. Since K is a strongly non-regular variety, the mapping $I \to J$, $i \to j$, if $\mathfrak{A}_i \subseteq \mathfrak{B}_j$, is a well-defined semilattice monomorphism. We shall henceforth suppose that $\mathfrak{I} = \langle I; \wedge \rangle$ is a subsemilattice of $\mathfrak{J} = \langle J; \wedge \rangle$. We remark once more that for $i, i' \in I$, with $i' \leq i$ in $\mathfrak{I}$ we have $\varphi_{ii'} = \psi_{ii'} | A_i$. Indeed, for every $a \in A_i$ and some $a' \in A_{i'}$ we must have $a\varphi_{ii'} = p(a_1, \dots, a_n) = a\psi_{ii'}$ where $a = a_l$ and $a_k = a'$ for all $k \in \{1, \dots, n\}$, $k \neq l$. Let ϱ be a congruence relation on $\mathfrak{A}$ and let θ be the intersection of all the congruence relations on $\mathfrak{B}$ that contain ϱ. We must show that θ induces ϱ on $\mathfrak{A}$.

Let us suppose that a and a' are θ-related elements of $\mathfrak{A}$. Then there exists a sequence

$$\begin{aligned} &a = p_1(a_{11}, \dots, a_{1m_1}), \quad p_1(b_{11}, \dots, b_{1m_1}) = p_2(a_{21}, \dots, a_{2m_2}), \dots \\ &\dots, p_{u-1}(b_{u-11}, \dots, b_{u-1m_{u-1}}) = p_u(a_{u1}, \dots, a_{um_u}), \quad p_u(b_{u1}, \dots, b_{um_u}) = a' \end{aligned} \tag{19}$$

where the above equalities hold in $\mathfrak{B}$, and where for each $1 \leq q \leq u$, and each $1 \leq r \leq m_q$ either $a_{qr} = b_{qr}$ or $a_{qr} \, \varrho \, b_{qr}$ in $\mathfrak{A}$. We may suppose that for each $1 \leq q \leq u$, $\mathbf{x}_1, \dots, \mathbf{x}_{m_q}$ occur in $\mathbf{p}_q(\mathbf{x}_1, \dots, \mathbf{x}_{m_q})$. For each $1 \leq q \leq u$, $1 \leq r \leq m_q$, we suppose

$a_{qr}\in B_{j_{qr}}$ and $b_{qr}\in B_{l_{qr}}$. Let us put $j_q=\bigwedge_{r=1}^{m_q} j_{qr}$ and $l_q=\bigwedge_{r=1}^{m_q} l_{qr}$. Observe that $l_q=j_{q+1}$ for $q=1, \dots, u-1$. Further, $a\in A_{j_1}$ and $a'\in A_{l_u}$.

We now proceed to show that $i=\left(\bigwedge_{q=1}^{u} j_q\right)\wedge l_u\in I$ and that $a\,\varrho\, a\varphi_{j_1 i}$ in $\mathfrak{A}$. Let us suppose that $a\,\varrho\, a\varphi_{j_1 i_s}$, where $i_s=\bigwedge_{q=1}^{s} j_q\in I$ (this statement holds trivially for $s=1$). Observe that $j_{s1}, \dots, j_{sm_s}\geqq i_s$. If $a_{s1}=b_{s1}$, then $j_{s1}=l_{s1}$ and $i_s\wedge l_{s1}=i_s\in I$, $a\,\varrho\, a\varphi_{j_1 i_s\wedge l_{s1}}$. If $a_{s1}\neq b_{s1}$, then $j_{s1}, l_{s1}\in I$, $a_{s1}, b_{s1}\in A$ and $a_{s1}\,\varrho\, b_{s1}$ in $\mathfrak{A}$. We can put

$$a\,\varrho\, a\varphi_{j_1 i_s} = p(a_1, \dots, a_n)\,\varrho\, p(a_1', \dots, a_n') = a\varphi_{j_1 i_s\wedge l_{s1}} \quad \text{in} \quad \mathfrak{A}$$

where $a_l=a_l'=a\varphi_{j_1 i_s}$ and $a_k=a_{s1}$, $a_k'=b_{s1}$ for $k\in\{1, \dots, n\}$, $k\neq l$. If $i_s\wedge l_{s1}\wedge \dots \wedge l_{sv}\in I$ for $v<m_s$ and $a\,\varrho\, a\varphi_{j_1 i_s\wedge l_{s1}\wedge\dots\wedge l_{sv}}$, then one can show in exactly the same way that $i_s\wedge l_{s1}\wedge\dots\wedge l_{sv}\wedge l_{sv+1}\in I$ and $a\,\varrho\, a\varphi_{j_1 i_s\wedge l_{s1}\wedge\dots\wedge l_{sv}\wedge l_{sv+1}}$. By induction we conclude that $i_s\wedge l_s\in I$ and $a\,\varrho\, a\varphi_{j_1 i_s\wedge l_s}$. If $s\leqq u-1$, then $l_s=j_{s+1}$, and so $i_s\wedge l_s=\bigwedge_{q=1}^{s+1} j_q=i_{s+1}$. Using induction on s, we have $i_u\in I$ and $a\,\varrho\, a\varphi_{j_1 i_u}$. Applying the above procedure, we again obtain $i=i_u\wedge l_u\in I$ and $a\,\varrho\, a\varphi_{j_1 i_u\wedge l_u}=a\varphi_{j_1 i}$. Observe that $i=j_1\wedge\left(\bigwedge_{q=1}^{u} l_q\right)\in I$. By symmetry we also obtain $a'\varrho\, a'\varphi_{l_u i}$.

For any $1\leqq q\leqq u$ and $1\leqq r\leqq m_q$, we put $a_{qr}\psi_{j_{qr}i}=a_{qr}'$ and $b_{qr}\psi_{l_{qr}i}=b_{qr}'$. If $a_{qr}=b_{qr}$, then $a_{qr}'=b_{qr}'$ in $\mathfrak{B}_i$. If $a_{qr}\,\varrho\, b_{qr}$ in $\mathfrak{A}$, then

$$a_{qr}' = p(a_1, \dots, a_n)\,\varrho\, p(b_1, \dots, b_n) = b_{qr}'$$

in $\mathfrak{A}$ where $a_l=a_{qr}$, $b_l=b_{qr}$ and $a_k=a_{qr}'=b_k$ for $k\in\{1, \dots, n\}$, $k\neq l$. Let us consider the sequence

$$\begin{aligned} &a\varphi_{j_1 i} = p_1(a_{11}', \dots, a_{1m_1}'), \quad p_1(b_{11}', \dots, b_{1m_1}') = p(a_{21}', \dots, a_{2m_2}'), \dots \\ (20)\qquad &\dots, p_{u-1}(b_{u-11}', \dots, b_{u-1\,m_{u-1}}') = p_u(a_{u1}', \dots, a_{um_u}'), \quad p_u(b_{u1}', \dots, b_{um_u}') = a'\varphi_{l_u i}. \end{aligned}$$

There exists a congruence relation on $\mathfrak{B}_i$ whose restriction to $\mathfrak{A}_i$ yields the restriction of ϱ to $\mathfrak{A}_i$ since K satisfies CEP. Therefore (20) shows that $a\varphi_{j_1 i}\,\varrho\, a'\varphi_{l_u i}$ in $\mathfrak{A}$. Consequently, $a\,\varrho\, a'$ in $\mathfrak{A}$. We conclude $\theta|A=\varrho$. We have proved that $K_{R(K)}$ satisfies CEP.

The converse part of the theorem is obvious.

COROLLARY 5. *Let K be a variety where $E(K)$ contains the strongly non-regular identity* (1). *Then*

(i) *if $K_{R(K)}$ satisfies WAP, then $K_{R(K)}$ satisfies CEP,*

(ii) *$K_{R(K)}$ satisfies WAP [SAP] if and only if K satisfies WAP [SAP] and CEP.*

COROLLARY 6. HSP$(\mathfrak{L})$ *satisfies SAP and CEP.*

The following theorem combines the two constructions we have discussed so far (see also [5] and [14]).

THEOREM 7. *Let K be a variety where $E(K)$ contains the strongly non-regular identity* (1). *Then $\overline{K}_{R(K)}$ is the variety which consists of the algebras which satisfy all* *the regular identities in* $E(K)$, *except perhaps the identities which are of the form* $\mathbf{x}=\mathbf{q}(\mathbf{x})$. *The variety* $\overline{K}_{R(K)}$ *consists of the algebras which are inflations of Płonka sums of semilattice ordered systems of algebras from* K. *Equivalently,* $\overline{K}_{R(K)}$ *consists of the algebras which are Płonka sums of semilattice ordered systems of inflations of algebras from* K.

Remark that in Theorem 7, $\overline{K}_{R(K)}$ contains $K_{R(K)}$ properly since $\mathbf{x}=\mathbf{p}(\mathbf{x}, \ldots, \mathbf{x})$ belongs to $R(K)$ but not to $E(\overline{K}_{R(K)})$. Using Theorem 2 and Corollary 5 we have the following

THEOREM 8. *Let* K *and* $\overline{K}_{R(K)}$ *be as in Theorem 7. Then*
(i) *if* $\overline{K}_{R(K)}$ *satisfies WAP, then* $\overline{K}_{R(K)}$ *satisfies CEP,*
(ii) $\overline{K}_{R(K)}$ *satisfies WAP [SAP] if and only if* K *satisfies WAP [SAP] and CEP.*

EXAMPLE 1. Semigroup varieties that consist of semigroups satisfying some fixed strongly non-regular identity have been characterized in [9], IV.2.17. (vi). The amalgamation properties and the congruence extension property for these varieties have been discussed in [2] and [1], respectively. It turns out ([2], 4.2.17) that varieties of semigroups that satisfy WAP must be either (i) varieties that consist of semigroups satisfying some fixed strongly non-regular identity, or (ii) varieties that can be constructed from (i) by applying Płonka sums and/or inflations.

EXAMPLE 2. Completely regular semigroups are semigroups that are the (disjoint) union of their maximal subgroups. They constitute a variety of algebras of type $\langle 2, 1\rangle$, the unary operation $^{-1}$ being the taking of the inverse within the same maximal subgroup. A completely regular semigroup that satisfies a strongly non-regular identity must be completely simple i.e. it must satisfy the strongly non-regular identity $\mathbf{x}=(\mathbf{xy})(\mathbf{xy})^{-1}\mathbf{x}$. We again see that the varieties of completely regular semigroups that satisfy WAP must be either (i) completely simple semigroup varieties or (ii) varieties that can be obtained from (i) by considering Płonka sums. Here the consideration of inflations leads to algebras which are not completely regular semigroups. Due to the fact that abelian group varieties are the only group varieties that satisfy CEP [1], the varieties constructed in (ii) must be rather special (see [2], Chapter II).

EXAMPLE 3. More examples of varieties satisfying WAP and CEP are obtained by considering Płonka sums of semilattice ordered systems of Boolean algebras [12] or of distributive lattices [11]. Quasilattices (Płonka sums of semilattice ordered systems of lattices [7]) however do not satisfy WAP, since lattices do not satisfy CEP.

REFERENCES

[1] BÍRÓ, B., KISS, E. W. and PÁLFY, P. P., On the congruence extension property, *Semigroup Forum* **15** (1977/78), 183—184. *MR* **57** # 6246.

[2] CLARKE, G. T., Semigroup varieties with the amalgamation property, Doctoral Thesis, Monash University, 1981.

[3] DUDEK, J. and GRACZYŃSKA, E., The lattice of varieties of algebras (to appear).

[4] GRÄTZER, G., *Universal algebra,* Second edition, Springer-Verlag, New York—Heidelberg, 1979. *MR* **80g**: 08001.

[5] HAŁKOWSKA, K., On some operator defined on equational classes of algebras, *Arch. Math. (Brno)* **12** (1976), 209—212. *MR* **55** # 5498.

[6] HALL, T. E., Inverse semigroup varieties with the amalgamation property, *Semigroup Forum* **16** (1978), 37—51. *MR* **58** # 11174.
[7] IMAOKA, T., Free products with amalgamation of bands, *Mem. Fac. Lit. Sci. Shimane Univ. Natur. Sci.* **10** (1976), 7—17. *MR* **55** # 3118.
[8] PADMANABHAN, R., Regular identities in lattices, *Trans. Amer. Math. Soc.* **158** (1971), 179—188. *MR* **43** # 7376.
[9] PETRICH, M., *Structure of regular semigroups*, Cahiers Mathématiques Montpellier, No. 11, Université des Sciences et Techniques du Languedoc, U.E.R. de Mathématiques, Montpellier, 1977. *Zbl* **385**. 20033; *MR* **58** # 28234.
[10] PŁONKA, J., On a method of construction of abstract algebras, *Fund. Math.* **60** (1967), 183—189. *MR* **37** # 1294.
[11] PŁONKA, J., On distributive quasi-lattices, *Fund. Math.* **60** (1967), 191—200. *MR* **36** # 85.
[12] PŁONKA, J., On sums of direct systems of Boolean algebras, *Colloq. Math.* **20** (1969), 209—214. *MR* **40** # 2590.
[13] PŁONKA, J., On equational classes of abstract algebras defined by regular equations, *Fund. Math.* **64** (1969), 241—247. *MR* **39** # 5450.
[14] PŁONKA, J., On the subdirect product of some equational classes of algebras, *Math. Nachr.* **63** (1974), 303—305. *MR* **50** # 12864.

(Received March 12, 1982)

DIENST HOGERE MEETKUNDE
RIJKSUNIVERSITEIT TE GENT
KRIJGSLAAN 281
B—9000 GENT
BELGIUM

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 113—127

ON DISTRIBUTION OF LINEAR RECURRENCES MODULO 1

PÉTER KISS and SÁNDOR MOLNÁR

To the memory of Professor László Rédei

Introduction

Let (G_n), $n=0, 1, 2, \ldots$, be a linear recurrence of order k $(k>1)$ defined by rational integers $A_1, A_2, \ldots, A_k$ $(A_k \neq 0)$ and by recursion

$$G_n = A_1 G_{n-1} + A_2 G_{n-2} + \ldots + A_k G_{n-k} \tag{1}$$

$(n \geqq k)$, where the initial values $G_0, G_1, \ldots, G_{k-1}$ are fixed not all zero rational integers. We suppose that the roots $\alpha_1, \alpha_2, \ldots, \alpha_k$ of the characteristic polynomial

$$g(x) = x^k - A_1 x^{k-1} - A_2 x^{k-2} - \ldots - A_k$$

are distinct and s is an integer such that

$$|\alpha_1| \geqq |\alpha_2| \geqq \ldots \geqq |\alpha_s| \geqq 1 > |\alpha_{s+1}| \geqq \ldots \geqq |\alpha_k|.$$

Let us introduce the notations

$$D = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ \alpha_1 & \alpha_2 & \ldots & \alpha_k \\ \alpha_1^2 & \alpha_2^2 & \ldots & \alpha_k^2 \\ \vdots & \vdots & & \vdots \\ \alpha_1^{k-1} & \alpha_2^{k-1} & \ldots & \alpha_k^{k-1} \end{vmatrix}$$

and

$$D_i(x_1, x_2, \ldots, x_k) = \begin{vmatrix} 1 & \ldots & 1 & x_1 & 1 & \ldots & 1 \\ \alpha_1 & \ldots & \alpha_{i-1} & x_2 & \alpha_{i+1} & \ldots & \alpha_k \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_1^{k-1} & \ldots & \alpha_{i-1}^{k-1} & x_k & \alpha_{i+1}^{k-1} & \ldots & \alpha_k^{k-1} \end{vmatrix},$$

where the determinant D is not zero since $\alpha_i \neq \alpha_j$ if $i \neq j$.

It is well-known that the explicit form of the terms of sequence (G_n) is

$$G_n = a_1 \alpha_1^n + a_2 \alpha_2^n + \ldots + a_k \alpha_k^n, \tag{2}$$

where

$$a_i = \frac{D_i(G_0, G_1, \ldots, G_{k-1})}{D}. \tag{3}$$

1980 *Mathematics Subject Classification.* Primary 10A35; Secondary 10K05.
Key words and phrases. Linear recurrence, distribution, density, limit points.

We say the sequence (G_n) of order k can be reduced if it satisfies a linear recursion of order less than k, too. In the following we assume the sequence (G_n) cannot be reduced. Furthermore, we assume that the ratio α_i/α_j of two roots of polynomial $g(x)$ is not a root of unity for any i and j $(i\neq j,\ 1\leq i, j\leq k)$. We say (G_n) is a non-degenerate sequence if these assumptions hold.

Some limit and distribution properties of special sequences (G_n) were studied by several authors. For example W. Gerdes [5, 6] gave conditions for the existence of limit $\lim_{n\to\infty} G_n$ in cases $k=2$ and $k=3$ if $A_1, A_2, \ldots, A_k$ and $G_0, G_1, \ldots, G_{k-1}$ are real numbers. M. B. Gregory and J. M. Metzger [7] studied the condition of the existence of limit $\lim_{n\to\infty} (\sin (G_n x\pi))$ for Fibonacci type second order recurrences, they showed that the limit exists if and only if x is an element of certain subset of $\mathbf{Q}(\sqrt{5})$. This result was extended for more general second order recurrences by S. Molnár [14]. An interesting result was obtained by A. Perelli and U. Zannier [15]: If in (2) the α_i's are rational numbers and $\lim_{n\to\infty} \|G_n\|=0$, where $\|x\|=\min|x-m|$ for $m\in\mathbf{Z}$, then $\alpha_1, \alpha_2, \ldots, \alpha_s$ are integers; furthermore, it remains true if the a_i's are polynomials of variable n.

Another type results concerning the distribution of the sequence (μG_n) modulo 1, where μ is a fixed real number, were obtained in special cases.

If $\alpha_1>1$ and $|\alpha_i|<1$ for $i=2, 3, \ldots, k$ then, by (2), the study of the sequence (μG_n) modulo 1 is equivalent to the study of the sequence $(\lambda\alpha^n)$, where λ is a real number and α is an algebraic integer such that the absolute values of its conjugates are all less then one. The algebraic integers having this property are called Pisot—Vijayaraghavan numbers (abbreviated PV numbers). Another characterization of PV numbers was given by M. Mendés France [13].

The properties of the sequences $(\lambda\alpha^n)$ modulo 1, where λ and α are real numbers, were studied by several authors; recently, for example, G. Choquet [3] and F. Beukers [1] obtained some results. We cite only a former result due to C. Pisot [16] which is closer to our results: (i) The set of all numbers α and λ, such that the set E of the limit points of $(\lambda\alpha^n)$ modulo 1 consists of a finite number of points, is denumerable; (ii) If α is algebraic, necessary and sufficient condition for E to be finite are that α is a PV number and $\lambda\in\mathbf{Q}(\alpha)$; (iii) If E is finite and t denotes the number of irrational numbers in E and if the convergence of sequence $(\lambda\alpha^n)$ modulo 1 towards its limit points is $o(n^{-t-1})$, then α is algebraic and therefore, by (ii), α is a PV number and $\lambda\in\mathbf{Q}(\alpha)$.

For general linear recurrences we know only a few results. A general result is the following one: If the terms of (G_n) are all distinct or $G_i\neq G_j$ for $\max(i, j)>n_0$, then the sequence (μG_n) is uniformly distributed modulo 1 for almost all real numbers μ (see L. Kuipers and H. Niederreiter [10], Theorem 4.1 on p. 32). The condition $G_i\neq G_j$ for large indices holds if $|\alpha_1|>|\alpha_i|$ for $i=2, 3, \ldots, k$, as it was shown in [8]. M. B. Levin and I. E. Sparlinski [12] gave a construction to find real numbers $a_1, a_2, \ldots, a_k$ in (2) such that the sequence (G_n) of real numbers was uniformly distributed modulo 1. L. Kuipers and J. S. Shiue [11] proved that the sequence $(\log G_n)$ is uniformly distributed modulo 1 if $|\alpha_1|>|\alpha_2|>\ldots>|\alpha_k|>0$ and $|\alpha_i|\neq 1$ for $i=1, 2, \ldots, k$. This is an extension of a result of J. L. Brown and R. L. Duncan [2] on Fibonacci-type sequences. Finally we quote a theorem of T. Vijayaraghavan [17] which has connection with linear recurrences, by (2): Let $\gamma_1, \gamma_2, \ldots, \gamma_r$ be algebraic numbers with $|\gamma_i|>1$ and let $d_1, d_2, \ldots, d_r$ be nonzero constants. If the limit points of the fractional parts of sequence $u_n = d_1\gamma_1^n + d_2\gamma_2^n + \ldots + d_r\gamma_r^n$ $(n=1, 2, \ldots)$ are finite in number, then the γ_i's are algebraic integers, and any conjugates of any of the γ_i not among them have absolute values less than one.

In the following we shall give necessary and sufficient conditions for number μ that the sequence (μG_n) has finitely many points of accumulation modulo 1. Furthermore in a special case ($g(x)$ is the minimal polynomial of a PV number) we show a construction to find uncountable many numbers μ such that the sequence (μG_n) has infinitely many limit points modulo 1 but it is not uniformly distributed.

Results

Throughout this paper $\{x\}$ denotes the fractional part of real number x, and $\|x\|$ denotes the distance from x to the nearest integer. Furthermore we say zero is an only point of accumulation of sequence $\{x_n\}$, or $\lim_{n\to\infty} \{x_n\} = 0$ if for any $\varepsilon > 0$ and $n > n(\varepsilon)$ we have $0 \leqq \{x_n\} < \varepsilon$ or $1 - \varepsilon < \{x_n\} < 1$, i.e. $\|x_n\| < \varepsilon$.

In (2) $\alpha_1, \alpha_2, \ldots, \alpha_k$ are algebraic integers and so, by the cited results of C. Pisot and T. Vijayaraghavan, the sequence $\{\mu G_n\}$ may have only finitely many points of accumulation for certain real numbers μ. The following theorem shows that in this case the limit points must be rational numbers and the μ's are determined.

THEOREM 1. *Let (G_n) be a non-degenerate linear recurrence and let μ be a real number. Suppose that the roots $\alpha_1, \alpha_2, \ldots, \alpha_k$ of the characteristic polynomial $g(x)$ are not roots of unity. Then the number of the limit points of sequence $\{\mu G_n\}$ is finite if and only if there are integers $N, P_N, P_{N+1}, \ldots, P_{N+k-1}$ and rational numbers $q_0, q_1, \ldots, q_{k-1}$ such that μ satisfies the equations*

$$\mu = \frac{D_i(P_N + q_0, P_{N+1} + q_1, \ldots, P_{N+k-1} + q_{k-1})}{\alpha_i^N \cdot D_i(G_0, G_1, \ldots, G_{k-1})}$$

simultaneously for indices $i = 1, 2, \ldots, s$. Furthermore, if the number of the limit points is finite then they are rational numbers.

If (μG_n) is a convergent sequence modulo 1, i.e. the sequence $\{\mu G_n\}$ has only one limit point, say q, then, as we shall see in the proof of Theorem 1, in the expression of μ we have $q_0 = q_1 = \ldots = q_{k-1} = q$. But it does not imply the existence of a single point of accumulation. Therefore the following consequence is not trivial.

COROLLARY 1. *Let $\sum_{i=1}^{k} A_i \neq 1$ and let μ be a real number. The sequence $\{\mu G_n\}$ is convergent if and only if μ satisfies simultaneously the equations*

$$\mu = \frac{D_i(P_N + q, P_{N+1} + q, \ldots, P_{N+k-1} + q)}{\alpha_i^N D_i(G_0, G_1, \ldots, G_{k-1})}$$

for all indices i with $1 \leqq i \leqq s$, where $N, P_N, P_{N+1}, \ldots, P_{N+k-1}$ are rational integers and q is a rational number of the form $q = t/(A_1 + A_2 + \ldots + A_k - 1)$ with integer t. If these conditions hold then $\lim_{n\to\infty} \{\mu G_n\} = \{q\}$.

Note that the conditions in Corollary 1 are weaker than conditions in Theorem 1. Namely condition $A_1+A_2+\ldots+A_k\neq 1$ gives the restriction $\alpha_i\neq 1$ instead of $\alpha_i^n\neq 1$ for any n and $1\leqq i\leqq k$.

Theorem 1 shows that the sequence $\{\mu G_n\}$ has infinitely many points of accumulation if (G_n) satisfies the conditions of the theorem and μ is not any element of algebraic number field $Q(\alpha_1, \alpha_2, \ldots, \alpha_k)$. However, in these cases it is not sure to be the sequence (μG_n) uniformly distributed modulo 1.

THEOREM 2. *Let (G_n) be a non-degenerate linear recurrence. Let us suppose that the characteristic polynomial $g(x)$ of (G_n) is the minimal polynomial of a PV number. Then there are uncountable many real number μ such that the numbers $\{\mu G_n\}$ are everywhere dense in the interval* $[0, 1)$ *but they are not uniformly distributed modulo 1.*

The proof of Theorem 2 implies the following result.

COROLLARY 2. *Let (G_n) be a sequence satisfying the conditions of Theorem 2. Then there are uncountable many real numbers μ such that the sequence $\{\mu G_n\}$ has infinitely many points of accumulation, but the terms of the sequence are not everywhere dense in interval* $[0, 1)$.

For the proof of the theorems we need some lemmas.

LEMMA 1. *If the sequence (G_n) cannot be reduced then for the coefficients in the explicit form* (2) *we have* $a_i\neq 0$ $(i=1, 2, \ldots, k)$.

LEMMA 2. *Let $z_0, z_1, \ldots, z_t, w_1, w_2, \ldots, w_t$ be complex numbers with condition* $|w_1|=|w_2|=\ldots=|w_t|=1$. *If w_i's are distinct, none of them is a root of unity, and*

$$\lim_{n\to\infty}\left(z_0+\sum_{i=1}^{t} z_i w_i^n\right)=0 \tag{4}$$

then $z_0=z_1=\ldots=z_t=0$.

The proofs of the results

PROOF OF LEMMA 1. It was shown in [9] that

$$D_i(G_0, G_1, \ldots, G_{k-1}) = c_i d(\alpha_i)$$

where

$$d(x) = \sum_{i=0}^{k-1}\left(G_{k-i-1}-\sum_{j=1}^{k-i-1} A_j G_{k-i-j-1}\right)x^i$$

and c_i is a non-zero number if the roots of $g(x)$ are distinct (see the proof of Theorem 2 in [9]) so $a_i=0$ if and only if $d(\alpha_i)=0$. It follows from a result of M. D'Ocagne [4] that the sequence (G_n) can be reduced if and only if the polynomials $g(x)$ and $d(x)$ are not coprime (see [4] pp. 156—159). Thus we have, if the terms of (G_n) do not satisfy any recursion of order less than k then $g(x)$ and

$d(x)$ are relatively prime polynomials and so $d(\alpha_i)\neq 0$ for $i=1, 2, \ldots, k$. From these, by

$$a_i = \frac{D_i(G_0, G_1, \ldots, G_{k-1})}{D} = \frac{c_i d(\alpha_i)}{D},$$

the statement follows since $D\neq 0$.

PROOF OF LEMMA 2. Clearly, the numbers $w_1, w_2, \ldots, w_t$, by the conditions, are not-real complex numbers. Let $\varphi_j=\arg(w_j)$ $(j=1, 2, \ldots, t)$. By the conditions $\varphi_j/2\pi$ is an irrational number for $1\leqq j\leqq t$ therefore there are infinitely many integer n such that for any $\varepsilon>0$ the inequality

$$\|n\varphi_j/2\pi\| < \varepsilon/2\pi \tag{5}$$

holds simultaneously for $j=1, 2, \ldots, t$. If n is an integer satisfying (5) then

$$w_j^n = e^{in\varphi_j} = e^{i(2\pi p_j+\varepsilon_j)} = e^{i\varepsilon_j},$$

where p_j is some integer and $|\varepsilon_j|<\varepsilon$, and so there are infinitely many integer n such that

$$|z_j w_j^n - z_j| < \varepsilon$$

simultaneously for $j=1, 2, \ldots, t$. Thus if limit (4) exists then it must be $(z_0+z_1+\ldots$ $\ldots+z_t)$, furthermore

$$\lim_{n\to\infty} \sum_{j=1}^{t} z_j w_j^n = \sum_{j=1}^{t} z_j. \tag{6}$$

By similar argument, substituting $z_j w_j^r$ for z_j, the existence of limit (4) implies

$$\lim_{n\to\infty} \sum_{j=1}^{t} (z_j w_j^r) w_j^n = \sum_{j=1}^{t} z_j w_j^r, \tag{7}$$

where r is some integer, furthermore the left sides of (6) and (7) are equal. Using this result in cases $r=1, 2, \ldots, t$, by (6) and (7) we get the system of equations

$$\begin{aligned} z_1w_1+z_2w_2+\ldots+z_tw_t &= z_1+z_2+\ldots+z_t \\ z_1w_1^2+z_2w_2^2+\ldots+z_tw_t^2 &= z_1+z_2+\ldots+z_t \\ &\vdots \\ z_1w_1^t+z_2w_2^t+\ldots+z_tw_t^t &= z_1+z_2+\ldots+z_t \end{aligned}$$

which implies the homogeneous system

$$\begin{aligned} (w_1-1)z_1+(w_2-1)z_2+\ldots+(w_t-1)z_t &= 0 \\ (w_1^2-1)z_1+(w_2^2-1)z_2+\ldots+(w_t^2-1)z_t &= 0 \\ &\vdots \\ (w_1^t-1)z_1+(w_2^t-1)z_2+\ldots+(w_t^t-1)z_t &= 0 \end{aligned}$$

in variables $z_1, z_2, \dots, z_t$. The determinant of this system is

$$d = \begin{vmatrix} w_1-1 & w_2-1 & \dots & w_t-1 \\ w_1^2-1 & w_2^2-1 & \dots & w_t^2-1 \\ \vdots & \vdots & & \vdots \\ w_1^t-1 & w_2^t-1 & \dots & w_t^t-1 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & w_1 & w_2 & \dots & w_t \\ 1 & w_1^2 & w_2^2 & \dots & w_t^2 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & w_1^t & w_2^t & \dots & w_t^t \end{vmatrix}.$$

It can be written in the form

$$d = \prod_{1 \leq i < j \leq t} (w_j - w_i) \prod_{j=1}^{t} (w_j - 1)$$

which is not zero by the conditions, therefore the system of equations has only the trivial solution $z_1 = z_2 = \dots = z_t = 0$ from which, by (4), $z_0 = 0$ follows.

PROOF OF THEOREM 1. Let μ be a real number and let S be the finite set of the accumulation points of sequence $\{\mu G_n\}$. Then the terms of the sequence (μG_n) is of the form

$$\mu G_n = P_n + q_n + \varepsilon_n, \tag{8}$$

where P_n is an integer, $q_n \in S$, and $\varepsilon_n \to 0$ as $n \to \infty$. Define numbers T_n by

$$T_n = P_{n+k} + q_{n+k} - \sum_{i=1}^{k} A_i (P_{n+k-i} + q_{n+k-i}) \tag{9}$$

for $n \geqq 0$. By (8), (9), and (1) we have

$$T_n = \mu G_{n+k} - \varepsilon_{n+k} - \sum_{i=1}^{k} A_i (\mu G_{n+k-i} - \varepsilon_{n+k-i}) = -\varepsilon_{n+k} + \sum_{i=1}^{k} A_i \varepsilon_{n+k-i} \tag{10}$$

and so

$$\lim_{n \to \infty} T_n = 0.$$

But, by (9),

$$\{T_n\} = \left\{ q_{n+k} - \sum_{i=1}^{k} A_i q_{n+k-i} \right\}$$

has only finitely many values different from zero, therefore there is an integer N such that

$$T_n = 0$$

for $n \geqq N$; furthermore, by (9) and (10), the sequences $(P_n + q_n)$ and (ε_n), $n = N, N+1, N+2, \dots$, are linear recurrences of order k with characteristic polynomial $g(x)$. So we have

$$P_n + q_n = p_1 \alpha_1^{n-N} + p_2 \alpha_2^{n-N} + \dots + p_k \alpha_k^{n-N}$$

for $n \geqq N$, where

$$p_i = \frac{D_i(P_N + q_N, P_{N+1} + q_{N+1}, \dots, P_{N+k-1} + q_{N+k-1})}{D}$$

$(i=1, 2, \ldots, k)$, $P_N, P_{N+1}, \ldots, P_{N+k-1}$ are some integers and $q_j \in S$ for $j=N, N+1, \ldots$ $\ldots, N+k-1$. Using this explicit form from (8)

$$\varepsilon_n = \mu G_n - (P_n + q_n) = \mu \sum_{i=1}^{k} a_i \alpha_i^n - \sum_{i=1}^{k} p_i \alpha_i^{n-N} =$$

$$= \alpha_1^{n-N}\left[\mu a_1 \alpha_1^N - p_1 + \sum_{i=2}^{k} (\mu a_i \alpha_i^N - p_i)\left(\frac{\alpha_i}{\alpha_1}\right)^{n-N}\right]$$

follows for $n \geqq N$.

Since $\lim_{n\to\infty} \varepsilon_n = 0$ and $|\alpha_i/\alpha_1| \leqq 1$ for $i=2, 3, \ldots, k$, Lemma 2 implies the equalities

$$\mu a_i \alpha_i^N - p_i = 0 \tag{11}$$

for $i=1$ and for $i>1$ satisfying condition $|\alpha_i| = |\alpha_1|$. Using this result and repeating the argument, we get the validity of (11) for all indices i satisfying condition $|\alpha_i| \geqq 1$.

Thus if the number of the limit points of sequence $\{\mu G_n\}$ is finite then (11) holds for $1 \leqq i \leqq s$, which implies the statement of the theorem on the form of μ since $a_i \neq 0$ $(i=1, 2, \ldots, k)$ by Lemma 1.

We prove that the elements of the set S are rational. We have shown that if the set S is finite then we have a linear recurrence $(P_n + q_n)$, $n=N, N+1, N+2, \ldots$, where P_n is an integer and $q_n \in S$. From this follows that the limit points q_n are terms of a sequence (g_n), $n=0, 1, 2, \ldots$, for which $g_n \in S$ and

$$g_n \equiv A_1 g_{n-1} + A_2 g_{n-2} + \ldots + A_k g_{n-k} \pmod 1 \tag{12}$$

thus (g_n) is a linear recurrence modulo 1 with characteristic polynomial $g(x)$. This sequence is uniquely determined by k consecutive terms $g_i, g_{i+1}, \ldots, g_{i+k-1}$ and S is a finite set, therefore (g_n) is a periodic sequence. We can assume that (g_n) is purely periodic and, by (12), it is sufficient to prove that $g_0, g_1, \ldots, g_{k-1}$ are rationals. By the periodicity

$$\begin{aligned} g_n &= g_0 \\ g_{2n} &= g_0 \\ &\vdots \\ g_{kn} &= g_0 \end{aligned} \tag{13}$$

for some integer n.

We need a result due to D'Ocagne (see [4], p. 161): Let (G_n) be a linear recurrence with characteristic polynomial $g(x)$. If (R_n), $n=0, 1, 2, \ldots$, is also a linear recurrence defined by the same characteristic polynomial $g(x)$ and by initial values $R_0 = R_1 =$ $= \ldots = R_{k-2} = 0$, $R_{k-1} = 1$ then

$$G_n = \sum_{i=0}^{k-1} G_i (R_{n+k-i-1} - A_1 R_{n+k-i-2} - \ldots - A_{k-i-1} R_n),$$

or in other form, by recursion (1),

$$G_n = \sum_{i=0}^{k-2} G_i (A_{k-i} R_{n-1} + A_{k-i+1} R_{n-2} + \ldots + A_k R_{n-i-1}) + G_{k-1} R_n$$

for $n \geqq k$. Using this result and that (g_n) is a linear recurrence modulo 1, (13) can be written in form

$$(A_k R_{n-1}-1)\, g_0 + \sum_{i=1}^{k-2}\left(\sum_{j=0}^{i} A_{k-i+j} R_{n-1-j}\right) g_i + R_n g_{k-1} = t_1$$

$$(A_k R_{2n-1}-1)\, g_0 + \sum_{i=1}^{k-2}\left(\sum_{j=0}^{i} A_{k-i+j} R_{2n-1-j}\right) g_i + R_{2n} g_{k-1} = t_2$$

$$\vdots \qquad\qquad\qquad\qquad \vdots$$

$$(A_k R_{kn-1}-1)\, g_0 + \sum_{i=1}^{k-2}\left(\sum_{j=0}^{i} A_{k-i+j} R_{kn-1-j}\right) g_i + R_{kn} g_{k-1} = t_k,$$

where $t_1, t_2, \ldots, t_k$ are some integers. This is a system of equations in variables $g_0, g_1, \ldots, g_{k-1}$ and the coefficients of variables are integers. So we have to prove that the $k \times k$ determinant $\mathbf{M}$ of this system is not zero and so the solutions are rational. Let us denote the i-th column vector of the determinant $\mathbf{M}$ by $\mathbf{v}_i$ $(i=1, 2, \ldots, k)$. Then the j-th coordinates of these vectors $\mathbf{v}_i$ are

$$\begin{aligned}
\mathbf{v}_{1,j} &= A_k R_{jn-1}-1\\
\mathbf{v}_{2,j} &= A_{k-1} R_{jn-1}+A_k R_{jn-2}\\
&\vdots\\
\mathbf{v}_{i,j} &= A_{k-i+1} R_{jn-1}+A_{k-i+2} R_{jn-2}+\ldots+A_k R_{jn-i}\\
&\vdots\\
\mathbf{v}_{k-1,j} &= A_2 R_{jn-1}+A_3 R_{jn-2}+\ldots+A_k R_{jn-k+1}\\
\mathbf{v}_{k,j} &= R_{jn}
\end{aligned}$$

which, with (1), imply the equalities

$$\begin{aligned}
&\mathbf{v}_{k-1}+A_1\mathbf{v}_k = (R_{n+1}, R_{2n+1}, \ldots, R_{kn+1}) = \mathbf{v}'_{k-1}\\
&\mathbf{v}_{k-2}+A_1\mathbf{v}'_{k-1}+A_2\mathbf{v}_k = (R_{n+2}, R_{2n+2}, \ldots, R_{kn+2}) = \mathbf{v}'_{k-2}\\
&\vdots\\
&\mathbf{v}_1+A_1\mathbf{v}'_2+A_2\mathbf{v}'_3+\ldots+A_{k-2}\mathbf{v}'_{k-1}+A_{k-1}\mathbf{v}_k =\\
&\qquad = (R_{n+k-1}-1, R_{2n+k-1}-1, \ldots, R_{kn+k-1}-1).
\end{aligned}$$

By these equalities we get

$$\mathbf{M} = \begin{vmatrix}
R_{n+k-1}-1 & R_{n+k-2} & \ldots & R_{n+1} & R_n\\
R_{2n+k-1}-1 & R_{2n+k-2} & \ldots & R_{2n+1} & R_{2n}\\
\vdots & \vdots & & \vdots & \vdots\\
R_{kn+k-1}-1 & R_{kn+k-2} & \ldots & R_{kn+1} & R_{kn}
\end{vmatrix}.$$

The terms of the sequence (R_n) are of the form

$$R_n = r_1\alpha_1^n + r_2\alpha_2^n + \ldots + r_k\alpha_k^n$$

$(n=0, 1, 2, \ldots)$, where $r_1, r_2, \ldots, r_k$ are some algebraic numbers and $r_i \neq 0$ for $i=1, 2, \ldots, k$ which can be seen by the proof of Lemma 1. We have

$$\sum_{i=1}^{k} r_i \alpha_i^j = 0$$

for $0 \leqq j \leqq k-2$ and

$$\sum_{i=1}^{k} r_i \alpha_i^{k-1} = 1$$

since $R_0 = R_1 = \ldots = R_{k-2} = 0$ and $R_{k-1} = 1$. From this it follows that

$$R_{tn+j} = \sum_{i=1}^{k} r_i \alpha_i^{tn+j} = \sum_{i=1}^{k} r_i \alpha_i^j (\alpha_i^{tn} - 1) = \sum_{i=1}^{k} r_i \alpha_i^j E_{tn}^{(i)}$$

for $0 \leqq j \leqq k-2$ and

$$R_{tn+k-1} - 1 = \sum_{i=1}^{k} r_i \alpha_i^{tn+k-1} - \sum_{i=1}^{k} r_i \alpha_i^{k-1} =$$

$$= \sum_{i=1}^{k} r_i \alpha_i^{k-1} (\alpha_i^{tn} - 1) = \sum_{i=1}^{k} r_i \alpha_i^{k-1} E_{tn}^{(i)},$$

where $t \geqq 0$ is any integer and $E_m^{(i)}$ denotes the number

$$E_m^{(i)} = \alpha_i^m - 1.$$

So we can write

$$\mathbf{M} = \begin{vmatrix} \Sigma r_i \alpha_i^{k-1} E_n^{(i)} & \Sigma r_i \alpha_i^{k-2} E_n^{(i)} & \ldots & \Sigma r_i E_n^{(i)} \\ \Sigma r_i \alpha_i^{k-1} E_{2n}^{(i)} & \Sigma r_i \alpha_i^{k-2} E_{2n}^{(i)} & \ldots & \Sigma r_i E_{2n}^{(i)} \\ \vdots & \vdots & & \vdots \\ \Sigma r_i \alpha_i^{k-1} E_{kn}^{(i)} & \Sigma r_i \alpha_i^{k-2} E_{kn}^{(i)} & \ldots & \Sigma r_i E_{kn}^{(i)} \end{vmatrix}.$$

Each of the elements in $\mathbf{M}$ is a sum with k summands and so we can express the determinant $\mathbf{M}$ as a sum of k^k determinants which are of the form

$$\begin{vmatrix} r_{i_1} \alpha_{i_1}^{k-1} E_n^{(i_1)} & r_{i_2} \alpha_{i_2}^{k-2} E_n^{(i_2)} & \ldots & r_{i_k} E_n^{(i_k)} \\ r_{i_1} \alpha_{i_1}^{k-1} E_{2n}^{(i_1)} & r_{i_2} \alpha_{i_2}^{k-2} E_{2n}^{(i_2)} & \ldots & r_{i_k} E_{2n}^{(i_k)} \\ \vdots & \vdots & & \vdots \\ r_{i_1} \alpha_{i_1}^{k-1} E_{kn}^{(i_1)} & r_{i_2} \alpha_{i_2}^{k-2} E_{kn}^{(i_2)} & \ldots & r_{i_k} E_{kn}^{(i_k)} \end{vmatrix}. \tag{14}$$

But a determinant of the form (14) can be different from zero only if $i_1, i_2, \ldots, i_k$ is a permutation of elements $1, 2, \ldots, k$; therefore

$$\mathbf{M} = (r_1 \cdot r_2 \cdot \ldots \cdot r_k) \cdot \begin{vmatrix} E_n^{(1)} & E_n^{(2)} & \ldots & E_n^{(k)} \\ E_{2n}^{(1)} & E_{2n}^{(2)} & \ldots & E_{2n}^{(k)} \\ \vdots & \vdots & & \vdots \\ E_{kn}^{(1)} & E_{kn}^{(2)} & \ldots & E_{kn}^{(k)} \end{vmatrix} \cdot B,$$

where

$$B = \Sigma (-1)^p \alpha_{i_1}^{k-1} \alpha_{i_2}^{k-2} \ldots \alpha_{i_k}^0,$$

the summation is extended to all permutations $i_1, i_2, \ldots, i_k$ of elements $1, 2, \ldots, k$, and p is the number of inversions in the permutation $i_1, i_2, \ldots, i_k$. On the other hand it can be easily seen that

$$B = \begin{vmatrix} \alpha_1^{k-1} & \alpha_2^{k-1} & \ldots & \alpha_k^{k-1} \\ \alpha_1^{k-2} & \alpha_2^{k-2} & \ldots & \alpha_k^{k-2} \\ \vdots & \vdots & & \vdots \\ \alpha_1 & \alpha_2 & & \alpha_k \\ 1 & 1 & & 1 \end{vmatrix} = \prod_{1 \leq i < j \leq k} (\alpha_i - \alpha_j)$$

and

$$\begin{vmatrix} E_n^{(1)} & E_n^{(2)} & \ldots & E_n^{(k)} \\ E_{2n}^{(1)} & E_{2n}^{(2)} & \ldots & E_{2n}^{(k)} \\ \vdots & \vdots & & \vdots \\ E_{kn}^{(1)} & E_{kn}^{(2)} & \ldots & E_{kn}^{(k)} \end{vmatrix} = \begin{vmatrix} \alpha_1^n - 1 & \alpha_2^n - 1 & \ldots & \alpha_k^n - 1 \\ \alpha_1^{2n} - 1 & \alpha_2^{2n} - 1 & \ldots & \alpha_k^{2n} - 1 \\ \vdots & \vdots & & \vdots \\ \alpha_1^{kn} - 1 & \alpha_2^{kn} - 1 & \ldots & \alpha_k^{kn} - 1 \end{vmatrix} =$$

$$= \begin{vmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & \alpha_1^n & \alpha_2^n & \ldots & \alpha_k^n \\ 1 & \alpha_1^{2n} & \alpha_2^{2n} & \ldots & \alpha_k^{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \alpha_1^{kn} & \alpha_2^{kn} & \ldots & \alpha_k^{kn} \end{vmatrix} = \prod_{1 \leq j < i \leq k} (\alpha_i^n - \alpha_j^n) \prod_{i=1}^{k} (\alpha_i^n - 1)$$

hence $\mathbf{M} \neq 0$; namely $\alpha_i^n \neq 1$ and $\alpha_i^n \neq \alpha_j^n$ for any i, j $(i \neq j)$ since neither α_i/α_j nor α_i is a root of unity by the conditions. This implies, as we have seen, that the limit points are really rational numbers.

The proof of the theorem will be complete if we show that the sequence (μG_n) has only finitely many limit points modulo 1 if μ satisfies the conditions given in the theorem.

Let $N, P_N, P_{N+1}, \ldots, P_{N+k-1}$ be integers and let $q_0, q_1, \ldots, q_{k-1}$ be rational numbers. Let us suppose that μ is a number satisfying the conditions given in the theorem. (We note $D_i = D_i(G_0, G_1, \ldots, G_{k-1}) \neq 0$ since $D_i = a_i D$, where $a_i \neq 0$ by Lemma 1 and $D \neq 0$ by the condition $\alpha_i \neq \alpha_j$ for $i \neq j$.) These assumptions imply the equality

$$\mu G_n = \mu \sum_{i=1}^{k} a_i \alpha_i^n = \sum_{i=1}^{s} p_i \alpha_i^{n-N} + \mu \sum_{i=s+1}^{k} a_i \alpha_i^n = \sum_{i=1}^{k} p_i \alpha_i^{n-N} + \delta_n \tag{15}$$

for $n \geq N$, where

$$p_i = \frac{D_i(P_N + q_0, P_{N+1} + q_1, \ldots, P_{N+k-1} + q_{k-1})}{D}$$

and

$$\delta_n = \mu \sum_{i=s+1}^{k} a_i \alpha_i^n - \sum_{i=s+1}^{k} p_i \alpha_i^{n-N} \to 0 \tag{16}$$

as $n \to \infty$, since $|\alpha_i| < 1$ for $i > s$. By an elementary property of determinants we get

$$p_i = \frac{D_i(P_N, P_{N+1}, \ldots, P_{N+k-1})}{D} + \frac{D_i(q_0, q_1, \ldots, q_{k-1})}{D},$$

which, together form (2), implies the equation

$$\sum_{i=1}^{k} p_i \alpha_i^n = H_n + K_n, \tag{17}$$

where (H_n) and (K_n), $n=0, 1, 2, \ldots$, are linear recurrences with common characteristic polynomial $g(x)$ and with initial values $H_0=P_N$, $H_1=P_{N+1}, \ldots, H_{k-1}=P_{N+k-1}$ and $K_0=q_0$, $K_1=q_1, \ldots, K_{k-1}=q_{k-1}$, respectively. The terms of the sequence (H_n) are integers and the numbers K_n $(n=0, 1, 2, \ldots)$ are rationals. The denominators of the numbers K_n, by the definition of the sequence (K_n), are bounded from above, since the initial values are fixed rational numbers and the numbers $A_1, A_2, \ldots, A_k$ are integers. From this follows, using (15), (16) and (17), that the sequence $\{\mu G_n\}$ can have only finitely many limit points, which completes the proof of Theorem 1.

Proof of Corollary 1. The necessity of the conditions follows from the proof of Theorem 1. Namely if $\lim_{n\to\infty} \{\mu G_n\}=q$ then $q_0=q_1=\ldots=q_{k-1}=q$ and the form of q follows from (12). Thus we have only to prove that q is a single limit point if the conditions hold for μ and q.

Let us suppose that μ and q have the forms given in Corollary 1. By (15), (16) and (17) it is sufficient to prove that $\lim_{n\to\infty} \{K_n\}=\{q\}$, where (K_n), $n=0, 1, 2, \ldots$, is a linear recurrence defined by characteristic polynomial $g(x)$ and initial values $K_0=K_1=\ldots=K_{k-1}=q$. If (R_n), $n=0, 1, 2, \ldots$, is a linear recurrence defined by characteristic polynomial $g(x)$ and initial values $R_0=R_1=\ldots=R_{k-1}=1$ then

$$K_n = qR_n$$

for $n\geqq 0$ since $D_i(q, q, \ldots, q)=qD_i(1, 1, \ldots, 1)$. Let $r=\left(\sum_{i=1}^{k} A_i\right)-1$. We can suppose $0<q=\frac{t}{r}<1$ and we shall prove that $\left\{\frac{t}{r} R_n\right\}=t/r$ for every $n\geqq 0$. This statement is obvious if R_n is of the form $mr+1$, where m is an integer, thus it is sufficient to prove

$$R_n \equiv 1 \pmod r \tag{18}$$

for $n\geqq 0$. However, (18) clearly holds for $n=0, 1, 2, \ldots, k-1$, and if

$$R_{n-1} \equiv 1 \pmod r$$
$$R_{n-2} \equiv 1 \pmod r$$
$$\vdots$$
$$R_{n-k} \equiv 1 \pmod r$$

for some integer $n\geqq k$ then, multiplying the i-th congruence by A_i $(i=1, 2, \ldots, k)$ and summing them, we obtain

$$R_n = \sum_{i=1}^{k} A_i R_{n-i} \equiv \sum_{i=1}^{k} A_i \equiv 1 \pmod r,$$

which completes the proof of Corollary 1.

NOTE. In case $\sum_{i=1}^{k} A_i=1$ similar conditions hold for the existence of $\lim_{n\to\infty} \{\mu G_n\}$. Now one of the roots of $g(x)$ is $x=1$, say $\alpha_s=1$. The following result can be proved similarly as Corollary 1: The sequence $\{\mu G_n\}$ is convergent if and only if μ satisfies simultaneously the equations

$$\mu = \frac{D_i(P_N, P_{N+1}, \dots, P_{N+k-1})}{\alpha_i^N \cdot D_i(G_0, G_1, \dots, G_{k-1})}$$

for all indices i with $1 \leqq i < s$, where $N, P_N, P_{N+1}, \dots, P_{N+k-1}$ are integers. If μ satisfies the condition and $q=\mu a_s+p_s$, where

$$a_s = \frac{D_s(G_0, G_1, \dots, G_{k-1})}{D}$$

and

$$p_s = \frac{D_s(P_N, P_{N+1}, \dots, P_{N+k-1})}{D},$$

then $\lim_{n\to\infty} \{\mu G_n\} = \{q\}$.

PROOF OF THEOREM 2. Let (g_n), $n=1, 2, 3, \dots$, be an increasing sequence of positive integers satisfying the conditions $g_i-g_{i-1} \geqq i$ and $g_i/g_{i-1}<2$ for $i=2, 3, \dots$. Let (c_n), $n=1, 2, \dots$, be also a sequence of non-negative integers such that $c_i \leqq i$ for every $i \geqq 1$ and every natural number occurs infinite frequently in the sequence. Such sequences can be easily constructed. For example $g_n=F_{n+3}$, where F_i is the i-th Fibonacci number, and $(c_n)=(1, 2, 1, 2, 3, 1, 2, 3, 4, \dots)$, where we wrote the first two, the first three, the first four, ... natural numbers, satisfy the conditions.

Let (G_n) be a sequence satisfying the conditions of the theorem and let $\alpha=\alpha_1$. The number

$$\mu = \sum_{i=1}^{\infty} c_i \alpha^{-g_i} \tag{19}$$

exists since α is a PV number and so, by $|\alpha|>1$, the sum is convergent. The terms of the sequence (G_n) can be written in the form

$$G_n = a\alpha^n + b_n, \tag{20}$$

where $|b_n| = \left|\sum_{i=2}^{k} a_i \alpha_i^n\right| \to 0$ as $n\to\infty$, since $|\alpha_i|<1$ for $2 \leqq i \leqq k$, and $a=a_1 \neq 0$ by Lemma 1.

We shall show that $\{ta\alpha^e\}$ is an accumulation point of the sequence (μG_n) modulo 1 for every integer $t \geqq 1$ and $e \geqq 0$. It is sufficient to prove that for any $\varepsilon>0$ there are infinitely many integer n such that

$$|\{\mu G_n\} - \{ta\alpha^e\}| < \varepsilon.$$

Let r be an index for which $c_r=t$. We can suppose that $r>e$ and $g_r>e$. For $n=g_r+e$ we can write

$$\mu G_n = \mu a\alpha^n + \mu b_n = \sum_{i=1}^{r-1} c_i a\alpha^{n-g_i} + ta\alpha^e + \sum_{i=r+1}^{\infty} c_i a\alpha^{-(g_i-n)} + \mu b_n. \tag{21}$$

By (20) we have

$$\sum_{i=1}^{r-1} c_i a\alpha^{n-g_i} = \sum_{i=1}^{r-1} c_i G_{n-g_i} - \sum_{i=1}^{r-1} c_i b_{n-g_i} = P_n + \omega_n,$$

where P_n is an integer and, using the notations $b = \max(|a_2|, |a_3|, \ldots, |a_k|)$ and $\beta = \max(|\alpha_2|, |\alpha_3|, \ldots, |\alpha_k|)$,

$$|\omega_n| = \left|\sum_{i=1}^{r-1}\left(c_i \sum_{j=2}^{k} a_j \alpha_j^{n-g_i}\right)\right| \leqq \sum_{i=1}^{r-1} kbc_i\beta^{n-g_i} \leqq$$

$$\leqq kb\beta^{n-g_{r-1}} \sum_{i=1}^{r-1} i\beta^{g_{r-1}-g_i} \leqq kb\beta^r \sum_{i=0}^{r-2} r\beta^i = kbr\beta^r \frac{1-\beta^{r-1}}{1-\beta} \to 0$$

as $r \to \infty$ since $\beta < 1$. Furthermore $\mu b_n \to 0$ as $r \to \infty$ and

$$\sum_{i=r+1}^{\infty} c_i a\alpha^{-(g_i-n)} \leqq a\alpha^{-(g_{r+1}-n)} \sum_{i=r+1}^{\infty} i\alpha^{-(g_i-g_{r+1})} \leqq a\alpha^{-(r-e)}\left(r+1+\sum_{i=r+2}^{\infty} i\alpha^{-i}\right) =$$

$$= a\alpha^e\left(r\alpha^{-r} + \alpha^{-r} + \alpha^{-r}\sum_{i=r+2}^{\infty} i\alpha^{-i}\right) \to 0$$

as $r \to \infty$ since the sum is convergent, therefore

$$\{\mu G_n\} - \{ta\alpha^e\} \to 0$$

as $r \to \infty$ and so as $n \to \infty$. But, by the conditions for sequence (c_n), there are infinitely many integer r such that $c_r = t$, thus $\{ta\alpha^e\}$ is an accumulation point of the sequence $\{\mu G_n\}$, indeed.

The number $a\alpha^e$ is irrational for some e since α is an irrational PV number, therefore the sequence $(ta\alpha^e)$, $t = 2, 3, 4, \ldots$, is uniformly distributed modulo 1. This implies that the numbers $\{\mu G_n\}$, $n = 1, 2, 3, \ldots$, are everywhere dense in the interval $[0, 1)$.

We show that the sequence (μG_n) is not uniformly distributed modulo 1.

Let ε and δ be numbers with conditions $0 < \delta < \frac{1}{4}$ and $0 < \varepsilon < \frac{1}{8}$, and let $n = g_r + e$, where $e < r$. By (21), similarly as above, we get

$$\mu G_n = P_n + \varepsilon_n + c_r a\alpha^e + c_{r+1} a\alpha^{-(g_{r+1}-n)},$$

where P_n is an integer and $\varepsilon_n \to 0$ as $n \to \infty$. If $e = n - g_r > \delta r$ and $g_{r+1} - n > \delta r$ then

$$\|c_r a\alpha^e\| = \left\|c_r G_e - c_r \sum_{i=2}^{k} a_i \alpha_i^e\right\| \leqq \|c_r kb\beta^e\| \leqq \|rkb\beta^{\delta r}\| = \|kbr(\beta^\delta)^r\| < \frac{\varepsilon}{4}$$

and

$$\|c_{r+1} a\alpha^{-(g_{r+1}-n)}\| \leqq \|2ar(\alpha^{-\delta})^r\| < \frac{\varepsilon}{4}$$

if n and so if r is sufficiently large. Thus we obtain that $\{\mu G_n\}$ lies in one of the intervals $[0, \varepsilon)$ and $(1-\varepsilon, 1)$ if $g_r \leqq n < g_{r+1}$, $\min(n - g_r, g_{r+1} - n) > \delta r$, and $n > n(\varepsilon)$.

Let $A(N, \varepsilon)$ be the number of the terms of sequence (μG_n) for which $n \leqq N$ and $\{\mu G_n\}$ lies in one of the intervals $[0, \varepsilon)$ and $(1-\varepsilon, 1)$. By the above results, if $N > n(\varepsilon)$, $g_r \leqq N < g_{r+1}$, and $g_{v-1} < n(\varepsilon) \leqq g_v$ then

$$A(N, \varepsilon) \geqq \sum_{i=v}^{r-1} (g_{i+1} - g_i)(1-2\delta) = (1-2\delta)(g_r - g_v)$$

and

$$\frac{A(N, \varepsilon)}{N} \geqq (1-2\delta)\left(\frac{g_r}{N} - \frac{g_v}{N}\right) > (1-2\delta)\left(\frac{g_r}{g_{r+1}} - \frac{g_v}{N}\right) \geqq \left(1-\frac{1}{2}\right)\left(\frac{1}{2} - \frac{g_v}{N}\right) = \frac{1}{4} - \frac{g_v}{2N},$$

where $(g_v)/(2N) \to 0$ as $N \to \infty$, since g_v is a fixed number. From this follows that the sequence (μG_n) cannot be uniformly distributed modulo 1, namely ε can be arbitrary small.

By the conditions, the sequence (c_n) has infinitely many terms equal to one, say $c_{i_1} = c_{i_2} = \ldots = 1$. Let (γ_n), $n = 1, 2, 3, \ldots$, be a sequence of integers such that $\gamma_n = c_n$ for $n \neq i_j$ $(j = 1, 2, \ldots)$ and $\gamma_n = 0$ or 1 for $n = i_1, i_2, \ldots$. There are uncountable many such sequences (γ_n) and each of them implies a real number μ using the sequence (γ_n) in (19) instead of (c_n). It can be easily seen that the μ's are distinct if $\gamma_n \neq 0$ for $n < n_0$, where n_0 is a constant depending only on α. Furthermore, by the argument applied above, we can show that each of the μ's has the property detailed in the theorem. Thus the theorem is proved.

PROOF OF COROLLARY 2. Let (g_n) be the sequence defined in the proof of Theorem 2 and let $c_n = 1$ for $n = 1, 2, 3, \ldots$. Construct a real number μ by these sequences as in the proof of Theorem 2. Similarly, as above, we can see that for any integer e the numbers $\{a\alpha^e\}$ and $\{a\alpha^{-e}\}$ are accumulation points of the sequence $\{\mu G_n\}$, since infinitely many integer n is of the form $n = g_r + e$ or $n = g_{r+1} - e$. Furthermore it is easy to prove that the numbers $\{a\alpha^e\}$ and $\{a\alpha^{-e}\}$ are all the limit points. But $\lim_{e \to \infty} (a\alpha^{-e}) = 0$ and

$$a\alpha^e = G_e + \varepsilon_e,$$

where $\varepsilon_e \to 0$ as $e \to \infty$, therefore the limit points cannot be everywhere dense in the interval $[0, 1)$.

Similarly as in the proof of Theorem 2, uncountable many distinct μ can be constructed having such properties.

REFERENCES

[1] BEUKERS, F., Fractional parts of powers of rationals, *Math. Proc. Cambridge Philos. Soc.* **90** (1981), 13—20.

[2] BROWN, J. L. JR. and DUNCAN, R. L., Modulo one uniform distribution of certain Fibonacci-related sequences, *Fibonacci Quart.* **10** (1972), 277—280, 294. *MR* **46** # 3426.

[3] CHOQUET, G., θ-fermés et dimension de Hausdorff. Conjectures de travail. Arithmétique des θ-cycles (où $\theta = 3/2$), *C. R. Acad. Sci. Paris. Sér. I. Math.* **292** (1981), 339—344. *MR* **82c**: 10057.

[4] OCAGNE, M. M. D'., Mémoire sur les suites récurrentes, *J. École Polytechn.* **64** (1894), 151—224.

[5] GERDES, W., Generalized tribonacci numbers and their convergent sequences, *Fibonacci Quart.* **16** (1978), 269—275. *MR* **80a**: 10019.

[6] GERDES, W., Convergent generalized Fibonacci sequences, *Fibonacci Quart.* **15** (1977), 156—160. *MR* **56**#237.
[7] GREGORY, M. B. and METZGER, J. M., Fibonacci sine sequences, *Fibonacci Quart.* **16** (1978), 119—120. *MR* **58**#16588.
[8] KISS, P., On common terms of linear recurrences, *Acta Math. Acad. Sci. Hungar.* **40** (1982), 119—123.
[9] KISS, P., On some properties of linear recurrences, *Publ. Math. Debrecen* **30** (1983), 273—281.
[10] KUIPERS, L. and NIEDERREITER, H., *Uniform distribution of sequences,* Pure and Applied Mathematics, John Wiley et Sons, New York—London—Sydney—Toronto, 1974. *MR* **54**# 7415.
[11] KUIPERS, L. and SHIUE, J. S., Remark on a paper by Duncan and Brown on the sequence of logarithms of certain recursive sequences, *Fibonacci Quart.* **11** (1973), 292—294. *MR* **48**#11025.
[12] LEVIN, M. B. and SPARLINSKIĬ, I. E., Uniform distribution of fractional parts of recurrent sequences, *Uspehi Mat. Nauk.* **34** (1979), no. 3 (207), 203—204 (in Russian). *MR* **80k**: 10046.
[13] MENDÈS-FRANCE, M., A characterization of Pisot numbers, *Mathematika* **23** (1976), 32—34. *MR* **54**#7394.
[14] MOLNÁR, S., Sine sequence of second order linear recurrences, *Period. Math. Hungar.* **14** (1983), 259—267.
[15] PERELLI, A. and ZANNIER, U., Arithmetic properties of certain recurrence sequences (in preparation).
[16] PISOT, C., Répartition (mod 1) des puissances successives des nombres réels, *Comment Math. Helv.* **19** (1946), 153—160. *MR* **8**—194.
[17] VIJAYARAGHAVAN, T., On the fractional parts of the powers of number. II, *Proc. Cambridge Philos. Soc.* **37** (1941), 349—357. *MR* **3**—274.

(Received March 17, 1982)

HO SI MINH TANÁRKÉPZŐ FŐISKOLA
MATEMATIKAI TANSZÉKE
LEÁNYKA U. 4
POSTAFIÓK 43
H—3301 EGER
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 129—133

THE VOLUME OF SYMMETRIC DOMAINS, THE KOECHER GAMMA FUNCTION AND AN INTEGRAL OF SELBERG

ADAM KORÁNYI[1]

To the memory of my teacher László Rédei

The volumes of the bounded symmetric domains corresponding to Lie groups of classical type were found by Hua [4] via case-by-case computations. A formula for the volume of arbitrary bounded symmetric domains was proved, independently of classification, in [7]. The proof in [7], however, is rather complicated: It makes use of results from [6] tying the volume to Koecher's generalized Gamma function Γ^*, and then it uses Gindikin's evaluation of Γ^* which is based on an extensive theory of homogeneous (not necessarily symmetric) cones [2].

The first result of the present paper is a simple direct proof of the volume formula which does not even use Γ^*. The second result is a direct computation of Γ^* for symmetric cones, independent of Gindikin's difficult general theory. The key to both results is the evaluation of certain integrals which for many years I did not know how to do. It was when I showed them to R. Askey, to whom I express here my warmest thanks, that he immediately recognized them as special cases or immediate consequences of an integral computed by Selberg [10] in 1944.

§ 1. The Selberg integral

For easy reference we write down Selberg's integral [10] and two of its consequences (see e.g. [1] where a number of further comments about these integrals can be found):

$$\int_0^1 \cdots \int_0^1 \prod_{i=1}^{l} t_i^{x-1}(1-t_i)^{y-1} \Big| \prod_{1\le i<j\le l} (t_i-t_j)\Big|^{2z} dt_1 \dots dt_l = \tag{1.1}$$
$$= \prod_{j=1}^{l} \frac{\Gamma\big(x+(j-1)z\big)\Gamma\big(y+(j-1)z\big)\Gamma(jz+1)}{\Gamma\big(x+y+(l+j-2)z\big)\Gamma(z+1)}.$$

Substituting $t_i \to t_i/y$ and letting $y\to\infty$ it follows from Stirling's formula that

$$\int_0^\infty \cdots \int_0^\infty \prod_{i=1}^{l} t_i^{x-1}e^{-t_i} \Big| \prod_{1\le i<j\le l} (t_i-t_j)\Big|^{2z} dt_1 \dots dt_l = \tag{1.2}$$
$$= \prod_{j=1}^{l} \frac{\Gamma\big(x+(j-1)z\big)\Gamma(jz+1)}{\Gamma(z+1)}.$$

---

[1] Partially supported by the National Science Foundation.

1980 *Mathematics Subject Classification.* Primary 32M15; Secondary 22E30.
Key words and phrases. Symmetric domains, Cartan domains, self-dual homogeneous cones.

Setting in (1.1) $y=x$ and $2t_i=1+s_i(2x)^{-1/2}$, Stirling's formula gives

$$(1.3)\quad \frac{1}{(2\pi)^{l/2}}\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty}\exp\Big(-\frac{1}{2}\sum_{i=1}^{l}t_i^2\Big)\Big|\prod_{1\le i<j\le l}(t_i-t_j)\Big|^{2z}dt_1\dots dt_l=\prod_{j=1}^{l}\frac{\Gamma(jz+1)}{\Gamma(z+1)}.$$

§ 2. The volume formula

We denote by D an irreducible bounded symmetric domain in the canonical Harish-Chandra realization. As well known (e.g. [9, Ch. II, §§3—4]), D can be regarded as imbedded in the subspace $\mathfrak{p}$ of the Lie algebra $\mathfrak{g}$ of its holomorphic automorphisms; $\mathfrak{p}$ has the structure of a complex Euclidean space and the subgroup K of elements in G fixing the point 0 acts on it by unitary transformations via the adjoint representation. $\mathfrak{p}$ contains a Cartan subalgebra $\mathfrak{a}$ of the pair (G, K); this is a totally real l-dimensional subspace of $\mathfrak{p}$ such that $\mathfrak{p}=K\mathfrak{a}$. There is a natural orthonormal basis in $\mathfrak{a}$ such that, denoting the coordinates by $t_1, \dots, t_l$, D is equal to the image under K of the cube $0\le t_i<1$ (in fact, the simplex $1>t_1>t_2>\dots>t_l>0$ is a fundamental domain for the action of K on D). The restricted positive roots of $\mathfrak{g}$ are then the following: $t_i\pm t_j$ $(i<j)$ all with multiplicity a, $2t_i(1\le i\le l)$ with multiplicity 1, and t_i with multiplicity $2b$. The integers a, b are invariants of D, and together with the rank l they determine D completely. The (complex) dimension of D is (cf. [7])

$$(2.1)\qquad n=l(l-1)\frac{a}{2}+l(b+1).$$

Let $\Pi(t)$ be the product of all positive roots, i.e.

$$\Pi(t)=2^l\prod_{1\le i<j\le l}(t_i^2-t_j^2)^a\prod_{i=1}^{l}t_i^{2b+1}.$$

It is well-known [3, p. 380] that the ratio of the volume elements of $\mathfrak{p}$ and $\mathfrak{a}$ at a point of $\mathfrak{a}$ with coordinates $t=(t_1, \dots, t_l)$ is $c|\Pi(t)|$ with some constant c. (Another way to say this, since it is known that the K-orbit of a regular element meets the positive octant $l!$ times, is that the $(n-l)$-dimensional measure of the K-orbit of t is $(l!)c|\Pi(t)|$.) It follows that (writing $dt=dt_1\dots dt_l$)

$$(2.2)\qquad \mathrm{Vol}(D)=c\int_0^1\dots\int_0^1|\Pi(t)|\,dt.$$

To determine the value of c we can proceed as follows. Writing, for $r>0$,

$$J(r)=\int_{\substack{\Sigma t_i^2<r^2\\ t_1,\dots,t_l>0}}|\Pi(t)|\,dt$$

we have, since K operates by unitary transformations, that $cJ(r)$ is the volume of the ball of radius r in $\mathfrak{p}$. Hence

$$(2.3)\qquad cJ(r)=cJ(1)r^{2n}=\frac{\pi^n}{\Gamma(n+1)}r^{2n}.$$

On the other hand, defining

$$I = \int_{\mathbf{R}^l} e^{-\Sigma t_i^2} |\Pi(t)| \, dt \tag{2.4}$$

we have

$$I = \int_0^\infty e^{-r^2} dJ(r) = J(1) n \int_0^\infty e^{-r^2} r^{n-1} dr = J(1)\Gamma(n+1). \tag{2.5}$$

From (2.3) and (2.5) it follows that $c = \pi^n I^{-1}$. By the change of variable $t_i^2 \to t_i$, (2.4) reduces to the integral (1.2). It follows that

$$c = \pi^n \prod_{j=1}^{l} \frac{\Gamma\left(\frac{a}{2}+1\right)}{\Gamma\left(b+1+(j-1)\frac{a}{2}\right)\Gamma\left(j\frac{a}{2}+1\right)}.$$

With the same change of variable the integral in (2.2) becomes a special case of (1.1). So we obtain

$$\mathrm{Vol}(D) = \pi^n \prod_{j=1}^{l} \frac{\Gamma\left(1+(j-1)\frac{a}{2}\right)}{\Gamma\left(b+2+(l+j-2)\frac{a}{2}\right)}$$

in agreement with the result in [7].

§ 3. The Koecher Gamma function

By methods similar to those of §2 it is easy to evaluate the Koecher Gamma function for symmetric cones.

The symmetric cones Ω are in one-to-one correspondence with the symmetric domains D "of tube type", i.e. those for which $b=0$. As one see from the discussion in [6, §5] and [8, §2], Ω can be realized as an open cone in a real form $\mathrm{Re}\,\mathfrak{p}$ of $\mathfrak{p}$. One has $\mathfrak{a} \subset \mathrm{Re}\,\mathfrak{p}$ and there is a subgroup L^0 of K such that $\mathrm{Re}\,\mathfrak{p} = L^0\mathfrak{a}$; Ω is the image under L^0 of the positive octant in $\mathfrak{a}$. The action of L^0 on $\mathrm{Re}\,\mathfrak{p}$ is isometrically isomorphic with its adjoint action on the transvection space (denoted $i\mathfrak{q}$ in [8]) of a symmetric space $K^*/L^0 \cong \Omega$. Therefore the ratio of the volume elements of $\mathrm{Re}\,\mathfrak{p}$ and $\mathfrak{a}$ is given by $c_0\Pi_0(t)$ where c_0 is a constant to be determined, and

$$\Pi_0(t) = \prod_{i<j} (t_i - t_j)^a$$

is the product of positive roots of (K^*, L^0) transferred to $\mathrm{Re}\,\mathfrak{p}$.

The Koecher norm function of Ω is the L^0-invariant extension of $(t_1 \ldots t_l)^{n/l}$ to $\mathrm{Re}\,\mathfrak{p}$ [5]. The generalized Gamma function Γ^* of Ω is therefore given by

$$\Gamma^*(s) = c_0 \int_0^\infty \ldots \int_0^\infty e^{-\Sigma t_i} \Pi t_i^{\frac{n}{l}(s-1)} |\Pi_0(t)| \, dt. \tag{3.1}$$

To determine c_0 we set, analogously to §2,

$$J_0(r) = \int\limits_{\substack{\Sigma t_i^2 < r^2 \\ t_1, \ldots, t_l > 0}} |\Pi_0(t)| \, dt$$

and note that now $c_0 J_0(r)$ is the volume of the ball of radius r in Re $\mathfrak{p}$:

$$c_0 J_0(r) = c_0 J_0(1) r^n = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)} r^n. \tag{3.2}$$

Defining

$$I_0 = \int\limits_{\mathbf{R}^l} e^{-\Sigma t_i^2} |\Pi_0(t)| \, dt \tag{3.3}$$

we have

$$I_0 = \int\limits_0^\infty e^{-r^2} \, dJ_0(r) = J_0(1) \Gamma\left(\frac{n}{2}+1\right). \tag{3.4}$$

From (3.2) and (3.4) we get $c_0 = \pi^{n/2} I_0^{-1}$. The substitution $t_i \to t_i/\sqrt{2}$ carries (3.3) into (1.3); using (2.1) and remembering $b=0$ this gives

$$c_0 = (2\pi)^{\frac{n-l}{2}} \prod_{j=1}^{l} \frac{\Gamma\left(\frac{a}{2}+1\right)}{\Gamma\left(j\frac{a}{2}+1\right)}.$$

Substituting this into (3.1) and using (1.2) we obtain

$$\Gamma^*(s) = (2\pi)^{\frac{n-l}{2}} \prod_{j=1}^{l} \Gamma\left(\frac{n}{l}(s-1) + (j-1)\frac{a}{2}\right)$$

which agrees with the result in [7].

REFERENCES

[1] Askey, R., Some basic hypergeometric extensions of integrals of Selberg and Andrews, *SIAM J. Math. Anal.* **11** (1980), 938—951. *MR* **82e**: 33002.

[2] Gindikin, S. G., Analysis in homogeneous domains, *Uspehi Mat. Nauk* **19** (1964), no. 4, 3—92 (in Russian). *MR* **30** # 2167. English translation: *Russian Math. Surveys* **19** (1964), 1—90.

[3] Helgason, S., *Differential geometry and symmetric spaces*, Pure and Applied Mathematics Vol. XII, Academic Press, New York—London, 1962. *MR* **26** # 2986.

[4] Hua, L. K., *Harmonic analysis of functions of several complex variables in the classical domains*, American Mathematical Society Translations of Mathematical Monographs, American Mathematical Society, Providence, R. I., 1963. *MR* **30** # 2162; see also *MR* **22** # 11145 and *MR* **23** # A3277.

[5] Koecher, M., Positivitätsbereiche im R^n, *Amer. J. Math.* **79** (1957), 575—596. *MR* **19**—867.

[6] Korányi, A., The Poisson integral for generalized half-planes and bounded symmetric domains, *Ann. of Math.* (2) **82** (1965), 332—350. *MR* **34** # 371.

[7] Korányi, A., Analytic invariants of bounded symmetric domains, *Proc. Amer. Math. Soc.* **19** (1968), 279—284. *MR* **37** # 463.
[8] Korányi, A. and Vági, S., Rational inner functions on bounded symmetric domains, *Trans. Amer. Math. Soc.* **254** (1979), 179—193. *MR* **80h**: 32062.
[9] Satake, I., *Algebraic structures of symmetric domains*, Publications of the Mathematical Society of Japan, 14, Kanô Memorial Lectures, 4, Iwanami Shoten, Tokyo; Princeton University Press, Princeton, N. J., 1980. *MR* **82i**: 32003.
[10] Selberg, A., Bemerkninger om et Multipelt Integral, *Norsk Mat. Tidsskr.* **26** (1944), 71—78 (Remarks on a multiple integral, in Norwegian). *MR* **8**—269.

(Received April 12, 1982)

MATHEMATICS DEPARTMENT
WASHINGTON UNIVERSITY
CAMPUS BOX 1146
ST. LOUIS, MO 63130
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 135—136

SOME REMARKS ON GROUPS WITH GENERALIZED QUATERNION SYLOW SUBGROUPS

L. HÉTHELYI

To the memory of László Rédei

In this short note we prove some simple results concerning 2-nilpotence and solubility of finite groups with generalized quaternion Sylow 2-subgroups.

THEOREM 1. *Let G be a finite group,* $S \in Syl_2(G)$, *S a generalized quaternion group,* $|S| \geqq 16$. *Then G has a normal 2-complement if and only if G does not involve* S_4.

PROOF. (a) If G has a normal 2-complement then it is trivial that S_4 is not involved in G.

(b) Suppose that G does not have a normal 2-complement, then by virtue of Ito's theorem [4], there is a subgroup $T=LQ$, where L is a normal Sylow 2-subgroup of T, Q is a cyclic Sylow q-subgroup of T $(q \neq 2)$, Q induces an automorphism group of order q on L. Our assumptions allow the only possibility that L is a quaternion group of order 8 and $q=3$.

Let S be a Sylow 2-subgroup of G containing L. Then $|N_S(L)/C_S(L)|=8$, so $24 \mid |N_G(L)/C_G(L)|$. Hence $N_G(L)/C_G(L) \cong \mathrm{Aut}\, L \cong S_4$.

REMARK 1. If $|S|=8$ then it is trivial that G has a normal 2-complement if and only if G does not involve A_4.

REMARK 2. In fact, we can prove more: If G does not have a normal 2-complement, then it involves a certain group of order 48, namely $N_G(L)/O_{2'}(C_G(L))$. This group appears as a stabilizer in a sharply double transitive group of degree 49 (see [3], p. 391, case III).

REMARK 3. The fact if S_4 is involved plays a crucial role in numerous investigations (see e.g. [2]).

THEOREM 2. *Let G be a finite group. Let* $S \in Syl_2(G)$, *S dihedral,* $|S| \geqq 8$. *Then G has a normal 2-complement if and only if G does not involve* S_4.

PROOF. Similar to the proof of Theorem 1.

THEOREM 3. *Let G be a finite soluble group. Let* $S \in Syl_2(G)$, *S a generalized quaternion group,* $|S| > 16$. *Then there is a normal 2-complement in G.*

1980 *Mathematics Subject Classification.* Primary 20D20; Secondary 20F16.
Key words and phrases. Generalized quaternion Sylow 2-subgroups, normal 2-complement, finite soluble group.

PROOF. Let G be a counterexample of minimal order. $O_{2'}(G)=E$ can be supposed. Then $|O_2(G)|=8$ and $O_2(G)$ is quaternion. As $O_2(G)\lhd S$ thus $|S|=16$.

THEOREM 4. *Let G be a finite soluble group. Let $S\in Syl_2(G)$, S dihedral, $|S|>8$. Then there is a normal 2-complement in G.*

PROOF. Similar to the proof of Theorem 3.

THEOREM 5. *Let G be a finite group, $S\in Syl_2(G)$, S a generalized quaternion group. If $|O_2(G)|>2$ then G is soluble.*

PROOF. By Feit and Thompson [1], it is enough to prove that G is 2-soluble.
Let G be a counterexample of minimal order. $O_{2'}(G)=E$ can be supposed. $|\pi(G)|\geqq 3$ can also be supposed. Let $P\in Syl_p(G)$, $p\neq 3$, $p\neq 2$, then $P\leqq C_G(O_2(G))$. Moreover, $Syl_2(C_G(O_2(G)))$ is cyclic thus $O_{2'}(C_G(O_2(G)))\neq E$. However, $O_{2'}(C_G(O_2(G)))$ char $C_G(O_2(G))$ and $C_G(O_2(G))\lhd G$. Thus $O_{2'}(G)\neq E$ which is a contradiction.

COROLLARY 1. *Let G be a finite group, $S\in Syl_2(G)$. Let S be a generalized quaternion group. Let us suppose that G has a homomorphic image $\bar{G}$ such that $|O_2(\bar{G})|>2$. Then G is soluble.*

PROOF. Let N denote the kernel of this homomorphism. If N is of odd order or N is a 2-group then G is soluble by Theorem 5.
If N is of even order (but N is not a 2-group) then as $|O_2(\bar{G})|>2$, $Syl_2(N)$ is cyclic thus N has a normal 2-complement K. Now G/K fulfils the conditions of Theorem 5 thus G/K is soluble. Therefore G is soluble by the solvability of groups of odd order.

COROLLARY 2. *Let G be a finite group. Let $S\in Syl_2(G)$, S a generalized quaternion group. If there is a soluble normal subgroup N of G such that $4\,|\,|N|$ then G is soluble.*

PROOF. If $O_{2'}(N)\neq E$ then let $\bar{G}$ denote $G/O_{2'}(N)$, $\bar{N}$ denote $N/O_{2'}(N)$. Then $\bar{N}\lhd\bar{G}$, $4\,|\,|\bar{N}|$. Thus $\bar{G}$ is soluble by induction and so G is soluble, too. If $O_{2'}(N)=E$ then $|O_2(N)|>2$. Hence $|O_2(G)|>2$, thus G is soluble by Theorem 5.

REFERENCES

[1] FEIT, W. and THOMPSON, J. G., Solvability of groups of odd order, *Pacific J. Math.* (1963), 775—1029. *MR* **29** # 3538.
[2] GLAUBERMAN, G., Weakly closed elements of Sylow subgroups, *Math. Z.* **107** (1968), 1—20. *MR* **40** # 4372.
[3] HALL, M., *The theory of groups*, The Macmillan Co., New York, N. Y., 1959. *MR* **21** # 1996.
[4] ITO, N., Note on (*LM*)-groups of finite orders, *Kōdai Math. Sem. Rep.* (1951), 1—6. *MR* **13**—317.

(Received April 14, 1982)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR
FILOZÓFIA TANSZÉK
RÁKÓCZI ÚT 5
H—1088 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 137—141

ПЕРИОДИЧЕСКИЕ ГРУППЫ С ПОЧТИ РЕГУЛЯРНЫМ ИНВОЛЮТИВНЫМ АВТОМОРФИЗМОМ

В. В. БЕЛЯЕВ и Н. Ф. СЕСЕКИН

Доброй памяти Ласло Редеи посвящается

Автоморфизм φ произвольной группы G называется почти регулярным на G, если централизатор $C_G(\varphi)$ конечен. В 1972 году Шунков установил замечательный результат [1]: если периодическая группа G допускает инволютивный почти регулярный автоморфизм, то она почти разрешима и обладает полной частью. Хартли и Мейкснер в 1980 году показали, что в условиях теоремы Шункова группа G обладает нильпотентной подгруппой конечного индекса ступени $\leqq 2$ в G [2].

Продолжая исследования в этом направлении нами установлены следующие результаты.

Теорема А. *Пусть φ инволютивный почти регулярный автоморфизм периодической группы G. Тогда подгруппа $[G, \varphi]$ имеет конечный индекс в G, а ее коммутант конечен.*

Теорема Б. *Централизатор произведения двух инволютивных почти регулярных автоморфизмов периодической группы имеет конечный индекс во всей группе.*

Из теоремы А нетрудно получить результат Хартли и Мейкснера. Теорема Б показывает, что инволютивный почти регулярный автоморфизм периодической группы определяется на ней «почти однозначно».

§ 1

Здесь мы приведем некоторые определения и известные результаты, которые используются в данной работе.

Определение 1. Если $\varphi \in \mathrm{Aut}\, G$, то под $[G, \varphi]$ понимаем подгруппу из G, порожденную элементами вида $[g, \varphi] = g^{-1}\varphi(g)$ для $g \in G$. Известно, что $[G, \varphi]$ нормальна в G. Централизатор $C_G(\varphi)$ автоморфизма φ определяется формулой $C_G(\varphi) = \{x, x \in G | \varphi(x) = x\}$.

1980 *Mathematics Subject Classification.* Primary 20F50; Secondary 20E36.
Key words and phrases. Group, automorphism, centralizer, index, commutant, element, almost regular involutive automorphism, almost splitting group, φ-admissible subgroup, subgroup of finite order, locally finite group, almost equal automorphisms.

Определение 2. Автоморфизм φ группы G называется почти тождественным на G, если индекс $(G\colon C_G(\varphi))$ конечен. Автоморфизмы φ и ψ назовем почти равными, если $\varphi^{-1}\psi$ почти тождественен.

Определение 3. Подгруппа X называется FC-вложенной в группу Y, если для любого элемента y из Y индекс $(X\colon C_X(y))$ конечен.

Определение 4. Подгруппа X называется FC-подгруппой в группе Y, если для любого элемента x из X индекс $(Y\colon C_Y(x))$ конечен.

Определение 5. Группа называется квазиабелевой, если ее коммутант конечен.

Объединяя предложения 3.2 и 3.17 книги Кегеля и Верфрица [3] получим следующий результат: если периодическая разрешимая группа G содержит такой элемент x простого порядка p, что $C_G(x)$ удовлетворяет условию минимальности для p-подгрупп, то индекс $(G\colon O_{p'p}(G))$ конечен и сама группа G удовлетворяет условию минимальности для p-подгрупп. Отсюда и упомянутой выше теоремы Шункова вытекает:

Теорема 1. *Пусть φ инволютивный почти регулярный автоморфизм периодической группы G. Тогда G содержит такую разрешимую φ-допустимую подгруппу H конечного индекса в G, что $H=A\times B$ где $B=O(H)$, A делимая абелева 2-группа конечного ранга.*

Из леммы 0.3 лекций Гагена [4] легко следует:

Теорема 2. *Пусть φ p-автоморфизм локально конечной p'-группы G и H φ-допустимая нормальная в G подгруппа. Тогда $C_{G/H}(\varphi)=C_G(\varphi)H/H$.*

Теорема 3 [5]. *Почти абелева FC-группа квазиабелева.*

§ 2

Лемма 1. *Пусть φ почти регулярный автоморфизм конечного порядка локально конечной группы G. Тогда*

$$|G\colon [G,\varphi]| \leqq |C_G(\varphi)|.$$

Доказательство. Пусть сначала G конечная группа. Понятно, что $[g,\varphi]=[h,\varphi]$ тогда и только тогда, когда $gh^{-1}\in C_G(\varphi)$. Поэтому

$$|[G,\varphi]| \geqq \left|\{[g,\varphi] \mid g\in G\}\right| = |G\colon C_G(\varphi)| = \frac{|G|}{|C_G(\varphi)|}.$$

Отсюда

$$|C_G(\varphi)| \geqq \frac{|G|}{|[G,\varphi]|} = |G\colon [G,\varphi]|.$$

Допустим теперь, что G локально конечная группа, но $|G\colon [G,\varphi)|> >|C_G(\varphi)|$. Пусть n такое натуральное число, что $|G\colon [G,\varphi)|\geqq n>|C_G(\varphi)|$ и

$g_1, \ldots, g_n$ система элементов группы G несравнимых по модулю $[G, \varphi]$. Подгруппа $H=\langle g_1^{\varphi^i}, \ldots, g_n^{\varphi^i}; 0 \leqq i \leqq |\varphi|\rangle$ конечна и φ-допустима. Следовательно $|H: [H, \varphi]| \leqq |C_H(\varphi)|$. Отсюда

$$|H: [G, \varphi] \cap H| \leqq |H: [H, \varphi]| \leqq |C_H(\varphi)| \leqq |C_G(\varphi)|.$$

Так как $n > |C_G(\varphi)|$, то найдутся такие k и l, что $g_k^{-1} g_l \in [G, \varphi]$. Это противоречит построению элементов $g_1, \ldots, g_n$. Лемма 1 доказана.

Следствие 1. *Пусть φ и ψ почти регулярные инволютивные автоморфизмы периодической абелевой группы G. Тогда φ и ψ почти равны.*

Доказательство. Так как $\varphi([g, \varphi]) = \varphi(g)^{-1} g = [g, \varphi]^{-1}$ то для любого элемента x из $[G, \varphi]$ имеем $\varphi(x) = x^{-1}$. Аналогично $\psi(x) = x^{-1}$, если $x \in [G, \psi]$. В силу леммы 1, подгруппа $D = [G, \varphi] \cap [G, \psi]$ имеет конечный индекс в G. Но для любого x из D $\varphi\psi(x) = x$; следовательно $D \subseteqq C_G(\varphi\psi)$.

Лемма 2. *Пусть φ такой автоморфизм произвольной группы G, что $[G, \varphi]$ конечен; тогда*

$$|G: C_G(\varphi)| \leqq |[G, \varphi]|.$$

Доказательство. Предположим, что $|G: C_G(\varphi)| > |[G, \varphi]|$ и $x_1, \ldots, x_n$ система элементов из различных смежных классов разложения группы G по $C_G(\varphi)$, причем $n > |[G, \varphi]|$. Тогда найдутся такие индексы k и l, что $[x_k, \varphi] = [x_l, \varphi]$. Но тогда $x_k^{-1} x_l \in C_G(\varphi)$. Лемма 2 доказана.

Лемма 3. *Пусть φ автоморфизм группы G, H некоторая конечная φ-допустимая нормальная подгруппа из G. φ почти регулярен (почти тождественнен) на G тогда и только тогда, когда φ почти регулярен (почти тождественен) на G/H.*

Доказательство. Пусть φ почти регулярный автоморфизм группы G и $C_{G/H}(\varphi) = L/H$. Тогда $[L, \varphi] \subseteqq H$. По лемме 2 $|L: C_L(\varphi)| \leqq |[L, \varphi]| \leqq |H|$. Из конечности $C_L(\varphi)$ следует конечность L, а поэтому и $C_{G/H}(\varphi)$. Пусть теперь φ почти регулярен на G/H. Учитывая конечность H и включение $C_G(\varphi)H/H \subseteqq C_{G/H}(\varphi)$, получим конечность $C_G(\varphi)$.

Так как $|G/H: C_{G/H}(\varphi)| \leqq |G/H: C_G(\varphi)H/H|$, то из конечности $|G: C_G(\varphi)|$ следует почти тождественность φ на G/H.

Пусть, наконец, φ почти тождественнен на G/H и $C_{G/H}(\varphi) = L/H$. Тогда $[L, \varphi] \subseteqq H$ и $|G: L|$ конечен. Отсюда по лемме 2 $|L: C_L(\varphi)|$ конечен, но тогда и $|G: C_G(\varphi)|$ конечен. Лемма 3 доказана.

Лемма 4. *Пусть φ почти регулярный инволютивный автоморфизм периодической $2'$-группы G, A ее нормальная φ-допустимая абелева подгруппа. Если $[G, \varphi] = G$, то A является FC-вложенной подгруппой в G.*

Доказательство. Пусть x такой элемент из G, что $\varphi(x) = x^{-1}$. Тогда подгруппа $H = \langle x, A\rangle$ φ-допустима, A имеет конечный индекс в H и A по условию абелева; поэтому $C_A(\varphi) = C_H(\varphi)$ содержится в некоторой конечной подгруппе K нормальной в H. Так как по теореме 2 $C_{H/K}(\varphi) = 1$, то, как известно, H/K абелева; следовательно $H' \subseteqq K$ и H' конечен. Отсюда следует, что $C_A(x)$ имеет конечный индекс в A.

Пусть теперь g произвольный элемент из G, тогда найдутся такие элементы $g_1, \ldots, g_n$ из G, что $g=[g_1, \varphi]^{\varepsilon_1}\cdot \ldots \cdot [g_n, \varphi]^{\varepsilon_n}$, где $\varepsilon_i \in \{1, -1\}$. Если $x_i=[g_i, \varphi]^{\varepsilon_i}$, то $\varphi(x_i)=x_i^{-1}$ и, по доказанному, $C_A(x_i)$ имеет в A конечный индекс. Следовательно, и $C_A(g)$ имеет конечный индекс в A.

Лемма 5. *Пусть G разрешимая периодическая $2'$-группа, φ ее почти регулярный инволютивный автоморфизм. Если $[G, \varphi]=G$, то коммутант G' группы G конечен.*

Доказательство. Пусть сначала G двуступенно разрешима. Покажем, что G' является FC-подгруппой в G. Для произвольного элемента g из G' найдутся такие элементы $g_1, \ldots, g_n$ из G, что $g \in \langle g_1, \ldots, g_n\rangle'$. По лемме 4 G' FC-вложен в G, тогда $C_{G'}(g_i)$ имеет конечный индекс в G' для любого $i \in \{1, \ldots, n\}$. Отсюда вытекает, что подгруппа $\langle G', g_1, \ldots, g_n\rangle$ является почти абелевой FC-группой. По теореме 3 ее коммутант K конечен. Очевидно, K нормальна в G и содержит элемент g. Поэтому $C_G(g)$ имеет конечный индекс в G. Отсюда следует, что конечная подгруппа $C_{G'}(\varphi)=C_G(\varphi)$ может быть вложена в φ-допустимую конечную нормальную в G подгруппу N. По теореме 2 $C_{G/N}(\varphi)=1$ и, следовательно, G/N абелева. Но тогда $G' \subseteq N$.

Пусть теперь G разрешима и H последний нетривиальный член ряда коммутантов группы G. По индуктивному предположению G/H имеет конечный коммутант $(G/H)'=G'/H$. Так как H абелева и нормальна в G, то по лемме 4 H FC-вложена в G.

Таким образом G' является почти абелевой FC-группой. По теореме 3 G'' конечен. По доказанному выше двуступенно разрешимая группа G/G'' имеет конечный коммутант $(G/G'')'=G'/G''$. Отсюда G' будет конечным.

Лемма 6. *Пусть G периодическая группа, φ ее почти регулярный инволютивный автоморфизм. Тогда G обладает подгруппой конечного индекса коммутант которой конечен.*

Доказательство. В силу теоремы 1 группа G содержит такую разрешимую φ-допустимую подгруппу H конечного индекса в G, что $H=A\times B$, где $B=O(H)$ и A делимая абелева 2-группа конечного ранга. По лемме 1 подгруппа $[B, \varphi]$ имеет конечный индекс в B. Так как B является $2'$-группой и $[[b, \varphi], \varphi]= =[b, \varphi]^{-2}$, то $[[B, \varphi], \varphi]=[B, \varphi]$. Следовательно по лемме 5 $[B, \varphi]$ имеет конечный коммутант. Теперь $[B, \varphi]\times A$ является искомой подгруппой в G.

§ 3

Докажем сначала теорему Б. Пусть φ и ψ почти регулярные инволютивные автоморфизмы периодической группы G. По лемме 6 группа G почти квазиабелева. Поэтому совокупность H всех ее FC-элементов является подгруппой конечного индекса в G. Из теоремы 3 следует, что H имеет конечный коммутант H'. По лемме 3 автоморфизмы φ и ψ почти регулярны и на абелевой группе H/H'. Но тогда, в силу следствия к лемме 1 их произведение почти тождественно на H/H'. А теперь снова по лемме 3 $\varphi\psi$ почти тождественно на H, а потому и на G.

Переходим к доказательству теоремы А. Для инволютивного почти регулярного автоморфизма φ рассмотрим полупрямое произведение $\langle\varphi\rangle \leftthreetimes G$, определенное действием φ на группе G. Для любого элемента g из G коммутатор $[g, \varphi]=g^{-1}\varphi^{-1}g\varphi$ есть произведение почти регулярных на G инволютивных автоморфизмов $g^{-1}\varphi^{-1}g$ и φ. Согласно теореме Б их произведение $[g, \varphi]$ есть почти тождественный на G автоморфизм. Следовательно $C_G([g, \varphi])$ имеет конечный индекс в G. Отсюда следует, что $[G, \varphi]$ является FC-группой. Но по лемме 6 $[G, \varphi]$ почти квазиабелева и по теореме 3 она имеет конечный коммутант. Конечность индекса $|G: [G, \varphi]|$ следует из леммы 1.

ЛИТЕРАТУРА

[1] Шунков, В. П., О периодических группах с почти регулярной инволюцией (Šunkov, V. P., Periodic groups with an almost regular involution, in Russian), *Algebra i Logika* **11** (1972), 470—493, 496. *MR* **47** # 8706.

[2] Hartley, B. and Meixner, Th., Periodic groups in which the centralizer of an involution has bounded order, *J. Algebra* **64** (1980), 285—291. *MR* **81i**: 20049.

[3] Kegel, O. and Wehrfritz, B., *Locally finite groups,* North-Holland Mathematical Library, Vol. 3, North-Holland Publishing Co., Amsterdam—London; American Elsevier Publishing Co., New York, 1973. *MR* **57** # 9848.

[4] Gagen, T. M., *Topics in finite groups,* London Mathematical Society Lecture Notes Series, No. 16, Cambridge University Press, Cambridge—New York—Melbourne, 1976. *MR* **53** # 10910.

[5] Беляев, В. В., Группы типа Миллера—Морено (Beljaev, V. V., Groups of the Miller—Moreno type, in Russian), *Sibirsk. Mat. Ž.* **19** (1978), 509—514, 715. *MR* **58** # 28185.

(Поступила 30 апреля 1982)

УЛ. БАРДИНА, Д. 46, КВ. 265
SU—620146 СВЕРДЛОВСК 146

УЛ. ФРУНЗЕ, Д. 20, КВ. 218
SU—620142 СВЕРДЛОВСК Л - 142
SOVIET UNION

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 143—158

ISOPERIMETRY IN VARIABLE METRIC

ISTVÁN FÁRY and ENDRE MAKAI, JR.

Dedicated to the memory of Professor L. Rédei

1. If we put the sides of a polygonal domain in a new order, without changing their directions or lengths, they will surround a new domain of different area, in general, while the perimeter remains the same in *every* metric of the plane. Searching for a domain of largest area is the simplest instance of our isoperimetric problem in *variable* metric.

More precisely, we consider Euclidean metrics in R^2 which are compatible with its linear structure. Such a metric is defined by its unit circle which is an ellipse in the affine structure of R^2. Let $\mathcal{M}_E$ be the set of all such metrics. For given $\mathcal{L}\subset\mathcal{M}_E$, and for rectifiable curves J, K in R^2, we write

$$K \equiv J \pmod{\mathcal{L}}, \tag{1}$$

if J and K have the *same* length in every metric in $\mathcal{L}$. If $\mathcal{L}$ consists of a single metric, the classical isoperimetric theorem states that the circle of the metric in the equivalence class of J encloses the largest area. In this paper, we consider the other extreme case $\mathcal{L}=\mathcal{M}_E$, and prove

THEOREM 1. *Let J be a rectifiable Jordan curve in R^2. Then there is an up to translation unique centrally symmetric convex curve C_J, called symmetric convexification of J, such that:* 1° *the length of C_J is the same as the length of J in every Euclidean metric of R^2;* 2° *the area enclosed by C_J is $\geqq$ the area enclosed by J, and equality holds only if J itself is convex and symmetric.*

Following Busemann (see [4]—[7] and sources quoted in those papers) and Petty [18] a Banach norm $\| \;\|$ in R^2 defines a Minkowski metric m. The length of a rectifiable curve J is then measured by

$$\lambda_m(J) = \int_J \|ds\|. \tag{2}$$

$U=\{\mathbf{x}\in R^2\colon \|\mathbf{x}\|=1\}$ is the unit circle of the metric; every centrally symmetric, nondegenerate convex curve is the unit circle of a Minkowski metric. Let $U^{\perp}$ denote the *isoperimetrix,* i.e. the curve obtained from U by taking its polar with respect to a Euclidean unit circle, and rotating it by $\pi/2$. Busemann showed [4] that the solution of the isoperimetric problem in the Minkowski plane is unique, and is homothetic to

---

1980 *Mathematics Subject Classification.* Primary 52A40; Secondary 51M25.

Key words and phrases. Isoperimetric problem, Minkowski metric, isoperimetrix, mixed volumes, convexification.

$U^{\perp}$. In this paper, we also consider "variable Minkowski metric", and, denoting $\mathcal{M}$ the set of all Minkowski metrics, we show

THEOREM 2. *Let C_J be the symmetric convexification of the rectifiable curve J. Then $\lambda_m(C_J)=\lambda_m(J)$ for all $m\in\mathcal{M}$. Consequently, C_J encloses the largest area in the class of rectifiable Jordan curves*

(3) $$\{K\subset R^2 \colon \lambda_m(K)=\lambda_m(J),\ \forall m\in\mathcal{M}\},$$

hence is the solution of an isoperimetric problem.

Interestingly, to prove Theorem 1 concerning Euclidean metrics, we first prove Theorem 2 concerning Minkowski metrics. The reason is that Minkowski geometry, as developed by Busemann, is the natural setting for the theory of convex bodies, which is our main tool here.

All questions discussed above have natural generalizations to n-dimensions, however, the latter are much more difficult. Even the construction of convexification requires Minkowski's theory in n-dimensions. However, we prove the following result about uniqueness:

THEOREM 3. *If in each Euclidean metric with unit sphere (which is an ellipsoid in R^n) near to a fixed ellipsoid the surface area of a centrally symmetric convex surface S is given, then S is determined up to translation.*

In general terms, one deals with two functions $\lambda\colon \mathcal{M}\times\mathcal{J}\to R$ and $A\colon \mathcal{M}\times\mathcal{J}\to R$: in R^2 (in R^n) $\mathcal{M}$ is the set of Minkowski metrics of R^2 (of R^n) and $\mathcal{J}$ is the set of rectifiable curves (objects for which $(n-1)$-area is defined), while $\lambda(m, J)$ is the length (is the $(n-1)$-area) and $A(m, J)$ is the enclosed area (enclosed volume). Usually, m is fixed, and λ, A are functions of J. Here J is fixed, and λ, A are functions of m. For example, $\mathcal{M}_E$ is an analytic manifold (in fact, positive definite symmetric matrices form an open subset of $R^{n(n+1)/2}$) and $\lambda(m, S)$ is analytic in m, e.g. for convex, or smooth S (see the proof of Theorem 3).

2. A located vector $\vec{\mathbf{ab}}$, $\mathbf{a}, \mathbf{b}\in R^2$, determines a vector $\vec{\mathbf{u}}$, which is considered to be an equivalence class of located vectors. $\vec{\mathbf{ab}}+\vec{\mathbf{bc}}=\vec{\mathbf{ac}}$, and $\mathbf{a}+\vec{\mathbf{u}}=\mathbf{b}$, iff $\vec{\mathbf{u}}$ is the equivalence class of $\vec{\mathbf{ab}}$.

CONSTRUCTION 1. *Convexification of a closed polygon.* Let $\mathbf{a}_0, \dots, \mathbf{a}_n$ be a sequence of points in R^2 determining the closed polygon P with edges $\mathbf{a}_{i-1}\mathbf{a}_i$, $i=1, \dots, n$ (also $i=n+1$, with $\mathbf{a}_{n+1}=\mathbf{a}_0$). Let $\vec{\mathbf{u}}_i$ be the class of $\overline{\mathbf{a}_{i-1}\mathbf{a}_i}$. Given a permutation π of $\{1, \dots, n+1\}$, $\pi 1=1$, we denote πP the polygon with vertices $\mathbf{b}_0=\mathbf{a}_0$, $\mathbf{b}_i=\mathbf{b}_{i-1}+$ $+\vec{\mathbf{u}}_{\pi i}$ $(\mathbf{b}_{n+1}=\mathbf{a}_0)$. If $\vec{\mathbf{u}}_{\pi 1}, \dots, \vec{\mathbf{u}}_{\pi(n+1)}$ is the positive, i.e. counterclockwise order of these vectors, then $Q=\pi P$, or a translate of Q, is called a *convexification* of P. If Q is the boundary of the convex domain D, the boundary Q_s of $\frac{1}{2}[D+(-D)]$, Minkowski sum, is called *symmetric convexification* of P.

In elementary geometric terms: each side of the convexification Q of P is parallel to some side of P and has length equal to the total length of the sides of P directly parallel to it. For the symmetric convexification Q_s of P, each side of the centrally symmetric Q_s is parallel to some side of P, and has length half the total length of the sides of P parallel to it. Clearly

(4) $$\lambda_m(\pi P) = \lambda_m(Q) = \lambda_m(Q_s) = \lambda_m(P)$$

for all Minkowski metrics m.

If we consider the sides of a polygon as singular 1-simplices, we have an integral chain (cycle). Specifically, for the closed polygon P

(5) $$t = \sum_{i=0}^{n} s[\mathbf{a}_i, \mathbf{a}_{i+1}],$$

where $s[\mathbf{a}, \mathbf{b}]$ denotes the map $s[\mathbf{a}, \mathbf{b}]\,(t_0, t_1) = t_0\mathbf{a} + t_1\mathbf{b}$. $S(t)$ is the support of t. For $\mathbf{x} \in R^2 - S(t)$ the winding number $w(t; \mathbf{x})$ is well-defined, and

(6) $$\sigma(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w(t; (\xi, \eta))\, d\xi\, d\eta$$

will be called signed area enclosed by t. If P is a Jordan polygon, i.e. a polygon which is a Jordan curve, then $\pm\sigma(t) = A(P)$ elementary geometric area enclosed by P. Now we prove a lemma also following from the method of the recent [23] which we received after essentially completing our paper. Our proof is different.

Lemma 1. *Let P be a closed polygon, t the cycle* (5), *and c the cycle constructed for the convexification Q of P with the positive orientation of the boundary. Then*

(7) $$\sigma(t) \leqq \sigma(c),$$

where equality holds only if P is convex and positively oriented, or is contained in a straight line. Hence

(8) $$A(P) \leqq A(Q),$$

if P is a Jordan polygon.

Proof. For $n=3$ we have a quadrilateral and its convexification, hence the statement is easy to prove directly. (Take the closed convex hull of the vertices, and consider the cases when this is a segment, a triangle, or a quadrilateral.) We now suppose that $n \geqq 4$, and that there is a non-convex quadrilateral

(9) $$\mathbf{a}_i, \mathbf{a}_{i+1}, \mathbf{a}_j, \mathbf{a}_{j+1}, \qquad 0 \leqq i < j \leqq n.$$

We will replace this by a convex quadrilateral and show that the σ-value is increased thereby. (We note that this operation may produce self-intersections, and that is the reason why we work with σ.) Specifically, we set

$$t_1 = \sum_{p=0}^{i-1} s[\mathbf{a}_p, \mathbf{a}_{p+1}] - s[\mathbf{a}_{j+1}, \mathbf{a}_i] + \sum_{p=j+1}^{n} s[\mathbf{a}_p, \mathbf{a}_{p+1}];$$

$$t_2 = s[\mathbf{a}_i, \mathbf{a}_{i+1}] + s[\mathbf{a}_{i+1}, \mathbf{a}_j] + s[\mathbf{a}_j, \mathbf{a}_{j+1}] + s[\mathbf{a}_{j+1}, \mathbf{a}_i];$$

$$t_3 = \sum_{p=i+1}^{j-1} s[\mathbf{a}_p, \mathbf{a}_{p+1}] - s[\mathbf{a}_{i+1}, \mathbf{a}_j].$$

Then $t_1+t_2+t_3=t$, consequently $\sigma(t_1)+\sigma(t_2)+\sigma(t_3)=\sigma(t)$. Let us convexify (9) and denote t_2^* the corresponding cycle: $\sigma(t_2^*)>|\sigma(t_2)|$, by construction. Now $s[\mathbf{a}_{j+1}, \mathbf{a}_i]$ was shifted in a new position when we constructed t_2^*, hence t_1 can be shifted in a new position t_{12}, so that this simplex is cancelled in $t_{12}+t_2^*$. Similarly, t_3 can shifted to t_{32}, so that $t_{12}+t_2^*+t_{32}$ does not contain $s[\mathbf{a}_{i+1}, \mathbf{a}_j]$. Finally, we may shift this cycle in new position $t_{13}+t_{23}^*+t_{33}$ where $s[\mathbf{a}_0, \mathbf{a}_1]$ belongs to it again.

Suppose now P is that polygon, for which $\sigma(t)$ attains its maximum, among all polygons πP, and P is not contained in a line. What we have shown above, implies about the side vectors $\vec{\mathbf{u}}_i$ that no $\vec{\mathbf{u}}_i, \vec{\mathbf{u}}_{i+1} (\vec{\mathbf{u}}_{n+2}=\vec{\mathbf{u}}_1)$ have opposite directions, and also that the angle (of absolute value $<\pi$) between the directions of $\vec{\mathbf{u}}_{i+1}$ and $\vec{\mathbf{u}}_i$ is $\geqq 0$ for all i, or is $\leqq 0$ for all i. Also if $\mathbf{a}_i, \mathbf{a}_{i+1}, \mathbf{a}_j$ (resp. $\mathbf{a}_{j+1}$) lie on a line, but $\mathbf{a}_{j+1}$ (resp. $\mathbf{a}_j$) does not, then their order on the line is $\mathbf{a}_i, \mathbf{a}_{i+1}, \mathbf{a}_j$ (resp. $\mathbf{a}_{j+1}, \mathbf{a}_i, \mathbf{a}_{i+1}$). These imply the following: there are no collinear, oppositely oriented sides, and a self-intersection can occur only if there are collinear, intersecting sides.

Hence P is either Jordan, and then (positively oriented) convex, or has two collinear, intersecting, similarly oriented sides $\vec{\mathbf{u}}_i, \vec{\mathbf{u}}_j$. By the above facts, if we pass from $\vec{\mathbf{u}}_i$ in the given orientation to further sides, lying on the same line, the end-point of the last one, $\mathbf{a}_{i'}$, will coincide with the similarly defined $\mathbf{a}_{j'}$, further also $\vec{\mathbf{u}}_{i'+1}$ and $\vec{\mathbf{u}}_{j'+1}$ are collinear. Hence P is a k times $(k>1)$ traversed convex polygon, contradicting our maximality assumption.

3. It is easy to see that Construction 1 gives a convexification in the sense of the following definition.

DEFINITION 1. *Given a rectifiable Jordan curve J in R^2, the boundary K of a convex domain D is called a convexification of J if there is a sequence J_n of Jordan polygons inscribed in J, such that the length of the longest side of J_n tends to* 0 *with n, and that the convexification K_n of J_n tends to K. The boundary C_J of $\frac{1}{2}[D+(-D)]$, Minkowski sum, is called symmetric convexification of J.*

Given an arbitrary Jordan curve J, and $\varepsilon>0$, there is an inscribed Jordan polygon J_ε, such that the longest side of J_ε is of length $<\varepsilon$ and the distance between corresponding points on J and J_ε is $<\varepsilon$ (in the standard metric of R^2, see [20], [22]). Consequently, if J is rectifiable, there is a sequence of Jordan polygons J_n inscribed in J, such that: 1° the standard length of the longest side of J_n is $<1/n$; 2° $\lim \lambda(J_n)=\lambda(J)$, where $\lambda(J)$ is the standard length (and similarly for λ_m); 3° $\lim A(J_n)=A(J)$. Clearly, $\lambda_m(J_n)=\lambda_m(K_n)$, $m\in\mathcal{M}$, where K_n is a convexification of J_n. By (8) and 3°

$$A(J) \leqq A(K). \tag{10}$$

Later we will see the uniqueness of the convexification. For the moment we state:

PROPOSITION 1. *Every rectifiable Jordan curve has a convexification. If K is the convexification of J, $\lambda_m(K)=\lambda_m(J)$ for every Minkowski metric m, and $A(K)\geqq A(J)$, with equality only for convex J.*

PROOF. We need to show only $A(J)<A(K)$ for non-convex J. There is a supporting line through $\mathbf{x}\neq\mathbf{y}\in J$, the open segment $\mathbf{xy}$ not intersecting J. Replace one of the arcs $\mathbf{xy}$ of J by its centrosymmetric image w.r.t. $\frac{\mathbf{x}+\mathbf{y}}{2}$, obtaining a Jordan

curve J' enclosing a larger area, and having a same convexification. Thus $A(J) < A(J') \leqq A(K)$.

4. In the sequel, we use the conventions and notations of the theory of convex bodies [3]. In particular, if C is a convex domain in R^2, $A(C)$ is its area, $\lambda_m(C)$ its perimeter in the metric $m \in \mathcal{M}$. If C_1, C_2 are convex, and $A(C_2) = A(C_1)$, then by the Brunn—Minkowski inequality [3; 48, p. 88],

$$A\left(\frac{1}{2}(C_1+C_2)\right) \geqq A(C_1), \tag{11}$$

where equality holds only, if C_1, C_2 are homothetic.

On the other hand, for arbitrary convex domains C_1, C_2,

$$\lambda_m(C_1+C_2) = 2A(C_1+C_2, U^\perp), \tag{12}$$

where we have mixed area A on the right, and $U^\perp$ is the isoperimetrix of m as above [4; p. 864]. The right-hand side of (12) is equal to

$$2A(C_1, U^\perp)+2A(C_2, U^\perp) = \lambda_m(C_1)+\lambda_m(C_2). \tag{13}$$

In particular,

$$\lambda_m(C) = \lambda_m\left(\frac{1}{2}[C+(-C)]\right). \tag{14}$$

From these observations we will easily obtain the following result:

Proposition 2. *Given a rectifiable Jordan curve J, its convexification K and symmetric convexification C_J belong to* (3). *There is an up to translation unique curve in* (3) *enclosing the maximal area; this curve is convex, and centrally symmetric, hence it is C_J.*

Proof. If J' is in (3), then any convexification K' of J' is also in (3). Now $A(J')$, J' in (3), is a bounded set of numbers, consequently there is a sequence K_n of convex curves in (3) such that $\lim A(K_n) = \sup A(J')$. As $\lambda(K_n)$ is bounded, some translates of the K_n's lie in a fixed compact subset of the plane. As $\lambda_m(K)$ is continuous on the set of all convex domains, by the usual compactness argument, we see that the maximum area is obtained for the boundary K of some convex domain C. However, by Proposition 1 no non-convex curve has a maximum area. If C_1, C_2 would be two convex sets of maximum area, $\frac{1}{2}(C_1+C_2)$ would be in (3) by (12), (13), while in virtue of the Brunn—Minkowski inequality (11), C_1, C_2 would be homothetic, hence equivalent under translation. This applies to $C_1 = C$, $C_2 = -C$, and proves that C is centrally symmetric.

Corollary 1. *Let P be a Jordan polygon. Then among the rectifiable Jordan curves C satisfying $\lambda_m(C) = \lambda_m(P)$ for all Minkowski metric m, the one with maximal enclosed area is the following centrally symmetric, convex polygon Q_s: each side of Q_s is parallel to some side of P, and has length half the total length of the sides of P parallel to it.*

PROOF. Notice that Q_s equals C_p of Proposition 2.

PROPOSITION 3. *If* $\lambda_m(C_1)=\lambda_m(C_2)$ *for all* $m\in\mathcal{M}$, *where* C_1, C_2 *are non-degenerate convex domains, then* $\frac{1}{2}[C_2+(-C_2)]$ *is a translate of* $\frac{1}{2}[C_1+(-C_1)]$.

PROOF. By considering $\frac{1}{2}[C_i+(-C_i)]$ instead of C_i (which satisfy the same conditions), we may suppose C_i symmetric. We have $\lambda_m(C_i)=2A(C_i, U^\perp)$, as above. Hence, as $U^\perp$ may be any symmetric convex domain D,

$$A(C_1, D) = A(C_2, D). \tag{15}$$

Next, given any convex domain E, $A(C_1, E)=A(-C_1, -E)=A(C_1, -E)= =A\left(C_1, \frac{1}{2}[E+(-E)]\right)=A\left(C_2, \frac{1}{2}[E+(-E)]\right)=A(C_2, E)$. In other words, we just proved (15) for every convex domain D. Thus by [7; (9.2), (9.4), pp. 69—70] we conclude that C_2 is a translate of C_1.

REMARK 1. Instead of symmetric metrics ($\|-\vec{\mathbf{x}}\|=\|\vec{\mathbf{x}}\|$, or, equivalently, U centrally symmetric) occurring above, we may consider asymmetric metrics (U arbitrary convex, containing the origin in its interior). Then the results above remain valid without essential changes (measuring the perimeter in (2) in positive orientation; by λ_m (negatively oriented J)$=\lambda_m$ (positively oriented $-J$) this is no restriction). The curve enclosing maximal area in Proposition 2 need not be symmetric, hence is the convexification. In Corollary 1 in place of Q_s we have the convexification Q of P, i.e. a polygon whose sides are sums of parallel similarly traversed sides of P. In Proposition 3 C_2 is a translate of C_1. A similar remark applies to the following Lemma 2 and Proposition 4, dropping the symmetry conditions, resp. considering the convexification of J. [For Lemma 2 proceed as follows. Let μ be the maximal number, such that $\mu C\subset D$, up to translation. Then either (1) μC and D have a supporting strip in common, or (2) μC possesses an inscribed triangle $\mathbf{c}_1\mathbf{c}_2\mathbf{c}_3$, D possesses a circumscribed triangle $\mathbf{d}_1\mathbf{d}_2\mathbf{d}_3$, $\mathbf{c}_i$ lying on the side $\overline{\mathbf{d}_j\mathbf{d}_k}$, $j, k\neq i$. In case (2) define a metric m with unit circle a triangle $\mathbf{u}_1\mathbf{u}_2\mathbf{u}_3$, with $\overrightarrow{\mathbf{O}\mathbf{u}_i}$ parallel and similarly oriented to $\mathbf{d}_j\mathbf{d}_k$ ($\mathbf{d}_1\mathbf{d}_2\mathbf{d}_3$ taken in positive orientation). Then $\lambda_m(\mathbf{c}_1\mathbf{c}_2\mathbf{c}_3)=\lambda_m(\mathbf{d}_1\mathbf{d}_2\mathbf{d}_3)$, hence $\lambda_m(\mu C)=\lambda_m(D)$, which implies the statement. In case (1) use e.g. $m(\varepsilon)$ of Lemma 3, for which $\lambda_{m(\varepsilon)}(\mu C)/\lambda_{m(\varepsilon)}(D)\to 1$, if $\varepsilon\to 0$.]

REMARK 2. Let us use the notations of Definition 1. Let us denote K_{sn} the symmetric convexification of J_n; we select them so that they all have a common center. Then the sequence K_{sn} converges to the symmetric convex curve enclosing the largest area, hence to the symmetric convexification of J. In fact, if the sequence would not converge, then it would contain two subsequences converging to two different symmetric convex curves, both enclosing maximal area in the class (3) of curves. Thus one would be a translate of the other, which is a contradiction. A similar remark applies to the case of asymmetric metric.

REMARK 3. If the Jordan curve J has continuous curvature with a finite number of zeros, or, more generally, consists of a finite number of arcs, each either convex or concave, then the convexification K of J can be described as follows. Suppose that the directions of the positively traversed half-tangents vary on these arcs from φ_i to ψ_i. Let (α_j) be all these angles in cyclic order $0 \leqq \alpha_j < 2\pi$. For each $[\alpha_j, \alpha_{j+1}]$, consider the parts of the arcs in which the positively traversed half-tangents have angles in $[\alpha_j, \alpha_{j+1}]$, and consider the Minkowski sum K' of these arcs. These, for all j, fit together to K, since, evidently, $\lambda_m(K') = \lambda_m(K)$, for all $m \in \mathcal{M}$.

REMARK 4. Let us consider asymmetric metrics. Hitherto, we only dealt with $\lambda_m(K)$, $m \in \mathcal{M}$. We have $\lambda_m(K) = \lambda_m(J)$ for all asymmetric metrics m, and even (using elements of Lebesgue integration) we see for any continuous $f\colon [0, 2\pi] \to R$, $f(0) = f(2\pi)$

$$\int_J f(\varphi)\, ds = \int_K f(\varphi)\, ds. \tag{16}$$

(φ is the angle of the tangent at ds).

COROLLARY 2. *The convexification K of a rectifiable Jordan curve J is unique up to translation.*

5. The classical isoperimetric theorem can be formulated concerning rectifiable Jordan curves K, such that $\lambda(K) \leqq$ const. Similarly, for a given rectifiable Jordan curve J, we can consider the class of all rectifiable Jordan curves K, such that

$$\lambda_m(K) \leqq \lambda_m(J), \quad \forall m \in \mathcal{L}, \tag{17}$$

where $\mathcal{L} \subset \mathcal{M}$ is given. In the classical case, the new formulation is trivially equivalent to the old. This is also true in our case, although less trivially:

LEMMA 2. *If for centrally symmetric convex domains C, D we have $\lambda_m(C) \leqq \lambda_m(D)$ for all $m \in \mathcal{M}$, then a translate of C is contained in D.*

PROOF. We have $A(C, U^\perp) \leqq A(D, U^\perp)$ for mixed areas by $\lambda_m(C) = 2A(C, U^\perp)$ [4; p. 864]. Here $U^\perp$ can be an arbitrary centrally symmetric convex curve, thus by a limit procedure we get $A(C, E) \leqq A(D, E)$ for $E =$ any unit segment. By [3; p. 45] $2A(C, E)$ is the length of (any) projection of C, whence a translate of C is contained in D.

From the above, we get easily:

PROPOSITION 4. *Among the rectifiable Jordan curves K such that $\lambda_m(K) \leqq \lambda_m(J)$, for all $m \in \mathcal{M}$, where J is a fixed curve, the largest area is enclosed by the symmetric convexification C_J of J.*

Clearly, with above results, we have the proof of Theorem 2 of the introduction. Theorem 3 will be proven later, when we discuss the n-dimensional case. We anticipate this proof, and use the special case of the theorem for $n=2$ in the

PROOF OF THEOREM 1. By 1° of Theorem 1, and by the special case $n=2$ of Theorem 3, C_J is the symmetric convexification of J, and $\lambda_m(C_J) = \lambda_m(J)$ for all Minkowski metrics m. Hence 2° follows.

Sharper form of Theorem 1. Let us denote $\mathcal{L}$ the set of Euclidean metrics, whose unit circles are near a fixed ellipse in R^2. Replacing 1° of Theorem 1 by

$$\lambda_m(C_J) = \lambda_m(J), \quad \forall m \in \mathcal{L}, \tag{18}$$

we have a weaker condition, with the same conclusion, hence a sharper form of Theorem 1.

We can improve Theorem 1 in another direction:

PROPOSITION 5. *In the class of all rectifiable Jordan curves K, such that* $\lambda_m(K) \leqq \lambda_m(J)$ *for all Euclidean metrics m, the symmetric convexification* C_J *of J encloses the largest area.*

(This cannot be improved by requiring the inequality to hold near a fixed metric only, as seen by taking J = any symmetric non-circular domain, K = circle, of a larger area and a smaller perimeter than those of J (in Euclidean sense).)

For the proof of this we need the simple

LEMMA 3. *Let* $m(\varepsilon)$ *be the Euclidean metric with unit circle* $\varepsilon^2x^2+y^2=1$, *and let C be a convex domain. Then, for* $\varepsilon\to 0$, $\frac{1}{2}\lambda_{m(\varepsilon)}(C)$ *converges to the width of the supporting strip parallel to the x-axis.*

PROOF. We have $\lambda_{m(\varepsilon)}(C)=\lambda_{m(1)}(C_\varepsilon)$, where C_ε is the image of C under the affinity $(x, y)\to(\varepsilon x, y)$. Hence, $\frac{1}{2}\lambda_{m(\varepsilon)}(C)$ is $\geqq$ than the width above. On the other hand, C is contained in some rectangle with sides parallel to the axes, thus the statement follows.

COROLLARY 3. *If for the centrally symmetric, convex J, K we have* $\lambda_m(K) \leqq \lambda_m(J)$ *for all Euclidean metrics m, then K is contained in a translate of J.*

PROOF OF PROPOSITION 5. We can replace both J and K by their respective symmetric convexifications, apply Theorem 1 and Corollary 3.

REMARK 1. Lemma 3, Corollary 3 and Proposition 5 have evident analogues, when, instead of considering all Euclidean metrics, we consider all Minkowski metrics with unit circle an affine image of a fixed centrally symmetric convex curve. (For asymmetric Minkowski metrics the analogue of Proposition 5 does not hold.)

REMARK 2. From the narrow point of view of this paper, the most efficient development would be to fix a J, and first consider all K's satisfying (16) for all continuous f, $f(0)=f(2\pi)$. This leads to the convexification of J. The above development is preferable, inasmuch as it shows connections with Euclidean and Minkowski metrics. However, in higher dimensions we will pay more attention to the generalization of (16).

6. In R^n $(n\geqq 2)$ we can pose the same problems as in the plane. We first consider geometric simplicial complexes, as natural generalizations of polygons (see [21; p. 357] under *simplicial complex*).

As in the plane, we need some remarks about orientations. A non-degenerate, oriented affine $(n-1)$-simplex in R^n spans an $(n-1)$-flat L, and defines an orientation of L. There is then a unique unit vector $\bar{\mathbf{u}}$, that is the exterior normal vector, if we think of the orientation of L as being induced by a convex polytope with one face in L.

CONSTRUCTION 2. *Convexification of an oriented geometric simplicial complex of dimension $n-1$ in R^n.* Let K be such a complex; each $(n-1)$-simplex is oriented (arbitrarily). Let

$$(19) \qquad \vec{\mathbf{u}}_1, \ldots, \vec{\mathbf{u}}_k \quad (\textit{if } i \neq j,\ \vec{\mathbf{u}}_i \neq \vec{\mathbf{u}}_j)$$

be the set of all the exterior normal unit vectors of the oriented $(n-1)$-simplices of K; we suppose that (19) spans R^n. Let us denote a_i the sum of the $(n-1)$-Lebesgue measures of all the $(n-1)$-simplices of K that have $\vec{\mathbf{u}}_i$ as exterior normal. By a classical theorem of Minkowski [17; II, p. 113], [3; 60), p. 118], the vectors

$$(20) \qquad a_1\vec{\mathbf{u}}_1, \ldots, a_k\vec{\mathbf{u}}_k, \quad a_{k+1}\vec{\mathbf{u}}_{k+1} = -\sum_{i=1}^{k} a_i\vec{\mathbf{u}}_i$$

(the last being omitted, if $\mathbf{0}$) determine an up to translation unique convex polyhedron P for which $\vec{\mathbf{u}}_i$ is the exterior normal for the i-th face A_i, and $\mu_{n-1}(A_i)=a_i$ (Lebesgue measure). The complement of the interior of A_{k+1} on the boundary of P (or P itself) is called a convexification of K (oriented, as given).

CONSTRUCTION 3. *Symmetric convexification of a geometric simplicial complex of dimension $n-1$ in R^n.* Let K be such a complex. Let (19) be the set of all unit vectors orthogonal to the $(n-1)$-simplices of K (if $\vec{\mathbf{u}}_i$ is in (19), $-\vec{\mathbf{u}}_i$ is also there); we suppose that (19) spans R^n. Let us denote b_i half the sum of the $(n-1)$-Lebesgue measures of all the $(n-1)$-simplices of K that are orthogonal to $\vec{\mathbf{u}}_i$; if $\vec{\mathbf{u}}_j=-\vec{\mathbf{u}}_i$, $b_j=b_i$. Then there is an up to translation unique centrally symmetric convex polytope Q whose j-th face B_j has $\vec{\mathbf{u}}_j$ as exterior normal, while $\mu_{n-1}(B_j)=b_j$. We call the boundary of Q (or Q itself) symmetric convexification of K.

LEMMA 4. *Let K be a geometric simplicial complex of dimension $n-1$ in R^n, whose $(n-1)$-simplices s_i, $i=1, \ldots, l$ are coherently oriented, i.e. such that*

$$(21) \qquad t = \sum_{i=1}^{l} s_i$$

is a cycle (here Σ stands for formal sum, which is an element of the simplicial chain group). Then $a_{k+1}\vec{\mathbf{u}}_{k+1}=\mathbf{0}$ in (20).

PROOF. The cone operator D centered to the origin defines an n-chain Dt, such that $t=\partial Dt$. We set

$$Dt = \sum_{j=1}^{m} c_j \bar{s}_j$$

where $\bar{s}_j$ is an n-simplex of R^n, and the formal sum is reduced. Let us denote $\vec{\mathbf{u}}_{j0}, \ldots, \vec{\mathbf{u}}_{jn}$ the exterior normal unit vectors of $\bar{s}_j$, and $a_{j0}, \ldots, a_{jn}$ the $(n-1)$-Lebesgue measures of the respective faces. By Minkowski's theorem, the vector sum

$$a_{j0}\vec{\mathbf{u}}_{j0} + \ldots + a_{jn}\vec{\mathbf{u}}_{jn}$$

is $\mathbf{0}$. Multiplying by c_j and summing gives the result.

COROLLARY 4. *Let M be a geometric simplicial complex of dimension $n-1$, and a connected C^0-manifold without boundary. Then the convexification of M (the simplices being taken with appropriate orientations) is the full boundary of the convex polyhedron P.*

PROOF. $R^n - M$ has two components by the generalized Jordan—Brouwer theorem; if U is the bounded component, $M = \overline{U} - U$. Let R^p be the subspace of R^n generated by the exterior normal unit vectors (19). Supposing $p \leqq n-1$, there is a direction $\pm\bar{\mathbf{u}}$ orthogonal to R^p, and if we project all the $(n-1)$-simplices of M in this direction into an open ball $D \subset U$, the complement of the projection will not be empty. A ray of direction $\bar{\mathbf{u}}$ starting in the complement will avoid all the $(n-1)$-simplices of M, hence M itself, which is a contradiction. Thus $R^p = R^n$, and P is defined. The intersections of any of the both half-spaces bounded by all the $(n-1)$-flats of R^n containing $(n-1)$-simplices of M are convex open sets, some of whose closures form a cell decomposition of $\overline{U}$; we thus construct a simplicial decomposition of $\overline{U}$. Taking all the n-simplices of the decomposition with positive orientations and with coefficients $+1$ gives an integral chain c of $\overline{U}$, and $t = \partial c$ is a fundamental cycle of M. The statement then follows by Lemma 4.

The above seem to be natural generalizations of our results concerning the plane. However, as yet, we were unable to generalize Lemma 1:

CONJECTURE 1 (cf. also the Remark added in proof). Convexification increases the volume enclosed, under appropriate conditions. Specifically, with the notations of the proof of Corollary 4,

$$\mu_n(\overline{U}) \leqq \mu_n(P), \tag{22}$$

where equality implies that $\overline{U}$ is convex.

Presumably, in order to prove this conjecture, one would need generalizations of the tools used in the plane. This is available: Given an affine $(n-1)$-cycle h in R^n (with non-negative coefficients) we can define the winding number $w(h; \mathbf{x})$, $\mathbf{x} \in R^n - S(h)$, and set

$$\sigma(h) = \int_{R^n} w(h; \mathbf{x})\, dv \tag{23}$$

where dv is the volume element of R^n (or $\sigma(h)$ = sum of the signed volumes of the simplices in the chain Dh, with the respective coefficients). By the proof of Lemma 4 we have $\Sigma \gamma_i a_i \bar{\mathbf{u}}_i = \mathbf{0}$ (with notations analogous to those in Construction 2, and γ_i's are the coefficients of the cycle). Supposing the $\bar{\mathbf{u}}_i$'s span R^n, one expects

$$\sigma(h) \leqq \mu_n(P), \tag{24}$$

where P is the convexification, defined in analogy to Construction 2. (24) would imply Conjecture 1, as $\sigma(h) = \mu_n(\overline{U})$, if h is the positive fundamental cycle of M (using the notations of the proof of Corollary 4).

If Conjecture 1 would be true, it also would be more clear, why do the convex bodies have such a great importance. Also the (classical or Minkowskian [5]) isoperimetric theorem would become clearer.

7. In the plane we had: existence of the convexification for all rectifiable Jordan curves, and area increasing property of this operation in the general case. In R^n we now have convexification for polytopes, hence for a very restricted class, and we do not know whether convexification increases the enclosed volume or not.

In this section we will discuss some conjectures, and partial results. We will say "surface S" meaning a triangulable $(n-1)$-manifold (with or without boundary) in R^n with various "smoothness" restrictions, as appropriate. For full generality, one would need a geometric integration theory [12], [21]. General C^k-class is certainly not always appropriate here, as the solutions of our isoperimetric problems may be general convex surfaces.

At some smoothness conditions the surface area can be defined as $\int dS$. If R^n is supplied with a Minkowski metric m with unit ball U, according to Busemann [6], S has a Minkowski area

$$\int \varkappa_{n-1} \frac{dS}{A(v(dS))}, \tag{25}$$

where $v(dS)$ is the normal unit vector to dS, $A(v)$ is the area of the intersection of U with a hyperplane through $\mathbf{0}$ normal to v, and $\varkappa_{n-1}$ is the Euclidean volume of the Euclidean unit $(n-1)$-ball. As in (16), we can also consider any integral of the form

$$\int f(v(dS))\,dS, \tag{26}$$

where f is continuous (real-valued) on S^{n-1}, S is orientable, and $v(dS)$ is a fixed (say, exterior) normal. Our main question will be:

Among the surfaces with equal Minkowski areas in all Minkowski metrics, or in the smaller class of orientable surfaces with equal integrals (26) *for all continuous f, which are the ones bounding the maximal volume?*

Let P be a convex polyhedron with inner points in R^n, A_i its i-th $(n-1)$-face, $\vec{\mathbf{u}}_i$ the exterior normal unit vector of A_i, and $a_i = \mu_{n-1}(A_i)$, $(n-1)$-Lebesgue measure, $i = 1, \ldots, k$. We now consider that these data determine a measure on the unit sphere S^{n-1} of R^n with weight a_i at $\vec{\mathbf{u}}_i \in S^{n-1}$. Then

$$\int f(v(dS))\,dS = \sum_{i=1}^{k} f(\vec{\mathbf{u}}_i)\,a_i. \tag{27}$$

Thus (26) for all continuous f determines the measure, and vice versa. Hence, from (27) and from Construction 1, 2, we get:

PROPOSITION 6. *If M, P are as in Corollary 4, and* (26) *is considered as an integral of measures, then this integral is the same for both surfaces M and ∂P (boundary of P). In particular, Euclidean surface areas in all Euclidean metrics, and areas in all Minkowski metrics agree. The last sentence applies also to an $(n-1)$-complex K and its symmetric convexification Q (see Construction 3).*

The uniqueness part of Minkowski's theorem quoted in Construction 2 means that the measure on S^{n-1} determines the convex polyhedron, up to translation. However, in this case it is not evident that if S_1, S_2 are convex, having equal Minkowski areas (for a set $\mathscr{L}$ of Minkowski metrics) or equal integrals (26) (for a set $\mathscr{F}$ of continuous functions f in (26)), are identical up to translation. We will come back to this question below.

PROPOSITION 7. *Among the (non-degenerate) convex surfaces with equal Minkowski areas — for a set* $\mathcal{L}\neq\emptyset$ *of Minkowski metrics — or equal integrals* (26) *— for a set* $\mathcal{F}$ *of continuous functions, containing at least one positive function — there is a unique one, bounding a maximal volume.*

PROOF. The existence of a convex S with maximal volume is immediate from the facts that (26) is continuous in S ([1], §6, Lemma and §5, (2)) and that, if the Euclidean surface area of S is bounded above, and its diameter tends to infinity, then the volume bounded by it tends to zero. This allows the usual compactness arguments.

Let us recall that for convex surfaces S_1, S_2 there exists a surface S (called their Blaschke sum in [14]) with

$$dS(v) = dS_1(v)+dS_2(v), \tag{28}$$

where the differentials are interpreted as Borel measures on S^{n-1} (surface area functions, [1], [7; (8.3), p. 62], [13]). (In R^2 this is the Minkowski sum, in R^n it is not.) In R^n, the inequality

$$V(S)^{\frac{n-1}{n}} \geq V(S_1)^{\frac{n-1}{n}} + V(S_2)^{\frac{n-1}{n}} \tag{29}$$

is proved in [15] for the volumina with equality only for homothetic S_1, S_2, for smooth bodies. But in fact the same proof, read in modern terminology ([1], [7], [13]), shows the general validity of (29) which implies uniqueness of an S enclosing maximal volume. In fact, if both S_1 and S_2 bound a maximal volume, then $2^{-\frac{1}{n-1}}S$ has equal integrals (26), but bounds a greater volume, unless $S_1=S_2$.

THEOREM 4. *If for a convex surface S the integral* $\int f(v)\,dS(v)$ (*v is exterior normal unit vector, $dS(v)$ is surface area function*) *is known for every continuous f, then S is determined, up to translation.*

PROOF. The values of these integrals determine the regular Borel measure $dS(v)$. Thus the surface S is determined by [7; Theorem 8.6, p. 64] or [2; (9.4), p. 70].

CONJECTURE 2. There is a class $\mathcal{S}$ of compact subspaces of R^n, and a geometric integration theory with following properties: All convex surfaces are in $\mathcal{S}$; for every $M\in\mathcal{S}$ a surface area is defined, which is the standard surface area, if M is a convex surface. All geometric simplicial complexes of dimension $n-1$ are in $\mathcal{S}$; every $M\in\mathcal{S}$ is homeomorphic to a geometric simplicial complex. Every $(n-1)$-dimensional C^r-submanifold of R^n is in $\mathcal{S}$, $r\geq 1$; there is a differential dM if $M\in\mathcal{S}$; dM is the classical area element if M is smooth. There is a generalized Gauss map v of Borel sets of S^{n-1} into Borel sets of $M\in\mathcal{S}$, such that

$$\int_A dM(v) = \int_{v(A)} dM \qquad (A\subset S^{n-1}) \tag{30}$$

defines a Borel measure $dM(v)$ on S^{n-1}, where differentials and integrals are taken in the sense of the selected geometric integration theory. If $M\in\mathcal{S}$ is a (connected)

C^0-manifold without boundary,

$$\int_{S^{n-1}} v\,dM(v) = 0. \tag{31}$$

For fixed M in $\mathscr{S}$ and for given set $\mathscr{F}$ of continuous functions f, we consider the class of $S \in \mathscr{S}$ such that

$$\int_S f(v)\,dS = \int_M f(v)\,dM \tag{32}$$

holds for every $f \in \mathscr{F}$. We conjecture, that if $\mathscr{F}$ = all continuous functions on S^{n-1} there is a unique convex surface S with $dS(v) = dM(v)$, which belongs to the class, and which also is the unique one enclosing the maximal volume (supposing (31) is satisfied). We also expect new phenomena in R^n, similar to the fact that *not* every centrally symmetric convex body is isoperimetrix of a Minkowski metric [18; Theorem 4.2, p. 63], etc.

CONJECTURE 3. If we consider those S's which have the same Minkowski surface area for all Minkowski metrics, the extremal S_M is $2^{-\frac{1}{n-1}}$ times the Blaschke sum of the above S, and $-S$. This would follow from Conjecture 2 and Theorem 3.

8. We can ask what milder conditions determine a *convex* S uniquely instead of its surface area with respect to all Minkowski metrics (compare Theorem 4). A specific result on this question is Theorem 3 formulated in the introduction.

PROOF OF THEOREM 3. Choose the fixed ellipsoid to be the unit ball in R^n. Consider an ellipsoid $\Sigma a_{ij} x_i x_j = 1$, $a_{ij} = a_{ji}$. Its intersection with the plane $\Sigma b_i x_i = 0$ has the area

$$\sqrt{\det(c_{ij})}\sqrt{\frac{\Sigma b_i^2}{\Sigma c_{ij} b_i b_j}}\,\varkappa_{n-1}, \tag{33}$$

where (c_{ij}) is the inverse of (a_{ij}). (To prove this formula, select new coordinate system x_i' in which the intersection is given by $x_n' = 0$, and a point with parallel tangent plane by $(0, \ldots, 0, 1)$.) Hence by (25) the surface area of S in this metric is

$$(\det(c_{ij}))^{-\frac{1}{2}} \int (\Sigma c_{ij} \cos\varphi_i \cos\varphi_j)^{\frac{1}{2}}\,dS, \tag{34}$$

where the $\cos\varphi_i$'s are the direction cosines of the unit normal vector at dS. Here $c_{ij} = \delta_{ij} + d_{ij}$, d_{ij} "small", $d_{ji} = d_{ij}$. Consequently,

$$\int (1 + \Sigma d_{ij} \cos\varphi_i \cos\varphi_j)^{\frac{1}{2}}\,dS \tag{35}$$

is known for all small values of d_{ij}. However, the integral is analytic in d_{ij}, and the power series expansion can be integrated term by term. Thus (35) is equal to

$$\sum_{k_{ij}} \frac{\frac{1}{2} \cdots \left(\frac{1}{2} - \sum\limits_{i \le j} k_{ij} + 1\right)}{\prod\limits_{i \le j} k_{ij}!}\, 2^{\sum\limits_{i \le j} (1-\delta_{ij}) k_{ij}} \prod_{i \le j} d_{ij}^{k_{ij}} \int \prod_{i \le j} (\cos\varphi_i \cos\varphi_j)^{k_{ij}}\,dS. \tag{36}$$

Here the outer summation extends over all $\binom{n}{2}$-tuplets of $k_{ij}\geqq 0$ integers, and i, j ranges over $1, \ldots, n$ with the condition $i\leqq j$ in all sums and products. By hypothesis, we know all coefficients of this power series, in particular, all

$$\int \prod_{i\leqq j} (\cos\varphi_i \cos\varphi_j)^{k_{ij}}\, dS = \int_{S^{n-1}} \left(\prod_{i\leqq j} \cos\varphi_i \cos\varphi_j\right)^{k_{ij}} dS(v) \tag{37}$$

for all $k_{ij}\geqq 0$.

Now consider the algebra $\mathcal{A}$ of continuous functions $f\colon S^{n-1}\to R$ satisfying $f(-v)=f(v)$; this is isomorphic to the algebra of all continuous functions on the real projective space, quotient of S^{n-1}. In this algebra, the linear combinations of all

$$\prod_{i\leqq j} (\cos\varphi_i \cos\varphi_j)^{k_{ij}} \tag{38}$$

form a subalgebra containing the constant functions ($k_{ij}=0$), which separates points of the projective space. In fact, if $x, x'\in S^{n-1}$ have coordinates $\cos\varphi_i$, $\cos\varphi_i'$ respectively, and $\cos^2\varphi_i=\cos^2\varphi_i'$, $\cos\varphi_i\cos\varphi_j=\cos\varphi_i'\cos\varphi_j'$, then $|\cos\varphi_i|= =|\cos\varphi_i'|$, and either $\operatorname{sg}\varphi_i=\operatorname{sg}\varphi_i'$ for all i, or $\operatorname{sg}\varphi_i=-\operatorname{sg}\varphi_i'$ for all i. That is: $x'=\pm x$. By the Stone-Weierstrass theorem [8; IV. 6, 16] applied to the real projective space, this subalgebra is dense in the supremum norm in $\mathcal{A}$. Thus we know all $\int f(v)\,dS(v)$, $f\in\mathcal{A}$. Using $dS(v)=dS(-v)$, hence for any continuous $f\colon S^{n-1}\to R$ we know $\int f(v)\,dS(v)=\int \frac{1}{2}\big(f(v)+f(-v)\big)\,dS(v)$. Thus by [8; IV. 6. 3] and [19] the regular Borel measure $dS(v)$ is known, and the surface is determined by [7] and [2].

REMARK. Suppose we know the surface area of a surface S, *symmetric to all coordinate planes,* with respect to all Euclidean metrics with unit sphere

$$\Sigma c_i x_i^2 = 1, \tag{39}$$

c_i near 1. Then, with the help of the expansion,

$$(1+\Sigma d_i \cos^2\varphi_i)^{\frac{1}{2}} = \sum_{k_i} \frac{\frac{1}{2}\cdots\left(\frac{1}{2}-\Sigma k_i+1\right)}{\Pi k_i!}\, \Pi d_i^{k_i}\cos^{2k_i}\varphi_i \tag{40}$$

one sees similarly that S is uniquely determined. If we omit the symmetry condition in Theorem 3 or above, we can determine only the Blaschke sum of S and $-S$, or that of the 2^n symmetric images, respectively.

9. One of the authors had the distinct privilege and great pleasure to do some joint work with the late Professor Rédei, and learn from him about the beautiful geometric work of Minkowski [16], [17]. Centrally symmetric convex bodies were as important in [10] as here, but possibly for different reasons: "They remind you of Number Theory, don't they?" said Rédei one day.

REMARK added in proof. In a joint paper by I. Bárány, K. Böröczky, J. Pach and the second named author it is proved that convexification increases volume in R^n, for polyhedra and for sufficiently smooth surfaces. Also an analogue of Theorem 4 is proved there, stating that greater integrals $\int f(v)\,dS(v)$ (for each $f \geqq 0$) imply a greater volume (although inclusion of a translate is false for $n \geqq 3$, even for boxes). So the main questions remaining are that concerning $\sigma(h)$ in (24), and the extension of the first mentioned result (or an eventual one about $\sigma(h)$) to possibly general surfaces. Also in *Period. Math. Hungar.* **14** (1983), 111—114 the present authors posed Problem 31 about questions, and a partial result related to possible generalizations of Proposition 4 to R^n.

REFERENCES

[1] Alexandroff, A. D., Zur Theorie der gemischten Volumina von konvexen Körpern I. Verallgemeinerung einiger Begriffe der Theorie von konvexen Körper, *Mat. Sb.* **2** (1937), 947—972 (Russian with German summary). *Zbl.* **17**, 426.

[2] Alexandroff, A. D., Zur Theorie der gemischten Volumina von konvexen Körpern II. Neue Ungleichungen zwischen den gemischten Volumina und ihre Anwendungen, *Mat. Sb.* **2** (1937), 1205—1238 (Russian with German summary). *Zbl.* **18**, 276.

[3] Bonnesen, T. und Fenchel, W., *Theorie der konvexen Körper*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Bd. 3, Heft 1, Springer-Verlag, Berlin, 1934. *Zbl.* **8**, 77.

[4] Busemann, H., The isoperimetric problem in the Minkowski plane, *Amer. J. Math.* **69** (1947), 863—871. *MR* **9**—372.

[5] Busemann, H., The isoperimetric problem for Minkowski area, *Amer. J. Math.* **71** (1949), 743—762. *MR* **11**—200.

[6] Busemann, H., The geometry of Finsler spaces, *Bull. Amer. Math. Soc.* **56** (1950), 5—16. *MR* **11**—400.

[7] Busemann, H., *Convex surfaces*, Interscience Tracts in Pure and Applied Mathematics, no. 6, Interscience Publishers, Inc., New York; Interscience Publishers Ltd., London, 1958. *MR* **21** # 3900.

[8] Dunford, N. and Schwarz, J. T., *Linear operators*, I. General Theory, Pure and Applied Mathematics, Vol. 7, Interscience Publishers, Inc., New York; Interscience Publishers, Ltd. London, 1958. *MR* **22** # 8302.

[9] Fáry, I., Absolute subcomplexes, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **4** (1977), 433—471. *MR* **57** # 7576.

[10] Fáry, I. und Rédei, L., Der zentralsymmetrische Kern und die zentralsymmetrische Hülle von konvexen Körpern, *Math. Ann.* **122** (1950), 205—220. *MR* **12**—526.

[11] Federer, H., Hausdorff measure and Lebesgue area, *Proc. Nat. Acad. Sci. USA* **37** (1951), 90—94. *MR* **13**—831.

[12] Federer, H., *Geometric measure theory*, Die Grundlehren der mathematischen Wissenschaften, Bd. 153, Springer-Verlag New York Inc., New York, 1969. *MR* **41** # 1976.

[13] Fenchel, W. und Jessen, B., Mengenfunktionen und konvexe Körper, *Danske Vid. Selsk., Mat.-Fys. Medd.* **16** (1938), nr. 3, 1—31. *Zbl.* **18**, 424.

[14] Firey, W. J., Blaschke sums of convex bodies and mixed bodies, *Proc. Colloquium on Convexity* (Copenhagen, 1965), Københavns Univ. Mat. Inst., Copenhagen, 1967, 94—101. *MR* **36** # 3231.

[15] Kneser, H. und Süss, W., Die Volumina in linearen Scharen konvexer Körper, *Mat. Tidsskr.* B 1 (1932), 19—25. *Zbl.* **4**, 130.

[16] Minkowski, H., *Diophantische Approximationen. Eine Einführung in die Zahlentheorie*, Chelsea Publishing Co., New York, 1957. *MR* **19**—124.

[17] Minkowski, H., *Gesammelte Abhandlungen von Hermann Minkowski*, Chelsea Publishing Co., New York, 1967.

[18] Petty, C. M., On Minkowski geometries, USC Doctoral Dissertation, Los Angeles, Calif., 1952.

[19] Saks, S., Integration in abstract metric spaces, *Duke Math. J.* **4** (1938), 408—411. *Zbl.* **19**, 170.

[20] Tverberg, H., A proof of the Jordan curve theorem, *Bull. London Math. Soc.* **12** (1980), 34—38. *MR* **81c**: 30045.

[21] WHITNEY, H., *Geometric integration theory*, Princeton University Press, Princeton, N. J., 1957. *MR* **19**—309.
[22] KOŚNIOWSKI, C., *A first course in algebraic topology*, Cambridge University Press, Cambridge—New York, 1980. *MR* **82c**: 57001.
[23] SCOTT, P. R., On a family of polygons, *Math. Mag.* **55** (1982), 104—105.

(Received June 28, 1982)

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CA 94720
USA

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

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 159—175

ON FACE-VECTORS OF MAPS WITH CONSTANT WEIGHT OF EDGES

S. JENDROL' and E. JUCOVIČ

By the weight of an edge AB of a graph we mean the sum of valencies of the vertices A, B. All regular graphs have of course constant weights of edges. If the connected graph with constant weight of edges is not regular then its vertex set decomposes into two disjoint sets; the vertices of one have valency a, the vertices of the second set have valency b; no two vertices of the same set are adjacent. Therefore such graphs or maps which are formed on surfaces by embedding of such graphs are called (a, b)-graphs or (a, b)-maps, respectively.

Basic combinatorial properties of 3-dimensional convex polytopes (in short polytopes in the sequel) with constant weight of edges are studied in Rosenfeld [4]. (a, b)-graphs (under a different name) are investigated in Acharya-Vartak [1].

By the face-vector or vertex-vector of a map M we mean the sequence $(p_i(M))$ or $(v_i(M))$ where $p_i(M)$ or $v_i(M)$ denotes the number of i-gonal faces or i-valent vertices of the map M, respectively.

The aim of the present paper is to contribute to the problem of characterizing face-vectors of (a, b)-maps on closed surfaces. The graphs of all maps investigated in the paper are 3-connected. (In case of planar maps these maps are combinatorially isomorphic with convex 3-polytopes by the well-known Steinitz's theorem — see Grünbaum [2]). In Section 1 we prove some lemmas allowing to construct (a, b)-maps on a closed orientable surface T_g of genus g disposing of (a, b)-maps on surfaces of lower genus and we state a general theorem concerning the existence of (a, b)-maps on T_g with mutually different face-vectors for given (a, b) and g. In Sections 2, 3 and 4 we deal with (3,6), (3,5) and (3,4)-maps, respectively.

1

From Euler's relation for maps on the closed orientable surface T_g of genus g, $f+v-h=2(1-g)$ (here f, v or h denotes the number of faces, vertices or edges of the map, respectively) there follows easily for an (a, b)-map M on T_g the following relation:

$$\sum_{i\geq 3}(2ab-abi+ai+bi)p_i(M)=4ab(1-g). \tag{1.1}$$

In case $g=0$ or $g=1$, (a, b)-maps do exist only for pairs $(a, b)\in\{(3, 4), (3, 5)\}$ or $\{(3,4), (3,5), (3,6)\}$, respectively. — From the bichromaticity of the graph of an

1980 *Mathematics Subject Classification.* Primary 05C10; Secondary 52A25.
Key words and phrases. Face-vector of a map, (a, b)-map, orientable surface.

(a, b)-map follows the evenness of the lengths of all its circuits (cf. Harary [3]); hence we have

LEMMA 1.1. *If the sequence* $(p_i, i \geqq 3)$ *is the face-vector of an* (a, b)*-map then for all odd* j *we have* $p_j = 0$.

From the obvious relations $h = av_a = bv_b$, $v = v_a + v_b$ and by means of Euler's relation mentioned before we get

$$f = \frac{(a-1)b-a}{b} v_a + 2(1-g). \tag{1.2}$$

By an α-configuration we mean a triple of quadrangles one of which separates the remaining two (Fig. 1.1 where the vertices marked A_i or B_i are a-valent or b-valent, respectively).

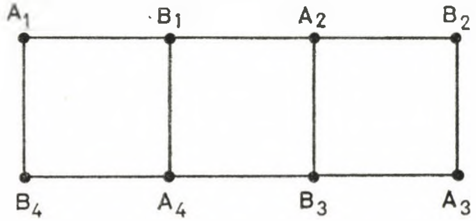


Fig. 1.1

LEMMA 1.2. *If an* (a, b)*-map* M *with the face-vector* $(p_i, i \geqq 3)$ *containing two* α*-configurations exists on the orientable closed surface* T_g *of genus* g *then on the closed orientable surface* T_{g+1} *there does exist an* (a, b)*-map* M' *with the face-vector* $(p'_i, i \geqq 3)$ *where*

$$p'_k = p_k \quad \text{for all} \quad k \neq 4, 6,$$
$$p'_4 = p_4 - 6,$$
$$p'_6 = p_6 + 4.$$

LEMMA 1.3. *Let* (a, b)*-maps* M_1 *or* M_2 *on orientable surfaces* T_{g_1} *or* T_{g_2} *with face-vectors* $(r_i, i \geqq 3)$ *or* $(s_i, i \geqq 3)$*, respectively, do exist and let each of these admit an* α*-configuration. Then there exists on the orientable surface* $T_{g_1+g_2}$ *of genus* g_1+g_2 *an* (a, b)*-map* M *with face-vector* $(p_i, i \geqq 3)$ *so that*

$$p_4 = r_4 + s_4 - 6,$$
$$p_6 = r_6 + s_6 + 4,$$
$$p_k = r_k + s_k \qquad \text{for all} \quad k \neq 4, 6 \quad \text{hold.}$$

PROOF OF LEMMA 1.2.

Let us denote the a-valent or b-valent vertices of one α-configuration A_1, A_2, A_3, A_4 or B_1, B_2, B_3, B_4, respectively, and analogously the vertices of the second α-configuration are marked A'_i and B'_i, $i = 1, 2, 3, 4$. (The marking of the vertices be such as in Fig. 1.1.) We choose inner points C_1, D_1 and C_2, D_2 on the edge A_2B_3 or A_4B_1, respectively, as new vertices, so that the vertex C_1 or C_2 is adjacent with the vertex A_2 or A_4, respectively. Analogously in the second α-configuration new vertices C_1', D_1', C_2', D_2' on the edges $A_2'B_3'$, $B_1'A_4'$ are chosen.

Delete from the surface T_g carrying the so arranged map the regions bounded by the graph-circuits $A_2C_1D_1B_3A_4C_2D_2B_1A_2$ and $A_2'C_1'D_1'B_3'A_4'C_2'D_2'B_1'A_2'$. After glueing these holes we get the desired map on T_{g+1}. The glueing is performed so that we identify the pairs of vertices A_2 and D_2', C_1 and B_1', D_1 and A_2', B_3 and C_1', A_4 and D_1', C_2 and B_3', D_2 and A_4' as well as the appropriate edges. The quadrangles $B_1A_2B_3A_4$, $B_1'A_2'B_3'A_4'$ disappear from the original map, and the quadrangles $A_1B_1A_4B_4$, $A_2B_2A_3B_3$, $A_1'B_1'A_4'B_4'$, $A_2'B_2'A_3'B_3'$ are changed into hexagons.

The proof of Lemma 1.3 proceeds analogously as the proof of Lemma 1.2.

THEOREM 1.1. *Let a, b be integers,* $(a, b) \notin \{(3, 4), (3, 5), (3, 6)\}$, $3 \leqq a < b$. *On every orientable closed surface* T_g *of genus g there exists at most a finite number of (a, b)-maps with mutually different face-vectors.*

To prove Theorem 1.1 we need some lemmas.

LEMMA 1.4. *Let* $3 \leqq a < b$ *be integers such that* $(a, b) \notin \{(3, 4), (3, 5), (3, 6)\}$. *Then for every integer* $i \geqq 4$

$$2ab - abi + ai + bi < 0$$

holds.

Suppose the contrary of Lemma 1.4, i.e. $2ab - abi + ai + bi \geqq 0$ holds. From this follows

$$(*) \qquad \frac{ai}{ai - 2a - i} \geqq b.$$

Two cases have to be distinguished. If $a = 3$ then $b \geqq 7$, and from $(*)$ we get $i \leqq \frac{42}{11}$ in contradiction with $i \geqq 4$. If $a \geqq 4$ then $b \geqq 5$ and from $(*)$ we get $i \leqq \frac{10a}{4a-5}$, again a contradiction to $i \geqq 4$.

As a corollary of Lemma 1.4 and the relation (1.1) we have

LEMMA 1.5. *For every (a, b)-map M on the surface* T_g, $3 \leqq a < b$, $(a, b) \notin \{(3, 4), (3, 5), (3, 6)\}$, *and every* $i \geqq 4$ *we have*

$$p_i(M) \leqq \frac{4ab(1-g)}{2ab - abi + ai + bi}.$$

Employing Lemma 1.4 and Lemma 1.5 we have

LEMMA 1.6. *For every (a, b)-map M on the surface* T_g, $3 \leqq a < b$, $(a, b) \notin \{(3, 4), (3, 5), (3, 6)\}$, *and every* $j > \frac{2ab(2g-1)}{ab-a-b}$, $p_j(M) = 0$ *holds.*

Theorem 1.1 follows now directly from Lemma 1.5 and Lemma 1.6.

In Sections 2—4 it will be shown that the couples of integers (3, 4), (3, 5) and (3, 6) are indeed exceptional; the assertion of Theorem 1.1 does not hold for them.

2

As follows from (1.1), the following is a necessary condition in order that a sequence of nonnegative integers (p_i, $i \geqq 3$) should be the face-vector of a (3, 6)-map on T_g:

$$\sum_{k \geqq 3} (4-k) p_k = 8(1-g). \tag{2.1}$$

Let us characterize first face-vectors of (3, 6)-maps on the torus (T_1).

Theorem 2.1. *A sequence* (p_i, $i \geqq 3$) *of non-negative integers is the face-vector of a* (3, 6)*-map on the torus if and only if*

$$p_k = 0 \quad \textit{for} \quad k \neq 4, \quad \textit{and}$$

$$p_4 \equiv 0 \pmod 3, \quad p_4 \neq 0 \quad \textit{hold.}$$

Proof. The necessity of the conditions mentioned follows from the relations (2.1) and (1.2). The sufficiency is proved by construction of the maps: In Fig. 2.1 is drawn a toroidal (3, 6)-map containing $3t$ quadrangles, $t=1, 2, \ldots$ (first identify the pairs of vertices A_0 and A_t, B_0 and B_t, and after this identify the equally labelled vertices).

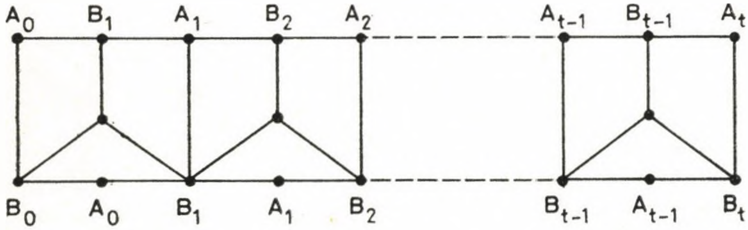


Fig. 2.1

A complete characterization of the face-vectors of (3, 6)-maps on orientable surfaces of genus $g \geqq 2$ is still lacking. The next two theorems yield partial results. From (2.1) follows directly

Theorem 2.2. *In the face-vector* (p_i, $i \geqq 3$) *of every* (3, 6)*-map on every orientable surface* T_g *we have* $p_k=0$ *for every* $k \geqq 8g-3$.

Theorem 2.3. *A sequence* (p_i, $i \geqq 3$) *of non-negative integers, for which* $p_k=0$ *for all* $k \neq 4, 6$ *holds, is the face-vector of a* (3, 6)*-map on the orientable surface* T_g *of genus* $g \geqq 2$ *if and only if*

$$p_6 = 4(g-1) \quad \textit{and} \quad p_4 \equiv 0 \pmod 3) \quad \textit{hold.}$$

Proof. The necessity of the conditions mentioned follows from (1.2) and (2.1). Sufficiency will again be proved by construction of the required map; an inductive procedure is used.

Denote M_1 the toroidal (3, 6)-map in Fig. 2.2, it consists of 6 quadrangles in two α-configurations (the quadrangles of one are marked α_1).

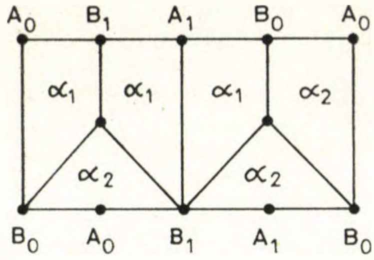


Fig. 2.2

Let M_k, $k=2, 3, \ldots, g-2$, denote the (3, 6)-map on the orientable surface T_k obtained by employing Lemma 1.3 to the (3, 6)-maps M_{k-1} and M_1.

We employ now Lemma 1.3 to the maps M_{g-2} and the toroidal (3, 6)-map containing $3n+6$ quadrangles, $n=0, 1, \ldots,$ (see Fig. 2.1) to get the map M_{g-1} on the surface T_{g-1}; it contains $4(g-2)$ hexagons and $3n+6$ quadrangles. The required map is obtained after using Lemma 1.2 with the map M_{g-1}.

Remark 1. The maps constructed in the proofs of Theorems 2.1 and 2.3 are in general not cell-complexes. If we require the maps to be cell-complexes some assumption concerning the number p_4 must be added. Our constructions (after small changes) work for all $p_4 \geqq 3(4g+1)$ where g is the genus of the surface carrying the map. For the torus $(g=1)$ this bound is sharp; there do not exist (3, 6)-maps on the torus with <15 quadrangles which would be cell-complexes.

3

From (1.1) there follows for the face-vector $(p_i, i \geqq 3)$ of a (3, 5)-map on the surface T_g the following relation:

$$\sum_{i \geqq 3} (30-7i)p_i = 60(1-g). \tag{3.1}$$

The following lemma is very useful in getting results about the existence of (3, 5)-maps on closed surfaces.

Lemma 3.1. *If there exists on T_g a map M_q with a 2-connected regular 3-valent graph and a face-vector $(q_i, i \geqq 3)$ with $q_i=0$ for odd i, then there exists on T_g a (3, 5)-map M with a 3-connected graph and a face-vector $(p_i, i \geqq 3)$ such that*

$$p_i = q_i \quad \textit{for all} \quad i \neq 4$$

and

$$p_4 = 30(1-g) + \frac{1}{2} \sum_{\substack{i \geqq 3 \\ i \neq 4}} (7i-30) q_i$$

hold.

PROOF. The map M_q is first changed into the map M' using the transformation replacing vertices by hexagons (see Grünbaum [2, p. 265]). The graph of the map M' is again regular 3-valent. To an adjacent pair of faces k-gon $-l$-gon in M_q there is associated in M' a pair of faces k-gon $-l$-gon separated by a couple of adjacent hexagons. (See Fig. 3.1 where a part of the original map is drawn by dashed lines.) The graph of the map M' is 3-connected because to two edge-disjoint paths of M_q there are associated in M' four paths from which at least three are edge-disjoint.

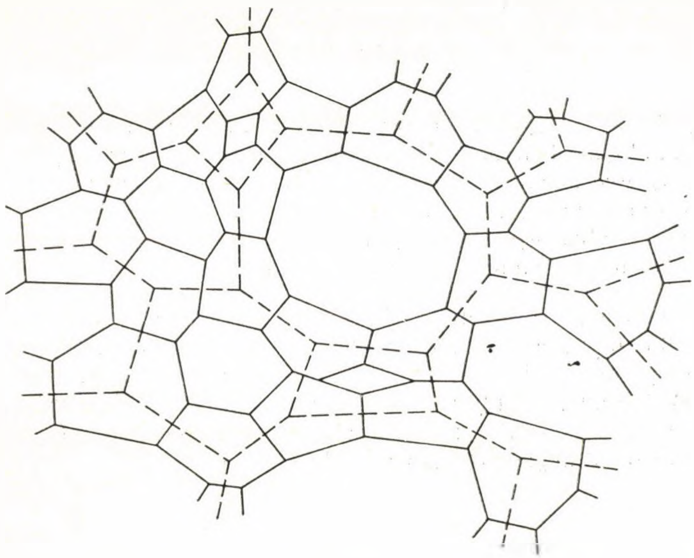


Fig. 3.1

All faces of M_q are even-gonal and the same holds for all faces of the map M'. The graph of M' is therefore bichromatic with the vertex set decomposed into two disjoint sets $\mathcal{T}_1$, $\mathcal{T}_2$ so that no two vertices of the same set are adjacent. To the vertices and edges of the map M' (which all remain vertices and edges of the constructed map M) add following new vertices and edges: Inside of every hexagon ω of M', which is by the transformation replacing vertices by hexagons associated to a vertex of M_q, choose a point, and join it with all vertices from the set $\mathcal{T}_1$ which are vertices of the face ω. In this way we decomposed all hexagons not occurring in the map M_q into quadrangles. Every vertex from the set $\mathcal{T}_1$ becomes 5-valent, and no vertex from the set $\mathcal{T}_2$ has changed its valency; the new vertices have valency 3 and are all joined by edges with vertices of valency 5. The graph of the map M is therefore a (3, 5)-graph. The types of all faces of the map M_q appear in M as well. In M we have more quadrangles; their number follows from (3.1).

THEOREM 3.1. *The sequence* $(p_i, i \geqq 3)$ *of non-negative integers is the face-vector of a toroidal* (3, 5)*-map if and only if it satisfies the conclusion of Lemma 1.1 and* (3.1).

The necessity of the conditions mentioned in Theorem 3.1 was already proved. Crucial in the proof of their sufficiency is the following

LEMMA 3.2. *To every sequence* $(q_i, i \geqq 3)$ *of non-negative integers satisfying the conditions*

$$q_j = 0 \quad \text{for all odd } j,$$

$$q_4 = \frac{1}{2} \sum_{\substack{k \geqq 3 \\ k \neq 4}} (k-6)\, q_k$$

there exists a map M *on the torus with a 3-connected 3-valent graph such that for all* i *we have* $q_i(M) = q_i$.

PROOF. It proceeds by construction of M. Let $d = \sum_{i \geqq 6} q_i$. The starting toroidal map containing d hexagons is in Fig. 3.2a or 3.2b for even or odd number $d \geqq 2$, respectively. (Equally labelled vertices on the marked "holes" are identified.) The case $d=1$ will be investigated at the end.

From each of the d hexagons $A_j A_{j+1} A_{j+2} B_{j+2} B_{j+1} B_j$, $j \leqq d-2$, in Fig. 3.2 a required $2t$-gon, $t \geqq 3$, will be formed as follows. Distinct inner points $C_1, C_2, \ldots$ $\ldots, C_{t-3}$ of the edge $A_j A_{j+1}$ are taken as new vertices and analogously we choose points $D_1, D_2, \ldots, D_{t-3}$ of the edge $B_j B_{j+1}$. Pairs of vertices $C_k, D_k, k = 1, 2, \ldots, t-3$ are joined by edges so that no two edges intersect. So a $2t$-gon is formed and "new" $t-3$ quadrangles appear in the map. — We proceed quite analogously with the other hexagons in Fig. 3.2. So we get the required toroidal map M in case $d \geqq 2$.

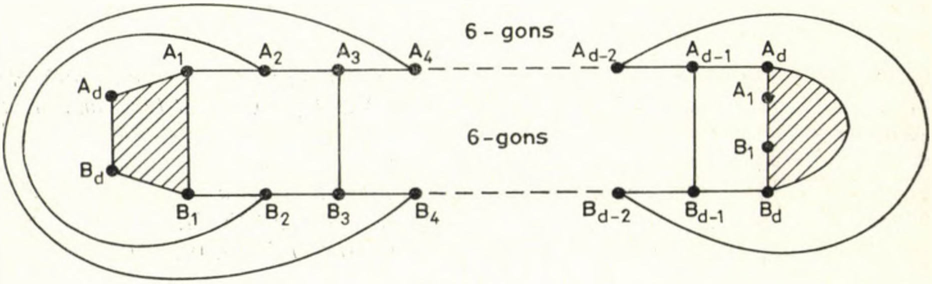


Fig. 3.2a

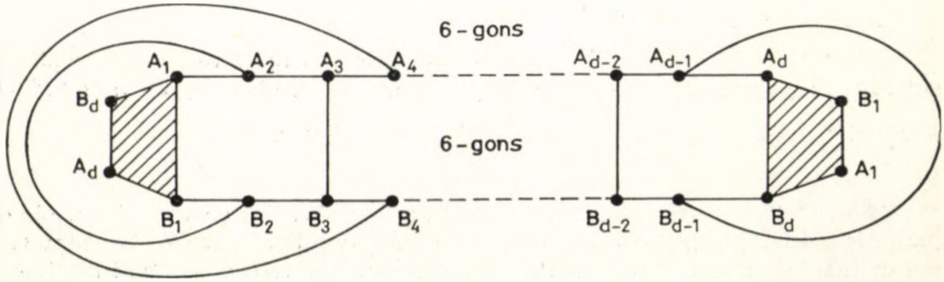


Fig. 3.2b

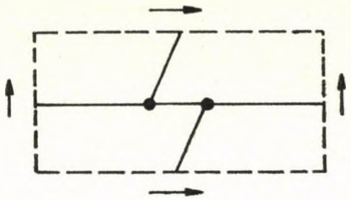


Fig. 3.3a

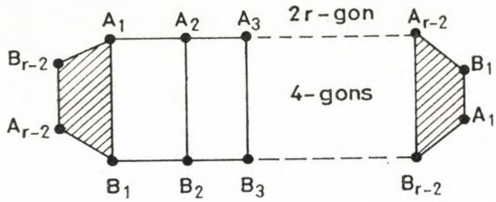


Fig. 3.3b

In case $d=1$, the required toroidal map containing one hexagon or one $2r$-gon, $r\geqq 4$, is drawn in Fig. 3.3a or 3.3b, respectively.

PROOF OF THEOREM 3.1. The sequence $(q_i, i\geqq 3)$ for which $q_k=p_k$ for all $k\geqq 3$, $k\neq 4$ and $q_4=\frac{1}{2}\sum_{\substack{k\geqq 3\\ k\neq 4}}(k-6)p_k$ holds satisfies the conditions of Lemma 3.2. Then we use Lemma 3.1.

THEOREM 3.2. *Every sequence* $(p_i, i\geqq 3)$ *of non-negative integers satisfying — besides the conclusions of Lemma 1.1 and* (3.1) *with* $g=0$ *— at least one of the conditions*

(A) $p_j\equiv 0 \pmod 2$ *for all* $j\geqq 6$,

(B) $p_6\geqq 2$ *with the exception of the sequences* $(p_4=55,\ p_6=2,\ p_8=1,\ p_j=0$ *for* $j\neq 4, 6, 8)$, $(p_4=68,\ p_6=3, p_{10}=1,\ p_j=0$ *for* $j\neq 4, 6, 10)$, *or*

(C) $$\sum_{j\geqq 1} p_{4j+2}\geqq 6$$

is the face-vector of a (3, 5)*-polytope.*

The proof leans heavily on Lemma 3.1. So we first prove the existence of a planar map with a regular 3-valent 2-connected graph and face-vector $(q_i, i\geqq 3)$ for which $q_j=p_j$ for $j\geqq 3$, $j\neq 4$ and $q_4=6+\frac{1}{2}\sum_{\substack{j\geqq 3\\ j\neq 4}}(j-6)p_j$ holds.

Case (A). The starting map is the map of the cube. We have in it a triple of quadrangles from which one separates the remaining two. Such a triple of quadrangles appear in the map at any step of the construction. This step consists of creating a pair of required t-gons, $t\geqq 6$, in such a way that the "middle" quadrangle of the triple of quadrangles is decomposed by $t-4$ new edges into $t-3$ quadrangles (see. Fig. 3.4). We proceed in this way until a trivalent planar map with face-vector (q_i) is obtained.

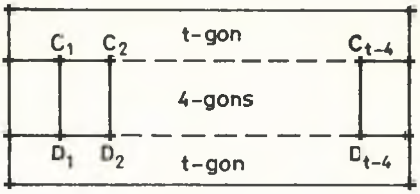


Fig. 3.4

Case (B). Let $d=\sum_{j\geq 3} q_{2j}$. If $\sum_{k\geq 4} q_{2k}\geq 2$, the starting maps are in Fig. 3.2a or 3.2b depending on whether d is even or odd, respectively. However, the marked regions are not "holes" but are decomposed as shown in Fig. 3.5. The planar map

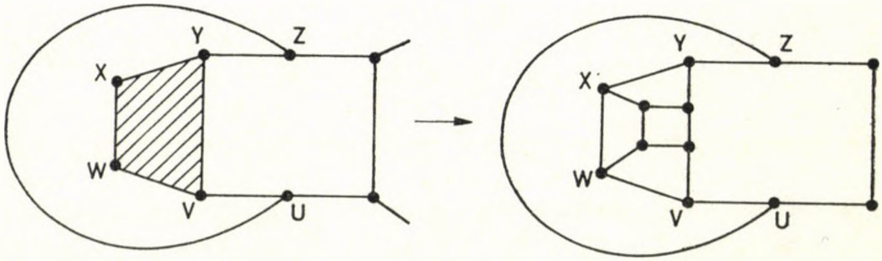


Fig. 3.5

obtained is 2-connected and admits two octagons and $d-2$ hexagons. The hexagon $XYZUVW$ as well as the second hexagon obtained analogously at decomposing the second hole are retained in the map. From the octagons and the remaining hexagons we form the other faces required using the procedure described in the proof of Lemma 3.2.

Let $\sum_{k\geq 4} q_{2k}=1$. If a $2r$-gon, $r\geq 6$, is required and q_6 is even or odd the starting maps are in Fig. 3.6a or 3.6b, respectively. If a 10-gon is needed and q_6 is even or odd

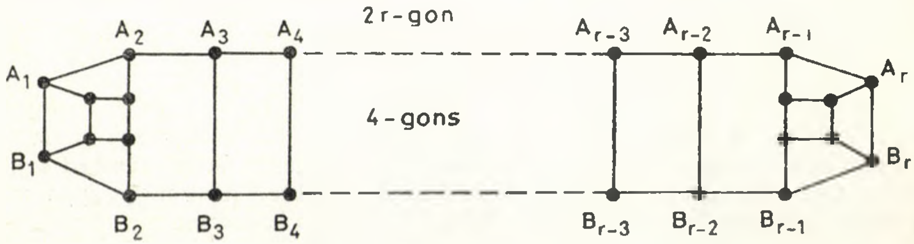


Fig. 3.6a

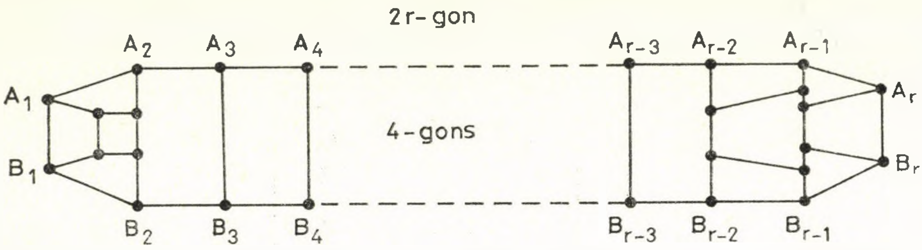


Fig. 3.6b

the starting map is that in Fig. 3.6a ($r=5$) or 3.7a, respectively. If an octagon is needed and q_6 is even or odd the starting map is in Fig. 3.7c or 3.7b, respectively. In all these cases q_6-2 or q_6-3 (or q_6-5 if $r=5$ or q_6-4 if $r=4$) required hexagons are formed in pairs from the triples of quadrangles which occur in the maps; the construction was described in case (A) (Fig. 3.4).

It remains to settle the cases when $\sum_{k\geq 4} q_{2k}=0$, i.e. when only hexagons and quadrangles are required. If q_6 is even we start with the map of the cube and use one triple of its adjacent quadrangles for creating the q_6 hexagons (Fig. 3.4). If q_6 is odd, the starting planar map containing three hexagons is in Fig. 3.7d (full lines). Notice

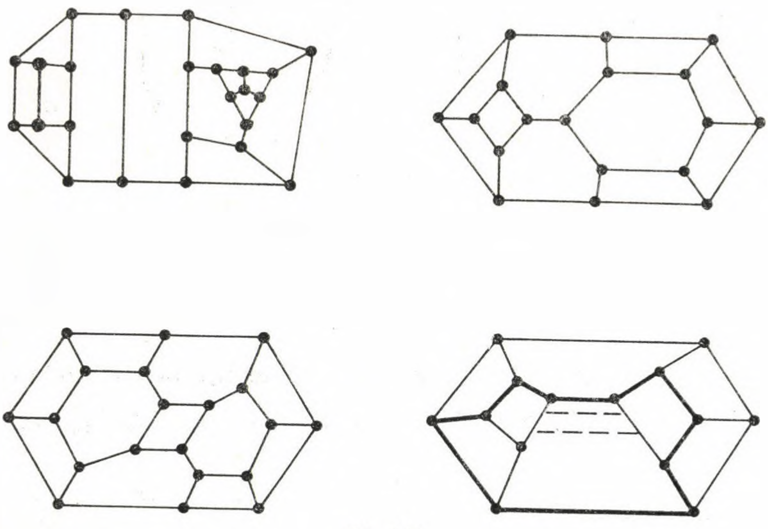


Fig. 3.7

that its submap bounded by the heavy lines is reproducable in course of increasing the number of hexagons by two. This proceeds by inserting the two dashed edges. Again a face-aggregate consisting of one hexagon and three quadrangles appears in the map and is used for further increasing the number of hexagons.

Case (C) splits into two subcases:

(C_1) If among the required faces there are at least four with numbers of vertices $k, l, m, n \equiv 2 \pmod 4$ such that $k \neq l, m \neq n$, then every pair of these required faces is used for filling of one hole in the map in Fig. 3.2a or 3.2b. (Fig. 3.2a or 3.2b is used in case $d \equiv 0 \pmod 2$ or $d \equiv 1 \pmod 2$, respectively, where $d = \sum_{j \geq 3} q_{2j}$.) Let us describe this procedure for $k = 4u+2 < l = 4v+2$, $u \geq 1$. First the "hole" is arranged as shown in Fig. 3.5. Then $4(u-1)$ inner points $E_1, F_1, G_1, H_1, E_2, F_2, \ldots, E_{u-1}, F_{u-1}, G_{u-1}, H_{u-1}$ of the edge YZ are chosen as new vertices. They are changed into trivalent vertices as it is drawn in Fig. 3.8. In the next step analogously on the edge $H_{u-1}Z$ (or YZ if $u=1$) z new vertices $C_1, C_2, \ldots, C_z$ are chosen, where $z = 2(v-u)-1$. On the edge VU new vertices $D_1, D_2, \ldots, D_z$ are chosen, and the vertices C_i, D_i are joined by mutually non intersecting edges. So a k-gon and an l-gon are created. — In the same way from the second hole in the maps in Fig. 3.2 an m-gon and an n-gon are obtained.

Fig. 3.8

At the end of this step we get a 3-valent planar map containing four of the faces required, $d-4$ hexagons and quadrangles. From the hexagons the remaining required faces with ≥ 6 edges are constructed as described in Lemma 3.2.

(C_2) If the situation of case (C_1) does not occur then either there is an integer $k \geq 2$ such that $q_{4k+2} \geq 6$ or there are two integers $k \geq 2$, $l \geq 1$, $k \neq l$, such that $q_{4k+2} = 5$, $q_{4l+2} = 1$ holds. In the first case we start with the maps in Fig. 3.2 containing $d-2$ hexagons. The "holes" are arranged as in Fig. 3.9. On the edge WX, $4(k-1)$ new vertices $E_1, F_1, G_1, H_1, \ldots, E_{k-1}, F_{k-1}, G_{k-1}, H_{k-1}$ are chosen and changed into trivalent vertices as before (Fig. 3.8). We get two of the $(4k+2)$-gons required. Then $2(k-2)$ edges "parallel" to the edge UZ with end-points on YZ and VU are inserted to get the third $(4k+2)$-gon. The second hole is arranged analogously.

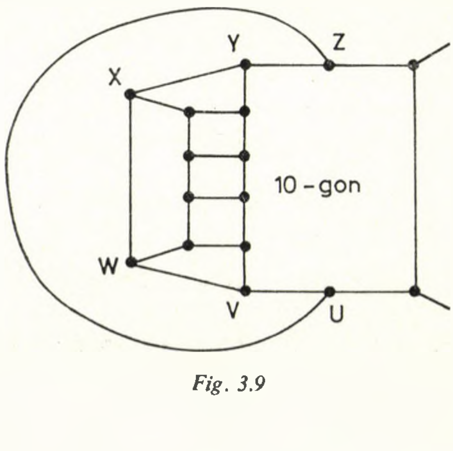


Fig. 3.9

If $q_{4k+2}=5$, $q_{4l+2}=1$, $k \neq l$, the starting map containing $d-1$ hexagons is in Fig. 3.2. One of the "holes" is replaced by three of the $(4k+2)$-gons required, as described above. The second hole is replaced by two such faces as in case (C_1). In both cases Lemma 3.2 can be applied further.

The maps constructed in all the cases (A), (B) and (C) are planar with a 3-valent 2-connected graph and face-vector $(q_i, i \geqq 3)$. Therefore by Lemma 3.1 there exists a planar (3,5)-map with face-vector $(p_i, i \geqq 3)$ and a 3-connected graph. By Steinitz's theorem mentioned at the beginning there exists a 3-polytope combinatorially isomorphic to that map. Theorem 3.2 is proved.

For (3,5)-maps on surfaces of higher genus we have

THEOREM 3.3. *Every sequence* $(p_i, i \geqq 3)$ *of non-negative integers satisfying besides the conclusion of Lemma 1.1 and* (3.1), *with* $g \geqq 2$ — *the conditions*

$$p_6 \geqq 4(g-1)$$

$$\sum_{k \geqq 7} (k-6) p_k \geqq 4(g-1)$$

is the face-vector of a (3,5)*-map on* T_g.

PROOF. First by Theorem 3.1 a toroidal (3,5) map is constructed with face-vector $(q_i, i \geqq 3)$ for which $q_4=p_4+6(g-1)$, $q_6=p_6-4(g-1)$, $q_k=p_k$ for all $k \neq 4, 6$ holds. The procedure of construction of this map ensures the existence of at least $2(g-1)$ α-configurations in it. So Lemma 1.2 can be employed further to get the desired map on the surface of desired genus g.

REMARK 2. Rosenfeld [4] has shown that there does not exist a (3, 5)-polytope P with $p_6(P)=1$, $p_4(P)=36$, $p_i(P)=0$ for $i \neq 4, 6$; therefore it is impossible to improve Theorem 3.2 (B). — The question of the existence of (3, 5)-polytopes with the face-vectors in the brackets is undecided because their constructibility by our procedure supposes the existence of trivalent planar maps with face-vectors ($p_8=1$, $p_6=2$, $p_4=7$, $p_i=0$ for $i \neq 4, 6, 8$), ($p_{10}=1$, $p_6=3$, $p_4=8$, $p_i=0$ for $i \neq 4, 6, 10$). However, it is not hard but boring to prove that these maps do not exist.

REMARK 3. We conjecture that Theorem 3.2 (C) can be improved by assuming $\sum_{j \geqq 1} p_{4j+2} \geqq 4$.

4

In this section no final results will be stated. It seems to be difficult to characterize face-vectors of (3, 4)-maps on T_g for any g. We present three procedures of construction of (3, 4)-maps allowing to obtain some partial results.

First of all let us remind that from (1.1) for the face-vector (p_i) of a (3, 4)-map on the orientable surface T_g there follows

$$\sum_{i \geqq 3} (24-5i) p_i = 48(1-g). \tag{4.1}$$

LEMMA 4.1. *Let there exist a map* M' *on the surface* T_g *with a 2-connected multigraph whose face-vector or vertex-vector is* $(q_i, i \geqq 2)$ *or* $(v_i, i \geqq 2)$, *respectively.*

Then there exists a (3, 4)*-map* M *on* T_g *with a* 3*-connected graph and a face-vector* $(p_i, i \geqq 4)$ *for which*

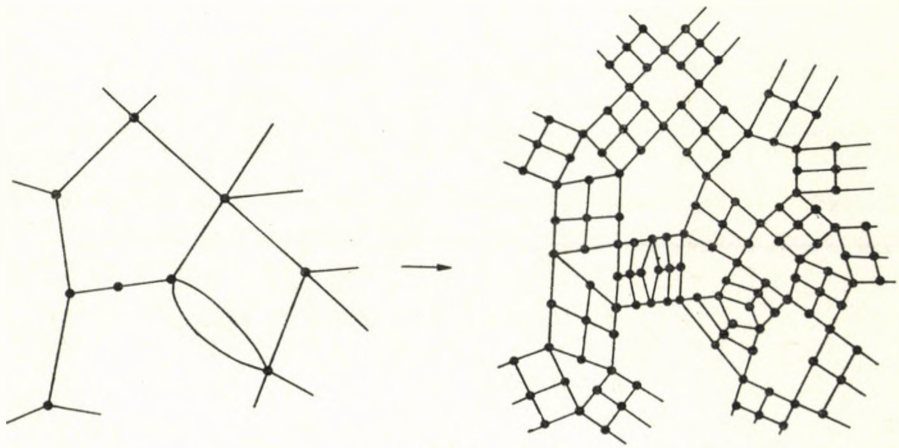
$$p_{2i} = q_i + v_i \quad \text{for all} \quad i \geqq 3,$$

$$p_{2i-1} = 0 \quad \text{for all} \quad i \geqq 2,$$

$$p_4 = 12(1-g) + \frac{1}{2} \sum_{i \geqq 3} (5i-12)(q_i+v_i)$$

hold.

PROOF. To every k-valent vertex and to every k-gonal face of the map M' associate a $2k$-gon in M, to every edge of M' associating a quadruple of quadrangles (a quadrangle decomposed by its two mid-lines). To an adjacent pair of faces k-gon — m-gon or a pair k-valent — m-valent vertex in M' there is associated a pair of faces $2k$-gon — $2m$-gon separated by the quadruple of quadrangles mentioned. If in M' an m-valent vertex is incident with a k-gon then their images in M are a $2k$-gon and a $2m$-gon which have precisely one 4-valent vertex in common. Clearly, M is a (3, 4)-map; the 3-connectedness of its graph is ensured by the fact that to every path in M' there are associated two distinct paths in M. The situation is well illustrated in Fig. 4.1 where in (a) there is a part of the map M' and in (b) its image in M. The numerical properties are clear from the procedure of construction and from (4.1).



(a) *Fig. 4.1* (b)

Our second procedure of construction is in fact a transformation of a regular 3-valent map M'' into a (3, 4)-map M with a "large" number of hexagons. It is contained in

LEMMA 4.2. *Let there exist on the surface* T_g *a map* N *with a regular* 3*-valent* 2*-connected graph and such a face-vector* $(q_i, i \geqq 3)$ *that* $q_i = 0$ *for all odd* i *holds. Then there exists on* T_g *a* (3, 4)*-map* M *with a* 3*-connected graph and face-vector*

$(p_i,\ i \geqq 3)$ *where*

$$p_k = q_k \quad \text{for all} \quad k \geqq 3,\ k \neq 4, 6,$$

$$p_6 = \frac{2}{3} q_4 + 2q_6 + \frac{1}{6} \sum_{i \geqq 7} iq_i,$$

$$p_4 = 12(1-g) + q_4 + 3q_6 + \frac{3}{2} \sum_{i \geqq 7} (i-4) q_i.$$

PROOF. The graph of the map N is bichromatic, i.e. its vertex set is decomposed into two disjoint sets T_1, T_2 such that no two vertices of the same set are adjacent. After performing on N the transformation replacing vertices by hexagons (see proof of Lemma 3.1) we get a map M' whose graph is regular 3-valent and 3-connected. Its each vertex is incident with two adjacent hexagons which are associated to two vertices, adjacent in N and therefore belonging to different sets T_i, $i=1, 2$. Denote the set of hexagons in M' associated with T_i by S_i, $i=1, 2$. So every vertex of M' is incident with one member of S_1 and one member of S_2. The graph of the map M' is bichromatic, too; its vertex set should be decomposed into independent sets R_1, R_2 (analogously as the vertex set of the map N). We choose an inner point of every hexagon from the set S_1 as a new vertex of the map M and join it by new edges with all those vertices from the set R_1 which are incident with that hexagon. So every hexagon from the set S_1 is decomposed into three quadrangles. It remains to verify that the map M obtained in such a way fulfils the assertion of Lemma 4.2.

3-connectedness of the graph of the map M follows from the 3-connectedness of the graph of the map M'. It is a (3, 4)-graph because of the following facts: Every vertex from the set R_1 became 4-valent, every vertex from the set R_2 remained 3-valent while both of these sets remained independent; the added new vertices are 3-valent all and they are joined with vertices from the set R_1 only. — For the map M there holds $p_k(M) = p_k(N) = q_k$ for all $k \geqq 3$, $k \neq 4, 6$. In the map M there remained all hexagons from the set S_2 (whose number is $\frac{v_3(N)}{2} = \frac{1}{6} \sum_{i \geqq 3} iq_i$) and all hexagons from the map N; therefore

$$p_6 = \frac{1}{6} \sum_{i \geqq 3} iq_i + q_6 = \frac{2}{3} q_4 + 2q_6 + \frac{1}{6} \sum_{i \geqq 7} iq_i.$$

The third procedure which has been employed already by Rosenfeld [4] is contained in

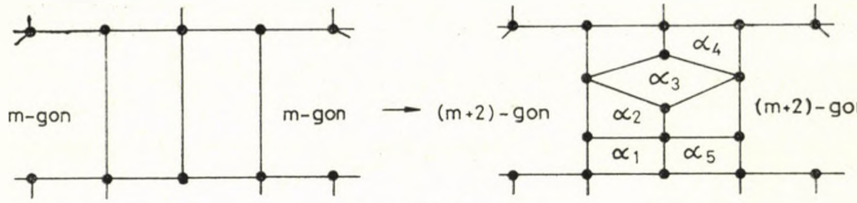
LEMMA 4.3. *Let there exist on the surface T_g a* (3, 4)*-map N with face-vector $(q_i,\ i \geqq 3)$, containing four quadrangles in a face-aggregate as in Fig. 4.2a (put $m=4$). Then there exists on T_g a* (3, 4)*-map M with face-vector $(p_i,\ i \geqq 3)$ where*

$$p_{2j} = q_{2j} + 2t_j, \quad j \geqq 3, \quad t_j = 0, 1, 2, \ldots$$

$$p_{2j+1} = q_{2j+1} = 0 \quad \text{for} \quad j \geqq 1,$$

$$p_4 = 12(1-g) + \frac{1}{2} \sum_{j \geqq 3} (5j-12)(q_{2j} + 2t_j)$$

PROOF. The quadruple of quadrangles is changed so that the two "middle" quadrangles are subdivided as drawn in Fig. 4.2 to increase the number of vertices of the lateral quadrangles by two. This procedure is repeated $(k-2)$-times to get two $2k$-gons (the quadrangles α_1, α_5 are used the second time etc.). To get another pair of required $2m$-gons, $m \geqq 3$, the quadrangles $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ are used.



(a) *Fig. 4.2* *(b)*

In the next three theorems the preceding lemmas are employed to get sufficient conditions for a sequence of numbers to be the face-vector of a (3, 4)-map.

THEOREM 4.1. *Every sequence* $(p_i, i \geqq 3)$ *of non-negative integers satisfying — besides the conclusion of Lemma 1.1 and* (4.1) *— the conditions*

$$p_i \equiv 0 \pmod 2 \quad \textit{for all} \quad i \geqq 6,$$

$$p_6 \geqq 4g$$

$$\sum_{k \geqq 3} (k-3) p_{2k} \geqq 4(g-1)$$

is the face-vector of a (3, 4)*-map on* T_g.

PROOF. The construction starts with a planar map consisting of two digons and two 2-valent vertices (and two edges). After using Lemma 4.1 we get the map of the rhombic dodecahedron. It contains two distinct chains each containing four quadrangles (Fig. 4.2a). Now use Lemma 4.3; we get a planar (3, 4)-map with a 3-connected graph and a face-vector $(q_i, i \geqq 3)$ for which

$$q_j = p_j \quad \text{for all} \quad j \neq 4, 6,$$

$$q_6 = p_6 - 4g, \quad \text{and}$$

$$q_4 = p_4 + 6g \quad \text{holds.}$$

Let us notice that in course of the construction of a pair of $2k$-gons $(k \geqq 3)$ by Lemma 4.3, $k-3$ distinct α-configurations (Fig. 1.1) and a chain of four quadrangles appear in the map. Within all transitions from a pair of m-gons to a pair of $(m+2)$-gons the quadrangles marked α_1 and α_5 (see Fig. 4.2b) are used, while the quadrangles marked $\alpha_2, \alpha_3, \alpha_4$ forming an α-configuration are retained in the map.

So the construction (taking place in pairs) of all the required faces with $\geqq 6$ vertices brings into the map $\frac{1}{2} \sum_{k \geqq 3} (k-3) p_{2k}$ α-configurations. The starting map contained two chains each containing four quadrangles, and these are not changed by the next construction. Each such chain contains one α-configuration. So we have together in the map $\frac{1}{2}\sum_{k\ge 3}(k-3)p_{2k}+2\ge 2g$ α-configurations. Now Lemma 1.2 is used to get the assertion of Theorem 4.1.

THEOREM 4.2. *Every sequence* $(p_i, i\ge 3)$ *of non-negative integers satisfying — besides the conclusion of Lemma 1.1 and* (4.1) *with* $g=1$ *— the conditions*

$$p_i \equiv 0 \pmod 2) \quad \textit{for all} \quad i$$

$$\sum_{i\ge 6} p_i \ne 0$$

is the face-vector of a toroidal (3, 4)*-map.*

PROOF. Lemma 3.2 ensures the existence of a 3-valent toroidal map M having one hexagon and two vertices (see Fig. 3.3a). From Lemma 4.2 then there follows the existence of a toroidal (3, 4)-map N having three quadrangles and two hexagons (one of them is associated to one of the vertices of the map M) and containing the face-aggregate in Fig. 4.2a with two hexagons ($m=6$). If for some $j\ge 4$, $p_{2j}\ne 0$ the two hexagons mentioned are used to get a pair of $2j$-gons. After this the construction proceeds as described in Lemma 4.3.

If $\sum_{j\ge 8} p_j=0$, first by Lemma 3.2 a toroidal 3-valent map with $\frac{p_6}{2}$ hexagons is obtained; after this use Lemma 4.2.

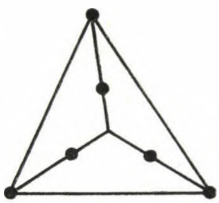


Fig. 4.3

THEOREM 4.3. *Every sequence* $(p_i, i\ge 3)$ *of non-negative integers satisfying — besides the conclusion of Lemma 1.1 and* (4.1) *with* $g=1$ *— the condition*

$$p_6 = \sum_{j\ge 4}(j-2)\left(p_{2j}-2\left[\frac{p_{2j}}{2}\right]\right)+2t, \qquad t=0, 1, \ldots,$$

is the face-vector of a toroidal (3, 4)*-map.*

PROOF. The case when $p_{2j}\equiv 0 \pmod 2$ for all $j\ge 4$ is settled in Theorem 4.2. Let for at least one j p_{2j} be odd. According to Lemma 3.2 there exists a toroidal

3-valent map M with face-vector $(q_i,\ i\geqq 3)$ such that

$$q_i = 0 \quad \text{for all odd} \quad i \geqq 3,$$

$$q_{2j} = p_{2j} - 2\left[\frac{p_{2j}}{2}\right] \quad \text{for all} \quad j \geqq 4,$$

$$q_6 = 0 \quad \text{and}$$

$$q_4 = \sum_{j\geqq 4}(j-3)q_{2j} = \sum_{j\geqq 4}(j-3)\left(p_{2j}-2\left[\frac{p_{2j}}{2}\right]\right) \quad \text{holds.}$$

By Lemma 4.2 there exists a toroidal (3, 4)-map N with face-vector $(r_i,\ i\geqq 3)$, for which

$$r_{2i+1} = q_{2i+1} = 0 \quad \text{for all} \quad i \geqq 1,$$

$$r_{2j} = q_{2j} = p_{2j} - 2\left[\frac{p_{2j}}{2}\right] \quad \text{for} \quad j \geqq 4 \quad \text{and}$$

$$r_6 = \frac{2}{3}q_4 + 2q_6 + \frac{1}{6}\sum_{j\geqq 4} 2jq_{2j} = \sum_{j\geqq 4}(j-2)\left(p_{2j}-2\left[\frac{p_{2j}}{2}\right]\right)$$

holds.

As the map M contains a quadrangle, the map N contains a quadruple of quadrangles as in Fig. 4.2a. Now Lemma 4.3 with map N can be applied to get the assertion of Theorem 4.3.

REFERENCES

[1] Acharya, B. D. and Vartak, M. N., On the construction of graphs with given constant valence-difference (S) on each of their lines, *Wiss. Z. Techn. Hochsch. Ilmenau* **23** (1977), no. 6, 33—60. *MR* **58** # 370.

[2] Grünbaum, B., *Convex polytopes,* Pure and Applied Mathematics, Vol. 16, Interscience Publishers [John Wiley & Sons, Inc.], New York, 1967. *MR* **37** # 2085.

[3] Harary, F., *Graph theory,* Addison-Wesley Publishing Co., Reading, Mass.—Menlo Park, Calif.—London, 1969. *MR* **41** # 1566.

[4] Rosenfeld, M., Polytopes of constant weight, *Israel J. Math.* **21** (1975), 24—30. *MR* **52** # 4132.

(Received October 16, 1979; in revised form September 2, 1981)

KATEDRA GEOMETRIE A ALGEBRY
PRÍRODOVEDECKÁ FAKULTA
UNIVERZITY P. J. ŠAFÁRIKA
KOMENSKÉHO UL. 14
ČS—041 54 KOŠICE
CZECHOSLOVAKIA

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 177—198

ÜBER DIE UNBEDINGTE KONVERGENZ DER MEHRFACHEN ORTHOGONALREIHEN

KÁROLY TANDORI

1. In der Arbeit [6] haben wir die unbedingte Konvergenz der Orthogonalreihen betrachtet. In dieser Note werden wir die entsprechenden Resultate über mehrfache Orthogonalreihen verallgemeinern. Die Sätze werden wir nur für doppelte Reihen verfassen und beweisen. Die Sätze können aber unmittelbar für beliebige $d(\geqq 2)$-fache Reihen ausgedehnt werden.

Es sei $(X, \mathcal{A}, \mu)$ ein endlicher Maßraum, $\{\varphi_{kl}(x)\}_{k,l=1}^{\infty}$ ein reelles orthonormiertes System in $(X, \mathcal{A}, \mu)$, und $\{a_{kl}\}_{k,l=1}^{\infty}$ eine Folge von reellen Zahlen.

Mit N_+^2 bezeichnen wir die Menge der geordneten Paare von positiven ganzen Zahlen:

$$N_+^2 = \{(k, l)\colon\ k, l = 1, 2, \ldots\}.$$

Es sei $(k, l) \to \big(i(k, l),\ j(k, l)\big)$ eine umkehrbar eindeutige Abbildung von N_+^2 auf sich selbst. Eine Anordnung der Orthogonalreihe

$$\sum_{k=1}^{\infty} \sum_{l=1}^{\infty} a_{kl}\varphi_{kl}(x) \tag{1}$$

ist die Reihe

$$\sum_{\alpha=1}^{\infty} \sum_{\beta=1}^{\infty} a_{k_\alpha, l_\beta}\varphi_{k_\alpha, l_\beta}, \tag{2}$$

wobei k_α, l_β diejenigen positiven ganzen Zahlen sind, für die $\alpha = i(k_\alpha, l_\beta)$, $\beta = j(k_\alpha, l_\beta)$ erfüllt sind. Die Anordnung nennen wir einfach, wenn

$$i(k, l) = i(k) \quad (l = 1, 2, \ldots;\ k = 1, 2, \ldots), \quad j(k, l) = j(l) \quad (k = 1, 2, \ldots;\ l = 1, 2, \ldots)$$

gelten.

Für die Indizes m, n setzen wir

$$s_{mn}(x) = \sum_{k=1}^{m} \sum_{l=1}^{n} a_{kl}\varphi_{kl}(x).$$

Wir sagen, daß die Reihe (1) in dem Punkt $x(\in X)$ konvergiert, wenn

$$\lim_{\substack{m\to\infty \\ n\to\infty}} s_{m,n}(x)$$

1980 *Mathematics Subject Classification.* Primary 42C20.
Key words and phrases. Unconditional convergence, multiple orthogonal series.

existiert. Für die Existenz dieses Grenzwertes ist notwendig und hinreichend, daß für jede Zahl $\varepsilon(>0)$ einen Index N derart existiert, daß im Falle $m, n, \bar{m}, \bar{n}>N$

$$|s_{mn}(x)-s_{\bar{m}\bar{n}}(x)|<\varepsilon$$

erfüllt ist.

Wir definieren noch eine Indexfolge $\{\nu(m)\}_{m=0}^{\infty}$. Es sei $\nu(0)=1$, und $\nu(m)=$ $=2^{2^{m-1}}$ $(m=1, 2, \ldots)$. Es sei

$$T(m, n)=\{(k, l)\in N_+^2:\ \nu(m)\leqq k<\nu(m+1),\ \nu(n)\leqq l<\nu(n+1)\}$$

$(m, n=0, 1, \ldots)$ und

$$\begin{aligned} T(m)=\{(k, l)\in N_+^2:\ &\nu(m)\leqq k<\nu(m+1),\ 1\leqq l<\nu(m+1),\quad \text{oder} \\ &1\leqq k<\nu(m+1),\ \nu(m)\leqq l<\nu(m+1)\}\qquad (m=0, 1, \ldots). \end{aligned} \tag{3}$$

2. Wir beweisen erstens die folgenden zwei Sätze.

SATZ 1. *Gilt*

$$\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}2^m 2^n\sqrt{\sum_{(k,l)\in T(m,n)}a_{kl}^2}<\infty, \tag{4}$$

dann konvergiert die Reihe (1) *bei jeder einfachen Anordnung in X fast überall.*

SATZ 2. *Gilt*

$$\sum_{m=0}^{\infty}2^{2m}\sqrt{\sum_{(k,l)\in T(m)}a_{kl}^2}<\infty, \tag{5}$$

dann konvergiert die Reihe (1) *bei jeder Anordnung in X fast überall.*

BEMERKUNG 1. Aus (5) folgt (4). Da, auf Grund von (3)

$$T(m)=\Big(\bigcup_{n=0}^{m}T(m, n)\Big)\cup\Big(\bigcup_{k=0}^{m-1}T(k, m)\Big)$$

ist, folgt

$$\sum_{n=0}^{m}2^m 2^n\sqrt{\sum_{(k,l)\in T(m,n)}a_{kl}^2}+\sum_{\mu=0}^{m-1}2^{\mu}2^m\sqrt{\sum_{(k,l)\in T(\mu,m)}a_{kl}^2}=$$

$$=2^{2m}\sum_{n=0}^{m}\frac{1}{2^{m-n}}\sqrt{\sum_{(k,l)\in T(m,n)}a_{kl}^2}+2^{2m}\sum_{\mu=0}^{m-1}\frac{1}{2^{m-\mu}}\sqrt{\sum_{(k,l)\in T(\mu,m)}a_{kl}^2}\leqq$$

$$\leqq\sqrt{2}\cdot 2^{2m}\left(\sqrt{\sum_{(k,l)\in\bigcup_{n=0}^{m}T(m,n)}a_{kl}^2}+\sqrt{\sum_{(k,l)\in\bigcup_{\mu=0}^{m-1}T(\mu,m)}a_{kl}^2}\right)\leqq$$

$$\leqq 2\cdot 2^{2m}\sqrt{\sum_{(k,l)\in T(m)}a_{kl}^2}\qquad (m=0, 1, \ldots),$$

woraus die Implikation (5)⇒(4) sich ergibt.

Zum Beweis dieser Sätze benötigen wir die folgenden Hilfssätze.

HILFSSATZ 1. *Es seien M, N positive ganze Zahlen, $\{b_{kl}\}$ $(k=1, \ldots, M, l=1, \ldots, N)$ eine reelle Zahlenfolge, und $\{\psi_{kl}(x)\}$ $(k=1, \ldots, M, l=1, \ldots, N)$ ein reelles ortho-*

normiertes System im endlichen Maßraum $(X, \mathscr{A}, \mu)$. Dann gilt

$$\int_X \max_{\substack{1\leq m\leq M\\ 1\leq n\leq N}} \left(\sum_{k=1}^{m}\sum_{l=1}^{n} b_{kl}\psi_{kl}(x)\right)^2 d\mu \leq C_1 \log^2 2M\cdot \log^2 2N \sum_{k=1}^{M}\sum_{l=1}^{N} b_{kl}^2$$

mit einer positiven Konstante C_1.

(Im folgenden bezeichnen $C_2, C_3, \ldots$ positive Konstanten.)
Dieser Hilfssatz ist bekannt (s. z. B. [3]).

HILFSSATZ 2. *Es seien M_1, N_1 positive ganze Zahlen, $\{c_{ij}\}$ $(i=1, \ldots, M_1, j=1, \ldots, N_1\}$ eine reelle Zahlenfolge, und $\{\chi_{ij}(x)\}$ $(i=1, \ldots, M_1, j=1, \ldots, N_1\}$ ein reelles orthonormiertes System im endlichen Maßraum $(X, \mathscr{A}, \mu)$. Ist die Anzahl der von Null verschiedenen Koeffizienten c_{ij} gleich mit $L(\geq 1)$, dann gilt*

$$\int_X \max_{\substack{1\leq m\leq M_1\\ 1\leq n\leq N_1}} \left(\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}\chi_{ij}(x)\right)^2 d\mu \leq C_1 \log^4 2L \sum_{i=1}^{M_1}\sum_{j=1}^{N_1} c_{ij}^2 .$$

BEWEIS des Hilfssatzes 2. Es seien $(1\leq)i_1<\ldots<i_M(\leq M_1)$ diejenigen Indizes i, für die

$$\sum_{j=1}^{N_1} b_{ij}^2 \neq 0$$

ist, weiterhin seien $(1\leq)j_1<\ldots<j_N(\leq N_1)$ diejenigen Indizes j, für die

$$\sum_{i=1}^{M_1} b_{ij}^2 \neq 0$$

ist. Offensichtlich gelten

$$M \leq L, \quad N \leq L. \tag{6}$$

Wir setzen

$$b_{kl} = c_{i_k, j_l}, \quad \psi_{kl}(x) = \chi_{i_k, j_l}(x) \qquad (k = 1, \ldots, M, l = 1, \ldots, N). \tag{7}$$

Auf Grund des Hilfssatzes 1 gilt

$$\int_X \max_{\substack{1\leq m\leq M\\ 1\leq n\leq N}} \left(\sum_{k=1}^{m}\sum_{l=1}^{n} b_{kl}\psi_{kl}(x)\right)^2 d\mu \leq C_1 \log^2 2M \log^2 2N \sum_{k=1}^{M}\sum_{l=1}^{N} b_{kl}^2. \tag{8}$$

Einerseits, auf Grund von (7) gilt

$$\max_{\substack{1\leq m\leq M_1\\ 1\leq n\leq N_1}} \left(\sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}\chi_{ij}(x)\right)^2 = \max_{\substack{1\leq m\leq M\\ 1\leq n\leq N}} \left(\sum_{k=1}^{m}\sum_{l=1}^{n} b_{kl}\psi_{kl}(x)\right)^2 \qquad (x\in X). \tag{9}$$

Andererseits, aus (6) und (7) folgt

$$\log^2 2M \log^2 2N \sum_{k=1}^{M}\sum_{l=1}^{N} b_{kl}^2 \leq \log^4 2L \sum_{i=1}^{M_1}\sum_{j=1}^{N_1} c_{ij}^2 . \tag{10}$$

Aus (8), (9) und (10) erhalten wir die Behauptung des Hilfssatzes 2.

BEWEIS des Satzes 1. Es sei (2) eine beliebige einfache Anordnung der Reihe (1). Weiterhin seien $m, n(\geqq 0)$ ganze Zahlen. Es sei $i(m,n;1)<\ldots<i(m,n;\nu(m+1)-\nu(m))$ die monoton wachsende Anordnung der Indizes $i(k)$ $(k=\nu(m),\ldots,\nu(m+1)-1)$, und $j(m,n;1)<\ldots<j(m,n;\nu(n+1)-\nu(n))$ die monoton wachsende Anordnung der Indizes $j(l)$ $(l=\nu(n),\ldots,\nu(n+1)-1)$. Wir setzen

$$\delta_{mn}^2(x)=\max_{t,u}\Big(\sum_{\substack{(k,l)\in T(m,n)\\ i(k)\leqq t,\, j(l)\leqq u}} a_{k,l}\varphi_{k,l}(x)\Big)^2=$$

$$=\max_{\substack{1\leqq p\leqq \nu(m+1)-\nu(m)\\ 1\leqq q\leqq \nu(n+1)-\nu(n)}}\Big(\sum_{r=1}^{p}\sum_{s=1}^{q} a_{k_{i(m,n,r)},l_{j(m,n,s)}}\varphi_{k_{i(m,n,r)}l_{j(m,n,s)}}(x)\Big)^2.$$

Dann gilt

$$\int_X \delta_{mn}^2(x)\,d\mu\leqq C_2\log^2(\nu(m+1)-\nu(m))\log^2(\nu(n+1)-\nu(n))\times$$

(11)

$$\times\sum_{r=1}^{\nu(m+1)-\nu(m)}\sum_{s=1}^{\nu(n+1)-\nu(n)} a^2_{k_{i(m,n,r)},l_{j(m,n,s)}}\leqq C_2 2^{2m}2^{2n}\sum_{(k,l)\in T(m,n)} a_{kl}^2,$$

auf Grund des Hilfssatzes 1. Aus (11) erhalten wir

$$\int_X \delta_{mn}(x)\,d\mu\leqq\sqrt{\mu(X)}\sqrt{\int_X\delta_{mn}^2(x)\,d\mu}\leqq\sqrt{C_2\mu(X)}\,2^m2^n\sqrt{\sum_{(k,l)\in T(m,n)} a_{kl}^2}.$$

Daraus und aus (4) ergibt sich

$$\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\int_X\delta_{mn}(x)\,d\mu<\infty,$$

und so folgt, daß die Reihe

(12) $$\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\delta_{mn}(x)$$

in X fast überall konvergiert.

Es sei $\varepsilon>0$ beliebig, und $x(\in X)$ ein Punkt, in welchem die Reihe (12) konvergiert. Dann gibt es einen Index N mit

(13) $$\sum_{m=N}^{\infty}\sum_{n=N}^{\infty}\delta_{mn}(x)<\varepsilon/2.$$

Es sei

$$\bar{s}_{\varkappa\lambda}(x)=\sum_{\alpha=1}^{\varkappa}\sum_{\beta=1}^{\lambda} a_{k_\alpha,l_\beta}\varphi_{k_\alpha,l_\beta}(x)=\sum_{\substack{(k,l)\in N_+^2\\ i(k)\leqq\varkappa,\, j(l)\leqq\lambda}} a_{k,l}\varphi_{k,l}(x)$$

die Partialsumme der Reihe (2). Dann gilt

$$\bar{s}_{\varkappa\lambda}(x)=\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\Big(\sum_{\substack{(k,l)\in T(m,n)\\ i(k)\leqq\varkappa,\, j(l)\leqq\lambda}} a_{k,l}\varphi_{k,l}(x)\Big).$$

Es seien $\varkappa, \lambda, \bar{\varkappa}, \bar{\lambda}$ $(>A = \max\{i(k), j(l)\colon (k,l)\in T(m,n), m, n<N\})$ beliebige Indizes. Dann gilt nach den obigen

$$\bar{s}_{\varkappa\lambda}(x)-\bar{s}_{\bar{\varkappa}\bar{\lambda}}(x) = \sum_{m=N}^{\infty}\sum_{n=N}^{\infty}\Big(\sum_{\substack{(k,l)\in T(m,n)\\ i(k)\le\varkappa,\, j(l)\le\lambda}} a_{k,l}\varphi_{k,l}(x) - \sum_{\substack{(k,l)\in T(m,n)\\ i(k)\le\bar{\varkappa},\, j(l)\le\bar{\lambda}}} a_{k,l}\varphi_{k,l}(x)\Big).$$

Daraus, nach der Definition von $\delta_{mn}(x)$, auf Grund von (13) ergibt sich

$$|\bar{s}_{\varkappa\lambda}(x)-\bar{s}_{\bar{\varkappa}\bar{\lambda}}(x)| \le 2\sum_{m=N}^{\infty}\sum_{n=N}^{\infty}\delta_{mn}(x) < \varepsilon,$$

für $\varkappa, \lambda, \bar{\varkappa}, \bar{\lambda} > A$. Daraus folgt, daß $\lim\limits_{\substack{\varkappa\to\infty\\ \lambda\to\infty}} \bar{s}_{\varkappa\lambda}(x)$ existiert.

Damit haben wir Satz 1 bewiesen.

Beweis des Satzes 2. Es sei (2) eine beliebige Anordnung der Reihe (1), und $m(\ge 0)$ eine ganze Zahl. Es seien weiterhin $M(m) = \max\{i(k,l)\colon (k,l)\in T(m)\}$, $N(m) = \max\{j(k,l)\colon (k,l)\in T(m)\}$, und $Z(m)$ die Anzahl der Elemente von $T(m)$. Wir setzen

$$c_{ij} = \begin{cases} a_{k,l}, & \text{für } i = i(k,l),\ \ j = j(k,l),\ \ (k,l)\in T(m),\\ 0, & \text{für } (i-i(k,l))^2+(j-j(k,l))^2 \ne 0,\ \ (k,l)\in T(m),\\ & 1\le i\le M(m),\ \ 1\le j\le N(m), \end{cases}$$

und

$$\chi_{ij}(x) = \varphi_{k,l}(x), \quad \text{für } i = i(k,l),\ \ j = j(k,l),\ \ (k,l)\in N_+^2,$$
$$1\le i\le M(m),\ \ 1\le j\le N(m).$$

Dann können wir den Hilfssatz 2 mit $L = Z(m) \le (2^{2^m})^2$, $M_1 = M(m)$, $N_1 = N(m)$ anwenden. So für die Funktion

$$\delta_m^2(x) = \max_{\substack{1\le p\le M(m)\\ 1\le q\le N(m)}}\Big(\sum_{\substack{(k,l)\in T(m)\\ i(k,l)\le p,\, j(k,l)\le q}} a_{k,l}\varphi_{k,l}(x)\Big)^2$$

folgt die Abschätzung

$$\begin{aligned}\int_X \delta_m^2(x)\,d\mu &= \int_X \max_{\substack{1\le p\le M(m)\\ 1\le q\le N(m)}}\Big(\sum_{i=1}^{p}\sum_{j=1}^{q} c_{ij}\chi_{ij}(x)\Big)^2 d\mu \le \\ &\le C_3 2^{4m}\sum_{i=1}^{M(m)}\sum_{j=1}^{N(m)} c_{ij}^2 = C_3 2^{4m}\sum_{(k,l)\in T(m)} a_{kl}^2.\end{aligned} \tag{14}$$

Aus (5) und (14) bekommen wir

$$\begin{aligned}\sum_{m=0}^{\infty}\int_X \delta_m(x)\,d\mu &\le \sqrt{\mu(X)}\sum_{m=0}^{\infty}\sqrt{\int_X \delta_m^2(x)\,d\mu} \le \\ &\le \sqrt{C_3\mu(X)}\sum_{m=0}^{\infty} 2^{2m}\sqrt{\sum_{(k,l)\in T(m)} a_{kl}^2} < \infty,\end{aligned}$$

und so folgt

$$\sum_{m=0}^{\infty} \delta_m(x) < \infty \tag{15}$$

in X fast überall.

Es sei $\varepsilon > 0$ beliebig, und $x (\in X)$ ein Punkt, in welchem die Reihe (15) konvergiert. Dann gibt es einen Index N mit

$$\sum_{m=N}^{\infty} \delta_m(x) < \varepsilon/2. \tag{16}$$

Wir setzen

$$\tilde{s}_{\varkappa\lambda}(x) = \sum_{\alpha=1}^{\varkappa} \sum_{\beta=1}^{\lambda} a_{k_\alpha, l_\beta} \varphi_{k_\alpha, l_\beta}(x) = \sum_{\substack{(k,l)\in N_+^2 \\ i(k,l)\leq\varkappa,\ j(k,l)\leq\lambda}} a_{k,l}\varphi_{k,l}(x).$$

Für diese Partialsumme gilt

$$\tilde{s}_{\varkappa\lambda}(x) = \sum_{m=0}^{\infty} \Big(\sum_{\substack{(k,l)\in T(m) \\ i(k,l)\leq\varkappa,\ j(k,l)\leq\lambda}} a_{k,l}\varphi_{k,l}(x) \Big).$$

Es seien $\varkappa, \lambda, \bar{\varkappa}, \bar{\lambda} (\geq B = \max \{i(k, l), j(k, l) : (k, l) \in T(m), m < N\})$. Dann ist

$$\tilde{s}_{\varkappa\lambda}(x) - \tilde{s}_{\bar{\varkappa}\bar{\lambda}}(x) = \sum_{m=N}^{\infty} \Big(\sum_{\substack{(k,l)\in T(m) \\ i(k,l)\leq\varkappa,\ j(k,l)\leq\lambda}} a_{k,l}\varphi_{k,l}(x) - \sum_{\substack{(k,l)\in T(m) \\ i(k,l)\leq\bar{\varkappa},\ j(k,l)\leq\bar{\lambda}}} a_{k,l}\varphi_{k,l}(x) \Big),$$

und so folgt auf Grund der Definition von $\delta_m(x)$ und der Abschätzung (16):

$$|\tilde{s}_{\varkappa\lambda}(x) - \tilde{s}_{\bar{\varkappa}\bar{\lambda}}(x)| \leq 2 \sum_{m=N}^{\infty} \delta_m(x) < \varepsilon,$$

für $\varkappa, \lambda, \bar{\varkappa}, \bar{\lambda} > B$. Daraus folgt, daß $\lim\limits_{\substack{\varkappa\to\infty \\ \lambda\to\infty}} \tilde{s}_{\varkappa\lambda}(x)$ existiert.

Damit haben wir Satz 2 bewiesen.

Bemerkung 2. Es ist klar, daß man diese Sätze für σ-endliche Maßräume auch beweisen kann.

Bemerkung 3. Den Satz 2 kann man offensichtlich im folgenden Sinne verschärfen.

Satz 2a. *Es sei*

$$\sum_{k=1}^{\infty} \sum_{l=1}^{\infty} a_{kl}^2 < \infty,$$

und es bezeichne $\{a_{k_\alpha, l_\beta}\}_{\alpha, \beta=1}^{\infty}$ *eine Anordnung der Folge* $\{a_{kl}\}_{k,l=1}^{\infty}$, *für die*

$$\min_{(\alpha,\beta)\in T(m)} |a_{k_\alpha, l_\beta}| \geq \max_{(\alpha,\beta)\in T(m+1)} |a_{k_\alpha, l_\beta}| \qquad (m = 0, 1, \ldots)$$

erfüllt ist. Gilt

$$\sum_{m=0}^{\infty} 2^{2m} \sqrt{\sum_{(\alpha,\beta)\in T(m)} a_{k_\alpha, l_\beta}^2} < \infty,$$

dann konvergiert die Reihe (1) *bei jeder Anordnung in X fast überall.*

3. Aus den Sätzen 1—2 kann man die folgenden Sätze erhalten.

SATZ 3. *Es sei* $\{\lambda_k\}_{k=1}^{\infty}$ *eine monoton nichtabnehmende Folge von positiven Zahlen mit*

(17) $$\sum_{m=0}^{\infty} \lambda_{\nu(m)}^{-1} < \infty.$$

Gilt

(18) $$\sum_{k=1}^{\infty} \sum_{l=1}^{\infty} a_{kl}^2 \log^2 2k \log^2 2l \, \lambda_k \lambda_l < \infty,$$

dann konvergiert die Reihe (1) *bei jeder einfacher Anordnung in X fast überall.*

Dieser Satz ist bekannt (s. [5]).

SATZ 4. *Es sei* $\{\lambda_k\}_{k=1}^{\infty}$ *eine monotone nichtabnehmende Folge von positiven Zahlen mit* (17). *Gilt*

(19) $$\sum_{k=1}^{\infty} \sum_{l=1}^{\infty} a_{kl}^2 \max(\log^4 2k, \log^4 2l) \max(\lambda_k, \lambda_l) < \infty,$$

dann konvergiert die Reihe (1) *bei jeder Anordnung in X fast überall.*

Diese Sätze sind die Analoge des Satzes von W. Orlicz [4].

BEWEIS des Satzes 3. Aus (17) und (18) folgt

$$\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} 2^m 2^n \sqrt{\sum_{(k,l)\in T(m)} a_{kl}^2} \leqq$$

$$\leqq C_4 \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \sqrt{\frac{1}{\lambda_{\nu(m)} \lambda_{\nu(n)}} \sum_{(k,l)\in T(m,n)} a_{kl}^2 \log^2 2k \log^2 2l \, \lambda_k \lambda_l} \leqq$$

$$\leqq C_4 \sqrt{\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{1}{\lambda_{\nu(m)} \lambda_{\nu(n)}}} \sqrt{\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \sum_{(k,l)\in T(m,n)} a_{kl}^2 \log^2 2k \log^2 2l \, \lambda_k \lambda_l} \leqq$$

$$\leqq C_4 \left(\sum_{m=0}^{\infty} \frac{1}{\lambda_{\nu(m)}} \right) \sqrt{\sum_{k=1}^{\infty} \sum_{l=1}^{\infty} a_{kl}^2 \log^2 2k \log^2 2l \, \lambda_k \lambda_l} < \infty,$$

und so ergibt sich die Behauptung auf Grund des Satzes 1.

Beweis des Satzes 4. Aus (19) folgt

$$\sum_{m=0}^{\infty} 2^{2m} \sqrt{\sum_{(k,l)\in T(m)} a_{kl}^2} \leqq$$

$$\leqq C_5 \sum_{m=0}^{\infty} \sqrt{\frac{1}{\lambda_{\nu(m)}} \sum_{(k,l)\in T(m)} a_{kl}^2 \max(\log^4 2k, \log^4 2l) \max(\lambda_k, \lambda_l)} \leqq$$

$$\leqq C_5 \sqrt{\sum_{m=0}^{\infty} \frac{1}{\lambda_{\nu(m)}}} \sqrt{\sum_{m=0}^{\infty} \sum_{(k,l)\in T(m)} a_{kl}^2 \max(\log^4 2k, \log^4 2l) \max(\lambda_k, \lambda_l)} =$$

$$= C_5 \sqrt{\sum_{m=0}^{\infty} \frac{1}{\lambda_{\nu(m)}}} \sqrt{\sum_{k=1}^{\infty} \sum_{l=1}^{\infty} a_{kl}^2 \max(\log^4 2k, \log^4 2l) \max(\lambda_k, \lambda_l)} < \infty,$$

und so bekommen wir die Behauptung auf Grund des Satzes 2.

4. Man kann zeigen, daß die Bedingung (18) im allgemeinen unverbesserbar ist.

Satz 5. *Ist* $\{\lambda_k\}_{k=1}^{\infty}$ *eine nichtabnehmende Folge von positiven Zahlen mit*

(20) $$\sum_{m=0}^{\infty} \frac{1}{\lambda_{\nu(m)}} = \infty,$$

dann gibt es ein endlicher Maßraum $(X, \mathcal{A}, \mu)$, *ein reelles orthonormiertes System* $\{\varphi_{kl}(x)\}_{k,l=1}^{\infty}$ *in diesem Raum und eine reelle Zahlenfolge* $\{a_{kl}\}_{k,l=1}^{\infty}$ *derart, daß*

$$\sum_{k=1}^{\infty} \sum_{l=1}^{\infty} a_{kl}^2 \log^2 2k \log^2 2l\, \lambda_k \lambda_l < \infty$$

gilt, und die Reihe (1) *in gewisser einfacher Anordnung in X fast überall divergiert.*

Beweis des Satzes 5. In der Arbeit [6] haben wir den folgenden Satz bewiesen.

Satz A. *Es sei* $\{a_k\}_{k=1}^{\infty}$ *eine monoton abnehmende Folge von positiven Zahlen. Gilt*

$$\sum_{m=0}^{\infty} 2^m \sqrt{\sum_{k=\nu(m)}^{\nu(m+1)-1} a_k^2} = \infty,$$

dann gibt es ein orthonormiertes System $\{\varphi_k(x)\}_{k=1}^{\infty}$ *im Intervall* (0, 1) *derart, daß die Reihe*

(21) $$\sum_{k=1}^{\infty} a_k \varphi_k(x)$$

in gewisser Anordnung

(22) $$\sum_{\alpha=1}^{\infty} a_{k_\alpha} \varphi_{k_\alpha}(x)$$

in (0, 1) *fast überall divergiert.*

Hier ist $k \to i(k)$ eine umkehrbar eindeutige Abbildung der Menge $\{1, 2, \dots\}$ auf sich selbst, und k_α bezeichnet diejenigen Index, für die $\alpha = i(k_\alpha)$ $(\alpha = 1, 2, \dots)$ gilt.

Durch einfacher Rechnung erhalten wir

$$\sum_{m=0}^{\infty} \frac{1}{\lambda_{\nu(m)} \left(\sum_{\mu=0}^{m} \lambda_{\nu(\mu)}^{-1} \right)} = \infty,$$

und

$$\sum_{m=0}^{\infty} \frac{1}{\lambda_{\nu(m)} \left(\sum_{\mu=0}^{m} \lambda_{\nu(\mu)}^{-1} \right)^2} < \infty$$

auf Grund von (20).

Es sei

$$a_k = \frac{1}{\sqrt{\nu(m+1)-\nu(m)}\, 2^m \lambda_{\nu(m+1)} \left(\sum_{\mu=0}^{m+1} \lambda_{\nu(\mu)}^{-1} \right)}$$

$(\nu(m) \leq k < \nu(m+1),\ m=0, 1, \ldots)$.

Dann gilt $a_k \geqq a_{k+1}$ $(k=1, 2, \ldots)$ und

$$\sum_{m=0}^{\infty} 2^m \sqrt{\sum_{k=\nu(m)}^{\nu(m+1)-1} a_k^2} = \sum_{m=0}^{\infty} \frac{1}{\lambda_{\nu(m+1)} \left(\sum_{\mu=0}^{m+1} \lambda_{\nu(\mu)}^{-1} \right)} = \infty.$$

Auf Grund des Satzes A gibt es ein orthonormiertes System $\{\varphi_k(x)\}_{k=1}^{\infty}$ im Intervall (0, 1) derart, daß die Reihe (21) in einer Anordnung (22) in (0, 1) fast überall divergiert.

Es sei $(X, \mathscr{A}, \mu)$ das Einheitsquadrat $E=(0, 1)\times(0, 1)$ mit dem gewöhnlichen Lebesgueschen Maß. Wir setzen

$$a_{kl} = a_k a_l, \quad \varphi_{kl}(x, y) = \varphi_k(x)\varphi_l(y) \quad (k, l = 1, 2, \ldots; (x, y) \in E).$$

Dann gilt

$$\sum_{k=1}^{\infty} \sum_{l=1}^{\infty} a_{kl}^2 \log^2 2k \log^2 2l \lambda_k \lambda_l \leqq$$

$$\leqq 4 \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{1}{\lambda_{\nu(m+1)} \left(\sum_{\mu=0}^{m+1} \lambda_{\nu(\mu)}^{-1} \right)^2 \lambda_{\nu(n+1)} \left(\sum_{\mu=0}^{n+1} \lambda_{\nu(\mu)}^{-1} \right)^2} < \infty,$$

weiterhin, nach obigen, divergiert die einfache Anordnung

$$\sum_{\alpha=1}^{\infty} \sum_{\beta=1}^{\infty} a_{k_\alpha, l_\beta} \varphi_{k_\alpha, l_\beta}(x, y)$$

der Reihe

$$\sum_{k=1}^{\infty} \sum_{l=1}^{\infty} a_{kl} \varphi_{kl}(x, y)$$

in E fast überall.

5. Endlich zeigen wir, daß Satz 4 auch unverbesserbar ist. Wir beweisen nähmlich den folgenden Satz.

SATZ 6. *Es sei* $\{\lambda_k\}_{k=1}^{\infty}$ *eine monoton nichtabnehmende Folge von positiven Zahlen mit* (20). *Dann gibt es ein endlicher Maßraum* $(X, \mathcal{A}, \mu)$, *ein reelles orthonormiertes System* $\{\varphi_{kl}(x)\}_{k,l=1}^{\infty}$ *in diesem Raum und eine reelle Zahlenfolge* $\{a_{kl}\}_{k,l=1}^{\infty}$ *derart, daß* (19) *gilt, und die Reihe* (1) *in gewisser Anordnung in X fast überall divergiert.*

Um diesen Satz zu beweisen, sollen wir weitere Hilfssätze vorausschicken.

HILFSSATZ 3. *Für beliebige positive ganze Zahl p und für beliebige disjunkte Intervalle* $I_1=(a_1, b_1)(\subseteqq(0, 1))$, $I_2, I_3(\subseteqq(1, 2))$ *gibt es orthonormiertes System der Treppenfunktionen* $\Phi_l(I_1, I_2, I_3; p; x)$ $(l=1, \ldots, 8p)$ *im Intervall* $(0, 2)$ *derart, daß*

$$\Phi_l(I_1, I_2, I_3; p; x) = 0 \quad (x \notin I_1 \cup I_2 \cup I_3), \tag{23}$$

$$\int_{I_i} \Phi_l(I_1, I_2, I_3; p; x)\, dx = 0 \quad (i = 1, 2, 3) \tag{24}$$

$(l=1, \ldots, 8p)$ *sind, weiterhin, in jedem Intervalle*

$$J_k = \left(a_1 + \frac{b_1-a_1}{8p}(k-1),\ a_1 + \frac{b_1-a_1}{8p}k\right)$$

sind die Funktionen $\Phi_l(I_1, I_2, I_3; p; x)$ *konstant, und für jedes Intervall* J_k *gibt es Indizes* l_k, l'_k $(1 \leqq l_k < 4p < l'_k \leqq 8p)$ *mit*

$$\sum_{l=l_k}^{4p} \Phi_l(I_1, I_2, I_3; p; x) \geqq C_6 \frac{\sqrt{p}\log 2p}{\sqrt{\operatorname{mes} I_1}} \quad (x \in J_k;\ k = 4p+1, \ldots, 8p), \tag{25}$$

$$\sum_{l=4p+1}^{l'_k} \Phi_l(I_1, I_2, I_3; p; x) \geqq C_6 \frac{\sqrt{p}\log 2p}{\sqrt{\operatorname{mes} I_1}} \quad (x \in J_k;\ k = 1, \ldots, 4p). \tag{26}$$

BEWEIS des Hilfssatzes 3. Wir betrachten das Menchoff—Kaczmarzsche Funktionensystem (s. [1], [2]). Für eine positive ganze Zahl p sei

$$\bar{f}_l(p; x) = \frac{1}{k-p-l-1/2}, \quad \text{für} \quad x \in \left(\frac{k-1}{p}, \frac{k}{p}\right) \quad (k = 1, \ldots, 8p,\ l = 1, \ldots, 4p).$$

Dann ist

$$\int_0^8 \bar{f}_l^2(p; x)\, dx = \sum_{k=1}^{8p} \frac{1}{(k-p-l-1/2)^2}\frac{1}{p} \leqq \frac{C_7}{p} \quad (l = 1, \ldots, 4p). \tag{27}$$

Für $i>j$ ergibt sich

$$\alpha_{ij}=\int_0^8 \bar f_i(p;x)\bar f_j(p;x)\,dx=\frac{1}{p}\sum_{k=1}^{8p}\frac{1}{(k-p-i-1/2)(k-p-j-1/2)}=$$

$$=\frac{1}{p}\,\frac{1}{i-j}\sum_{k=1}^{8p}\left(\frac{1}{k-p-i-1/2}-\frac{1}{k-p-j-1/2}\right)=$$

$$=\frac{1}{p}\,\frac{1}{i-j}\left(\sum_{k=1-p-i}^{7p-i}\frac{1}{k-1/2}-\sum_{k=1-p-j}^{7p-j}\frac{1}{k-1/2}\right)=$$

$$=\frac{1}{p}\,\frac{1}{i-j}\left(\sum_{k=1-p-i}^{-p-j}\frac{1}{k-1/2}-\sum_{k=7p-i+1}^{7p-j}\frac{1}{k-1/2}\right),$$

und so gilt

$$|\alpha_{ij}|\leqq\frac{1}{p}\,\frac{1}{i-j}\left(\frac{i-j}{p+j+1/2}+\frac{i-j}{7p-i+1/2}\right)\leqq\frac{C_8}{p^2}. \tag{28}$$

Wir teilen das Intervall (8, 10) in $N=4p(4p-1)$ Teilintervalle gleicher Länge: I_{ij} $(1\leqq i,j\leqq N;\ i\neq j)$. Es sei

$$\bar f_l(p;x)=\begin{cases}\sqrt{\frac{1}{2}N|\alpha_{lj}|} & (x\in I_{lj};\ j=1,\dots,4p,\ j\neq l),\\ -\sqrt{\frac{1}{2}N|\alpha_{lj}|}\,\operatorname{sign}\alpha_{lj} & (x\in I_{jl},\ j=1,\dots,4p,\ j\neq l),\\ 0 & \text{sonst}\end{cases}$$

$(l=1,\dots,4p)$. Es ist klar, daß die Treppenfunktionen $\bar f_l(p;x)$ $(l=1,\dots,4p)$ bilden im Intervall (0, 10) ein orthogonales System, aus (27) und (28) folgt

$$\int_0^{10}\bar f_l^2(p;x)\,dx=\int_0^8\bar f_l^2(p;x)\,dx+\frac{1}{2}\sum_{j=1}^{l-1}|\alpha_{lj}|+\frac{1}{2}\sum_{j=l+1}^{4p}|\alpha_{lj}|\leqq\frac{C_9}{p} \tag{29}$$

$(l=1,\dots,4p)$, weiterhin, auf Grund der Definition der Funktionen $\bar f_l(p;x)$ bestehen

$$\bar f_l(p;x)>0\qquad(l=1,\dots,p+k),$$

$$\sum_{l=1}^{p+k}\bar f_l(p;x)=\sum_{l=1}^{p+k}\frac{1}{2p+k+1-p-l-1/2}\geqq C_{10}\log 2p,$$

$$\bar f_l(p;x)<0\qquad(l=p+k+1,\dots,4p), \tag{30}$$

$$\sum_{l=p+k+1}^{4p}\bar f_l(p;x)=\sum_{l=p+k+1}^{4p}\frac{1}{2p+k+1-p-l-1/2}\leqq-C_{10}\log 2p$$

$$\left(x\in\left(\frac{2p+k}{p},\ \frac{2p+k+1}{p}\right);\ k=0,\dots,p-1\right).$$

Es sei

$$\bar{f}_l(p;x) = \bar{f}_l(p;x)\Big(\int\limits_0^{10} \bar{f}_l^2(p;x)\,dx\Big)^{-1/2} \qquad (l = 1, \dots, 4p).$$

Diese Treppenfunktionen bilden in (0, 10) ein orthonormiertes System, weiterhin aus (29) und (30) folgt, daß für jeden Index $k=0, \dots, p-1$ einen von k abhängigen Index $L_k (p \leqq L_k \leqq 2p)$ derart existiert, daß

$$(31)\qquad \sum_{l=1}^{L_k} \bar{f}_l(p;x) \geqq C_{11}\sqrt{p}\log 2p, \qquad \sum_{l=L_k+1}^{4p} \bar{f}_l(p;x) \leqq -C_{11}\sqrt{p}\log 2p$$
$$\Big(x\in\Big(\frac{2p+k}{p}, \frac{2p+k+1}{p}\Big);\ k = 0, \dots, p-1\Big)$$

erfüllt sind.

Wir setzen

$$\bar{g}_l(p;x) = \begin{cases} \frac{1}{\sqrt{2}}\bar{f}_l(p;x+2) & (x\in(0,1)), \\ -\frac{1}{\sqrt{2}}\bar{f}_l(p;x+1) & (x\in(1,2)), \\ 0 & \text{sonst} \end{cases}$$

und

$$\bar{h}_l(p;x) = \begin{cases} \frac{1}{\sqrt{2}}\bar{f}_l(p;x) & (x\in(0,2)), \\ -\frac{1}{\sqrt{2}}\bar{f}_l(p;x-2) & (x\in(2,4)), \\ \frac{1}{\sqrt{2}}\bar{f}_l(p;x-1) & (x\in(4,11)), \\ -\frac{1}{\sqrt{2}}\bar{f}_l(p;x-8) & (x\in(11,18)), \\ 0 & \text{sonst} \end{cases}$$

$(l=1, \dots, 4p)$. Offensichtlich gelten

$$\int\limits_0^2 \bar{g}_l(p;x)\,dx = \int\limits_0^{18} \bar{h}_l(p;x)\,dx = 0 \qquad (l = 1, \dots, 4p),$$

$$\int\limits_0^2 \bar{g}_k(p;x)\bar{g}_l(p;x)\,dx + \int\limits_0^{18} \bar{h}_k(p;x)\bar{h}_l(p;x)\,dx = \int\limits_0^{10} \bar{f}_k(p;x)\bar{f}_l(p;x)\,dx = \delta_{kl}$$
$$(k, l = 1, \dots, 4p),$$

wobei $\delta_{kl}=1$ $(k=l)$ und $\delta_{kl}=0$ $(k\neq l)$ ist. Dann seien

$$g_l(p;x) = \sqrt{2}\,\bar{g}_l(p;2x), \quad h_l(p;x) = \sqrt{18}\,\bar{h}_l(p;18x) \qquad (l = 1, \dots, 4p).$$

Nach obigen gelten

$$(32)\qquad \int_0^1 g_l(p;x)dx = \int_0^1 h_l(p;x)\,dx = 0 \qquad (l = 1, \dots, 4p),$$

$$(33)\qquad \int_0^1 g_k(p;x)g_l(p;x)\,dx + \int_0^1 h_k(p;x)h_l(p;x)\,dx =$$

$$= \int_0^2 \bar{g}_k(p;x)\bar{g}_l(p;x)dx + \int_0^{18} \bar{h}_k(p;x)\bar{h}_l(p;x)dx = \delta_{kl} \qquad (k, l = 1, \dots, 4p).$$

Weiterhin, auf Grund von (31) folgt, daß es Indizes l_k, l_k' $(1 \leqq l_k, l_k' \leqq 4p)$ mit

$$(34)\qquad \begin{aligned} &\sum_{k=1}^{l_k} g_l(p;x) \geqq C_{12}\sqrt{p}\log 2p \qquad \left\{x\in\left(\frac{k-1}{8p}, \frac{k}{8p}\right);\ k = 1, \dots, 4p\right\},\\ &\sum_{k=l_k'}^{4p} g_l(p;x) \geqq C_{12}\sqrt{p}\log 2p \qquad \left\{x\in\left(\frac{k-1}{8p}, \frac{k}{8p}\right);\ k = 4p+1, \dots, 8p\right\}\end{aligned}$$

existieren.

Wir teilen das Intervall I_3 in $8p$ disjunkte Teilintervalle gleicher Länge: $\bar{I}_i$ $(i = 1, \dots, 8p)$; die Indikatorfunktion des Intervalls $\bar{I}_i$ bezeichnen wir mit $\chi_i(x)$. Weiterhin, für eine im Intervall (0, 1) definierte Funktion $f(x)$ und für ein Intervall $I=(a;b)$ $(\subseteqq(0,1))$ setzen wir

$$f(I;x) = \begin{cases} f\left(\dfrac{x-a}{b-a}\right) & (x\in I),\\ 0 & \text{sonst.}\end{cases}$$

Es seien

$$\Phi_l(I_1, I_2, I_3; p; x) = \frac{1}{\sqrt{2\,\mathrm{mes}\,I_1}}\, g_l(I_1; p; x) +$$

$$+\frac{1}{\sqrt{2\,\mathrm{mes}\,I_2}}\, h_l(I_2; p; x) - \frac{\sqrt{2p}}{\sqrt{\mathrm{mes}\,I_3}}\,\chi_l(x) + \frac{\sqrt{2p}}{\sqrt{\mathrm{mes}\,I_3}}\,\chi_{l+4p}(x),$$

$$\Phi_{l+4p}(I_1, I_2, I_3; p; x) = \frac{1}{\sqrt{2\,\mathrm{mes}\,I_1}}\, g_l(I_1; p; x) +$$

$$+\frac{1}{\sqrt{2\,\mathrm{mes}\,I_2}}\, h_l(I_2; p; x) + \frac{\sqrt{2p}}{\sqrt{\mathrm{mes}\,I_3}}\,\chi_l(x) - \frac{\sqrt{2p}}{\sqrt{\mathrm{mes}\,I_3}}\,\chi_{l+4p}(x)$$

$(l=1, \dots, 4p)$. Auf Grund von (33) folgt, daß die Treppenfunktionen $\Phi_l(I_1, I_2, I_3; p; x)$ $(l=1, \dots, 8p)$ ein orthonormiertes System im Intervall (0, 2) bilden. Nach der Definition ist (23) offensichtlich, und sind die Funktionen $\Phi_l(I_1, I_2, I_3; p; x)$ $(l=1, \dots, 8p)$ in jedem Intervalle J_k $(k=1, \dots, 8p)$ konstant. Aus (32) folgt (24), weiterhin aus (34) bekommen wir (25) und (26).

Es sei $\{\lambda_k\}_{k=1}^{\infty}$ eine monoton nichtabnehmende Folge von positiven Zahlen mit (20). Dann gibt es eine Indexfolge $(4\leqq)m(1)<\ldots<m(s)<\ldots$ mit

$$\sum_{m=m(s)}^{m(s+1)-1} \frac{1}{\lambda_{\nu(m+1)}\left(\sum_{\mu=1}^{m+1} \lambda_{\nu(\mu)}^{-1}\right)} \geqq 1 \qquad (s=1,2,\ldots). \tag{35}$$

Es seien

$$R(s)=\bigcup_{m=m(s)}^{m(s+1)-1} T(m-1,m-1),$$

und

$$p(m)=\frac{\nu(m)-\nu(m-1)}{8(\nu(m-1)-\nu(m-2))},$$

$$a_{kl}=\begin{cases} \dfrac{1}{(\nu(m)-\nu(m-1))\,2^{2m}\lambda_{\nu(m+1)}\left(\sum_{\mu=1}^{m+1}\lambda_{\nu(\mu)}^{-1}\right)} \\ \qquad ((k,l)\in T(m-1,m-1),\ m=m(1),m(1)+1,\ldots), \\ 0 \qquad \text{sonst.}\end{cases} \tag{36}$$

Mit diesen Bezeichnungen können wir den folgenden Hilfssatz beweisen.

HILFSSATZ 4. *Für jeden Index s gibt es ein orthonormiertes System der Treppenfunktionen $\psi_s(x,y)$, $\psi_{kl}(s;x,y)$ $((k,l)\in R(s))$ in dem Quadrat $F=(0,2)\times(0,2)$ nach dem gewöhnlichen Lebesgueschen Maß mit folgenden Eigenschaften. Es gelten*

(a) $\iint_F \psi_s(x,y)\,dx\,dy=0$, (b) $\iint_F \psi_{kl}(s;x,y)\,dx\,dy=0 \qquad ((k,l)\in R(s))$,

es gibt eine Anordnung $(k,l)\to(i(s,k,l),j(s,k,l))(\in N_+^2)$ der Paare $(k,l)\in R(s)$ derart, daß

$$\text{(c)}\qquad \begin{aligned} &\max\big(i(s,k,l),j(s,k,l);(k,l)\in R(s)\big)<\\ &<\min\big(i(s,k,l),j(s,k,l):(k,l)\in R(s+1)\big)\end{aligned}$$

$(s=1,2,\ldots)$ ist, und für jeden Punkt $(x,y)\in E=(0,1)\times(0,1)$ gibt es Indizes $m_1(s,x,y)\leqq m_2(s,x,y)$, $n_1(s,x,y)\leqq n_2(s,x,y)$ derart, daß

$$\text{(d)}\qquad \sum_{\substack{(k,l)\in R(s)\\ m_1(s,x,y)\leqq i(s,k,l)\leqq m_2(s,x,y)\\ n_1(s,x,y)\leqq j(s,k,l)\leqq n_2(s,x,y)}} a_{k,l}\psi_{k,l}(s;x,y)\geqq C_{13}$$

besteht.

BEWEIS des Hilfssatzes 4. Es sei s_0 eine positive ganze Zahl. Wir nehmen an, daß für $s=1,\ldots,s_0-1$ die Funktionen $\psi_s(x,y)$, $\psi_{kl}(s;x,y)$ $((k,l)\in R(s))$ mit der erforderten Eigenschaften definiert sind. Es sei

$$N(s_0-1)=\max\big(i(s_0-1,k,l),j(s_0-1,k,l):(k,l)\in R(s_0-1)\big).$$

Wir setzen

$$R(s_0, \sigma) = \bigcup_{m=m(s_0)}^{m(s_0)+\sigma} T(m-1, m-1) \qquad (\sigma = 0, \dots, m(s_0+1)-m(s_0)-1).$$

Für jeden Index σ $(0 \leqq \sigma < m(s_0+1)-m(s_0))$ definieren wir das orthonormierte System der Treppenfunktionen $\psi_{kl}(s_0; x, y)$ $((k, l) \in R(s_0, \sigma))$ in F und eine Anordnung $(k, l) \to (i(s_0, \sigma, k, l), j(s_0, \sigma, k, l))$ $(\in N_+^2)$ der Paare $(k, l) \in R(s_0, \sigma)$ derart, daß

$$(37) \qquad i(s_0, \sigma, k, l), \quad j(s_0, \sigma, k, l) > N(s_0-1) \quad ((k, l) \in R(s_0, \sigma))$$

und

$$(38) \qquad \iint_F \psi_{kl}(s_0; x, y)\, dx\, dy = 0 \qquad ((k, l) \in R(s_0, \sigma))$$

gelten, in jedem Rechteck

$$F_{i,j}(s_0, \sigma) = I_i(s_0, \sigma) \times I_j(s_0, \sigma) \quad (i, j = 1, \dots, \nu(m(s_0)+\sigma)-\nu(m(s_0)+\sigma-1))$$

$$\left(I_i(s_0, \sigma) = \left(\frac{i-1}{\nu(m(s_0)+\sigma)-\nu(m(s_0)+\sigma-1)}, \quad \frac{i}{\nu(m(s_0)+\sigma)-\nu(m(s_0)+\sigma-1)}\right)\right)$$

die Funktionen $\psi_{k,l}(s_0; x, y)$ $((k, l) \in R(s_0, \sigma))$ konstant sind, und für jedes Rechteck $F_{i,j}(s_0, \sigma)$ $(i, j = 1, \dots, \nu(m(s_0)+\sigma)-\nu(m(s_0)+\sigma-1))$ Indizes $m_1(s_0, \sigma, i, j) \leqq m_2(s_0, \sigma, i, j)$, $n_1(s_0, \sigma, i, j) \leqq n_2(s_0, \sigma, i, j)$ mit

$$(39) \qquad \begin{aligned} &\sum_{\substack{(k,l) \in R(s_0, \sigma) \\ m_1(s_0,\sigma,i,j) \leqq i(s_0,\sigma,k,l) \leqq m_2(s_0,\sigma,i,j) \\ n_1(s_0,\sigma,i,j) \leqq j(s_0,\sigma,k,l) \leqq n_2(s_0,\sigma,i,j)}} a_{k,l} \Psi_{k,l}(s_0; x, y) \geqq \\ &\geqq C_{13} \sum_{m=m(s_0)}^{m(s_0)+\sigma} \frac{1}{\lambda_{\nu(m+1)} \left(\sum_{\mu=1}^{m+1} \lambda_{\nu(\mu)}^{-1} \right)} \qquad (x \in F_{ij}(s_0, \sigma)) \end{aligned}$$

existieren. Wir wenden den Hilfssatz 3 mit $p = p(m(s_0))$ an, und wir setzen

$$\begin{aligned} &\Psi_{k+(i-1)8p(m(s_0))+\nu(m(s_0)-1),\ l+(j-1)8p(m(s_0))+\nu(m(s_0)-1)}(s_0; x, y) = \\ &= \Phi_k(I_i(s_0, 0), (1, 3/2), (3/2, 2); x)\Phi_l(I_j(s_0, 0), (1, 3/2), (3/2, 2); y) \\ &(k, l = 1, \dots, 8p(m(s_0)), \quad i, j = 1, \dots, \nu(m(s_0)-1)-\nu(m(s_0)-2)). \end{aligned}$$

Weiterhin sei $i(s_0, 0, k, l) = k + N(s_0-1)$, $j(s_0, 0, k, l) = l + N(s_0-1)$. Auf Grund des Hilfssatzes 3 und (36) bilden die Treppenfunktionen $\Psi_{kl}(s_0; x, y)$ $((k, l) \in R(s_0, 0))$ ein orthonormiertes System in F, und werden (37), (38) und (39) für $\sigma = 0$ erfüllt. Weiterhin sind die Funktionen $\Psi_{k,l}(s_0; x, y)$ $((k, l) \in R(s_0, 0))$ in den Rechtecken $F_{i,j}(s_0, 0)$ $(i, j = 1, \dots, \nu(m(s_0))-\nu(m(s_0)-1))$ konstant.

Es sei σ_0 eine nichtnegative ganze Zahl $(\sigma_0 < m(s_0+1)-m(s_0)-1)$. Wir nehmen an, daß die Treppenfunktionen $\Psi_{kl}(s_0; x, y)$ $((k, l) \in R(s_0, \sigma_0))$ und die Anordnung $(k, l) \to (i(s_0, \sigma_0, k, l), j(s_0, \sigma_0, k, l))$ $(\in N_+^2)$ der Paare $(k, l) \in R(s_0, \sigma_0)$ schon derart definiert sind, daß (37), (38) und (39) für $\sigma = \sigma_0$ erfüllt sind. Weiterhin sind die Funktionen $\Psi_{k,l}(s_0; x, y)$ $((k, l) \in R(s_0, \sigma_0))$ in den Rechtecken $F_{i,j}(s_0, \sigma_0)$ $(i, j = 1, \dots, \nu(m(s_0)+\sigma_0)-\nu(m(s_0)+\sigma_0-1))$ konstant.

Es seien $I_2, I_3(\subseteq(1, 2))$ disjunkte Intervalle mit der Eigenschaften, daß alle Funktionen $\Psi_{kl}(s_0, x, y)$ $((k, l)\in R(s_0, \sigma_0))$ im Falle $(x, y)\in I_2\times I_3$ konstant sind. Wir wenden den Hilfssatz 3 im Falle $p=p(m(s_0)+\sigma_0+1)$ an, und wir setzen

$$\Psi_{k+(i-1)8p(m(s_0)+\sigma_0+1)+\nu(m(s_0)+\sigma_0),\, l+(j-1)8p(m(s_0)+\sigma_0+1)+\nu(m(s_0)+\sigma_0)}(s_0; x, y) =$$

$$= \Phi_k(I_i(s_0, \sigma_0), I_2, I_3; p(m(s_0)+\sigma_0+1); x)\Phi_l(I_j(s_0, \sigma_0), I_2, I_3; p(m(s_0)+\sigma_0+1); y)$$

$$\big(k, l = 1, \ldots, 8p(m(s_0)+\sigma_0+1); \quad i, j = 1, \ldots, \nu(m(s_0)+\sigma_0)-\nu(m(s_0)+\sigma_0-1)\big).$$

Auf Grund des Hilfssatzes 3 bilden die Treppenfunktionen $\Psi_{kl}(s_0; x, y)$ $((k, l)\in R(s_0, \sigma_0+1))$ ein orthonormiertes System in F, weiterhin ist (38) für $\sigma=\sigma_0+1$ erfüllt, und sind die Funktionen $\Psi_{k,l}(s_0; x, y)$ $((k, l)\in R(s_0, \sigma_0+1))$ in jedem Rechteck $F_{i,j}(s_0, \sigma_0+1)$ $(i, j=1, \ldots, \nu(m(s_0)+\sigma_0+1)-\nu(m(s_0)+\sigma_0))$ konstant.

Die Anordnung $(k, l)\to(i(s_0, \sigma_0+1, k, l), j(s_0, \sigma_0+1, k, l))$ $(\in N_+^2)$ der Paare $(k, l)\in R(s_0, \sigma_0+1)$ definieren wir durch Induktion. Die Rechtecke $F_{ij}(s_0, \sigma_0)$ $(i, j=1, \ldots, \nu(m(s_0)+\sigma_0)-\nu(m(s_0)+\sigma_0-1))$ ordnen wir in eine Reihenfolge $F_{i_\alpha, j_\alpha}(s_0, \sigma_0)$ $(\alpha=1, \ldots, (\nu(m(s_0)+\sigma_0)-\nu(m(s_0)+\sigma_0-1))^2)$ an. Es seien für $\alpha=1, \ldots, (\nu(m(s_0)+\sigma_0)-\nu(m(s_0)+\sigma_0-1))^2$

$$T_\alpha^{(1)} = \{k\colon (i_\alpha-1)8p(m(s_0)+\sigma_0+1)+\nu(m(s_0)+\sigma_0) < k \leqq$$
$$\leqq (i_\alpha-1)8p(m(s_0)+\sigma_0+1)+4p(m(s_0)+\sigma_0+1)+\nu(m(s_0)+\sigma_0)\},$$

$$T_\alpha^{(2)} = \{k\colon (i_\alpha-1)8p(m(s_0)+\sigma_0+1)+4p(m(s_0)+\sigma_0+1)+\nu(m(s_0)+\sigma_0) < k \leqq$$
$$\leqq i_\alpha 8p(m(s_0)+\sigma_0+1)+\nu(m(s_0)+\sigma_0)\},$$

$$\overline{T}_\alpha^{(1)} = \{l\colon (j_\alpha-1)8p(m(s_0)+\sigma_0+1)+\nu(m(s_0)+\sigma_0) < l \leqq$$
$$\leqq (j_\alpha-1)8p(m(s_0)+\sigma_0+1)+4p(m(s_0)+\sigma_0+1)+\nu(m(s_0)+\sigma_0)\},$$

$$\overline{T}_\alpha^{(2)} = \{l\colon (j_\alpha-1)8p(m(s_0)+\sigma_0+1)+4p(m(s_0)+\sigma_0+1)+\nu(m(s_0)+\sigma_0) < l \leqq$$
$$\leqq j_\alpha 8p(m(s_0)+\sigma_0+1)+\nu(m(s_0)+\sigma_0)\},$$

und $P_0=R(s_0, \sigma_0)$, $P_\alpha=P_{\alpha-1}\cup((T_\alpha^{(1)}\cup T_\alpha^{(2)})\times(\overline{T}_\alpha^{(1)}\cup\overline{T}_\alpha^{(2)}))$. Für jeden Index α $(1\leqq\alpha\leqq(\nu(m(s_0)+\sigma_0)-\nu(m(s_0)+\sigma_0-1))^2)$ definieren wir eine Anordnung $(k, l)\to(i(s_0, \sigma_0+1, \alpha, k, l), j(s_0, \sigma_0+1, \alpha, k, l))$ $(\in N_+^2)$ der Paare $(k, l)\in P_\alpha$ derart, daß

(40) $$i(s_0, \sigma_0+1, \alpha, k, l),\ j(s_0, \sigma_0+1, \alpha, k, l) > N(s_0-1) \qquad ((k, l)\in P_\alpha)$$

gilt, weiterhin Indizes $m_1(s_0, \sigma_0+1, \alpha, i, j)\leqq m_2(s_0, \sigma_0+1, \alpha, i, j)$, $n_1(s_0, \sigma_0+1, \alpha, i, j)\leqq n_2(s_0, \sigma_0+1, \alpha, i, j)$ mit

(41) $$\sum_{\substack{(k,l)\in P_\alpha \\ m_1(s_0,\sigma_0+1,\alpha,i,j)\leqq i(s_0,\sigma_0+1,\alpha,k,l)\leqq m_2(s_0,\sigma_0+1,\alpha,i,j) \\ n_1(s_0,\sigma_0+1,\alpha,i,j)\leqq j(s_0,\sigma_0+1,\alpha,k,l)\leqq n_2(s_0,\sigma_0+1,\alpha,i,j)}} a_{k,l}\Psi_{k,l}(s_0; x, y) \geqq$$
$$\geqq C_{13}\sum_{m=m(s_0)}^{m(s_0)+\sigma_0+1}\frac{1}{\lambda_{\nu(m+1)}\left(\sum_{\mu=1}^{m+1}\lambda_{\nu(\mu)}^{-1}\right)} \qquad ((x, y)\in F_{ij}(s_0, \sigma_0+1))$$

$((i_{\bar\alpha}-1)8p(m(s_0)+\sigma_0+1)<i\leqq i_{\bar\alpha}8p(m(s_0)+\sigma_0+1),\ (j_{\bar\alpha}-1)8p(m(s_0)+\sigma_0+1)<j\leqq$ $\leqq j_{\bar\alpha}8p(m(s_0)+\sigma_0+1),\ \bar\alpha=1,\dots,\alpha)$,
und Indizes $m_1(s_0,\sigma_0+1,\alpha,i_{\bar\alpha},j_{\bar\alpha})\leqq m_2(s_0,\sigma_0+1,\alpha,i_{\bar\alpha},j_{\bar\alpha})$, $n_1(s_0,\sigma_0+1,\alpha,i_{\bar\alpha},j_{\bar\alpha})\leqq$ $\leqq n_2(s_0,\sigma_0+1,\alpha,i_{\bar\alpha},j_{\bar\alpha})$ mit

$$\sum_{\substack{(k,l)\in P_\alpha \\ m_1(s_0,\sigma_0+1,\alpha,i_{\bar\alpha},j_{\bar\alpha})\leqq i(s_0,\sigma_0+1,\alpha,k,l)\leqq m_2(s_0,\sigma_0+1,\alpha,i_{\bar\alpha},j_{\bar\alpha}) \\ n_1(s_0,\sigma_0+1,\alpha,i_{\bar\alpha},j_{\bar\alpha})\leqq j(s_0,\sigma_0+1,\alpha,k,l)\leqq n_2(s_0,\sigma_0+1,\alpha,i_{\bar\alpha},j_{\bar\alpha})}} a_{k,l}\Psi_{k,l}(s_0;x,y)\geqq$$

$$\geqq C_{13}\sum_{m=m(s_0)}^{m(s_0)+\sigma_0}\frac{1}{\lambda_{\nu(m+1)}\left(\sum_{\mu=1}^{m+1}\lambda_{\nu(\mu)}^{-1}\right)}\qquad ((x,y)\in F_{i_{\bar\alpha},j_{\bar\alpha}}(s_0,\sigma_0)) \tag{42}$$

$$(\bar\alpha=\alpha+1,\dots,(\nu(m(s_0)+\sigma_0)-\nu(m(s_0)+\sigma_0-1))^2)$$

existieren.

Für $(k,l)\in R(s_0,\sigma_0)$ setzen wir

$$i(s_0,\sigma_0+1,1,k,l)=\begin{cases} i(s_0,\sigma_0,k,l), & \text{für } i(s_0,\sigma_0,k,l)<m_1(s_0,\sigma_0,i_1,j_1),\\ i(s_0,\sigma_0,k,l)+4p(m(s_0)+\sigma_0+1), & \text{für}\\ \qquad m_1(s_0,\sigma_0,i_1,j_1)\leqq i(s_0,\sigma_0,k,l)\leqq m_2(s_0,\sigma_0,i_1,j_1), \\ i(s_0,\sigma_0,k,l)+8p(m(s_0)+\sigma_0+1), & \text{für}\\ \qquad m_2(s_0,\sigma_0,i_1,j_1)<i(s_0,\sigma_0,k,l), \end{cases}$$

und

$$j(s_0,\sigma_0+1,1,k,l)=\begin{cases} j(s_0,\sigma_0,k,l), & \text{für } j(s_0,\sigma_0,k,l)<n_1(s_0,\sigma_0,i_1,j_1),\\ j(s_0,\sigma_0,k,l)+4p(m(s_0)+\sigma_0+1), & \text{für}\\ \qquad n_1(s_0,\sigma_0,i_1,j_1)\leqq j(s_0,\sigma_0,k,l)\leqq n_2(s_0,\sigma_0,i_1,j_1), \\ j(s_0,\sigma_0,k,l)+8p(m(s_0)+\sigma_0+1), & \text{für}\\ \qquad n_2(s_0,\sigma_0,i_1,j_1)<j(s_0,\sigma_0,k,l), \end{cases}$$

weiterhin für $(k,l)\in P_1-R(s_0,\sigma_0)$

$$i(s_0,\sigma_0+1,1,k,l)=\begin{cases} m_1(s_0,\sigma_0,i_1,j_1)+k-1, & \text{für } k\in T_1^{(1)}\\ m_2(s_0,\sigma_0,i_1,j_1)+4p(m(s_0)+\sigma_0+1)+k, & \text{für } k\in T_1^{(2)}\end{cases}$$

und

$$j(s_0,\sigma_0+1,1,k,l)=\begin{cases} n_1(s_0,\sigma_0,i_1,j_1)+l-1, & \text{für } l\in \bar T_1^{(1)}\\ n_1(s_0,\sigma_0,i_1,j_1)+4p(m(s_0)+\sigma_0+1)+l, & \text{für } l\in \bar T_1^{(2)}.\end{cases}$$

Für die Anordnung $(k,l)\to(i(s_0,\sigma_0+1,1,k,l),j(s_0,\sigma_0+1,1,k,l))$ $(\in N_+^2)$ der Paare $(k,l)\in P_1$ gilt (40) für $\alpha=1$ offensichtlich, weiterhin, auf Grund des Hilfssatzes 3 und (36) erhalten wir, daß (41) und (42) im Falle $\alpha=1$ bestehen.

Es sei α_0 eine positive ganze Zahl $(\alpha_0<(\nu(m(s_0)+\sigma_0)-\nu(m(s_0)+\sigma_0-1))^2)$. Wir nehmen an, daß die Anordnung $(k,l)\to(i(s_0,\sigma_0+1,\alpha_0,k,l),\ j(s_0,\sigma_0+1,\alpha_0,k,l))$ der Paare $(k,l)\in P_{\alpha_0}$ schon derart definiert sind, daß (40), (14) und (42) für $\alpha=\alpha_0$ erfüllt sind.

Für $(k, l)\in P_{\alpha_0}$ setzen wir

$$i(s_0, \sigma_0+1, \alpha_0+1, k, l) = \begin{cases} i(s_0, \sigma_0+1, \alpha_0, k, l), & \text{für } i(s_0, \sigma_0+1, \alpha_0, k, l) < \\ \quad < m_1(s_0, \sigma_0+1_0, i_{\alpha_0}, j_{\alpha_0}), \\ i(s_0, \sigma_0+1, \alpha_0, k, l)+4p(m(s_0)+\sigma_0+1), \\ \text{für } m_1(s_0, \sigma_0+1, \alpha_0, i_{\alpha_0}, j_{\alpha_0}) \leqq i(s_0, \sigma_0+1, \alpha_0, k, l) \leqq \\ \quad \leqq m_2(s_0, \sigma_0+1, \alpha_0, i_{\alpha_0}, j_{\alpha_0}), \\ i(s_0, \sigma_0+1, \alpha_0, k, l)+8p(m(s_0)+\sigma_0+1), & \text{für} \\ \quad m_2(s_0, \sigma_0+1, \alpha_0, i_{\alpha_0}, j_{\alpha_0}) < i(s_0, \sigma_0+1, \alpha_0, k, l), \end{cases}$$

und

$$j(s_0, \sigma_0+1, \alpha_0+1, k, l) = \begin{cases} j(s_0, \sigma_0+1, \alpha_0, k, l), & \text{für } j(s_0, \sigma_0+1, \alpha_0, k, l) < \\ \quad < n_1(s_0, \sigma_0+1, \alpha_0, i_{\alpha_0}, j_{\alpha_0}), \\ j(s_0, \sigma_0+1, \alpha_0, k, l)+4p(m(s_0)+\sigma_0+1), \\ \text{für } n_1(s_0, \sigma_0+1, \alpha_0, i_{\alpha_0}, j_{\alpha_0}) \leqq j(s_0, \sigma_0+1, \alpha_0, k, l) \leqq \\ \quad \leqq n_2(s_0, \sigma_0+1, \alpha_0, i_{\alpha_0}, j_{\alpha_0}), \\ j(s_0, \sigma_0+1, \alpha_0, k, l)+8p(m(s_0)+\sigma_0+1), & \text{für} \\ \quad n_2(s_0, \sigma_0+1, \alpha_0, i_{\alpha_0}, j_{\alpha_0}) < j(s_0, \sigma_0+1, \alpha_0, k, l), \end{cases}$$

weiterhin für $(k, l)\in P_{\alpha_0+1}-P_{\alpha_0}$

$$i(s_0, \sigma_0+1, \alpha_0+1, k, l) = \begin{cases} m_1(s_0, \sigma_0+1, \alpha_0, i_{\alpha_0}, j_{\alpha_0})+k-1, & \text{für } k\in T^{(1)}_{\alpha_0+1}, \\ m_2(s_0, \sigma_0+1, \alpha_0, i_{\alpha_0}, j_{\alpha_0})+4p(m(s_0)+\sigma_0+1)+k, & \text{für} \\ \quad k\in T^{(2)}_{\alpha_0+1}, \end{cases}$$

und

$$j(s_0, \sigma_0+1, \alpha_0+1, k, l) = \begin{cases} n_1(s_0, \sigma_0+1, \alpha_0, i_{\alpha_0}, j_{\alpha_0})+l-1, & \text{für } l\in \overline{T}^{(1)}_{\alpha_0+1}, \\ n_2(s_0, \sigma_0+1, \alpha_0, i_{\alpha_0}, j_{\alpha_0})+4p(m(s_0)+\sigma_0+1)+l, & \text{für} \\ \quad l\in \overline{T}^{(2)}_{\alpha_0+1}. \end{cases}$$

Für die Anordnung $(k, l)\to(i(s_0, \sigma_0+1, \alpha_0+1, k, l), j(s_0, \sigma_0+1, \alpha_0+1, k, l)\,(\in N_+^2))$ der Paare $(k, l)\in P_{\alpha_0+1}$ gilt (40) für $\alpha=\alpha_0+1$ offensichtlich, weiterhin, auf Grund des Hilfssatzes 3 und (36), nach der Voraussetzung erhalten wir, daß (41) und (42) im Falle $\alpha=\alpha_0+1$ bestehen. Durch Induktion für α bekommen wir die Anordnung

$$(43)\quad \begin{aligned} &(k, l) \to \big(i(s_0, \sigma_0+1, (v(m(s_0)+\sigma_0)-v(m(s_0)+\sigma_0-1))^2, k, l), \\ &j(s_0, \sigma_0+1, (v(m(s_0)+\sigma_0)-v(m(s_0)+\sigma_0-1))^2, k, l)\big) \quad (\in N_+^2) \end{aligned}$$

für die Paare $(k, l)\in P_{(v(m(s_0)+\sigma_0)-v(m(s_0)+\sigma_0-1))^2}=R(s_0, \sigma_0+1)$, und für diese Anordnung sind (40), (41) und (42) im Falle $\alpha=(v(m(s_0)+\sigma_0)-v(m(s_0)+\sigma_0-1))^2$ erfüllt. Die Anordnung $(k, l)\to(i(s_0, \sigma_0+1, k, l), j(s_0, \sigma_0+1, k, l))$ sei mit der Anordnung (43) gleich; dann werden (37) und (39) sich für $\sigma=\sigma_0+1$ auf Grund von (40), (41) und (42) erfüllt.

Durch Induktion für σ erhalten wir ein im F orthonormiertes System der Treppenfunktionen $\Psi_{kl}(s_0, x, y)$ $((k, l)\in R(s_0, m(s_0+1)-m(s_0)-1)=R(s_0)$ und eine Anordnung

$$(44)\quad (k, l) \to \big(i(s_0, (m(s_0+1)-m(s_0)-1, k, l), \quad j(s_0, m(s_0+1)-m(s_0)-1, k, l)\big)$$

der Paare $(k, l) \in R(s_0)$ derart, daß (37), (38) und (39) im Fall $\sigma = m(s_0+1) - m(s_0)$ erfüllt sind. Die Anordnung $(k, l) \to (i(s_0, k, l), j(s_0, k, l))$ sei mit der Anordnung (44) gleich. Dann werden die Forderungen (b), (c) und (d) des Hilfssatzes 4 für $s = s_0$ bestehen.

Die Funktionen $\Psi_{kl}(s_0, x, y)$ $((k, l) \in R(s_0))$ sind Treppenfunktionen in F. So gibt es eine Einteilung von F in paarweise disjunkte Rechtecke $K_1, \ldots, K_{\lambda(s_0)}$ derart, daß jede Funktion $\Psi_{kl}(s_0, x, y)$ $((k, l) \in R(s_0))$ in jedem Rechteck K_l $(l = 1, \ldots, \lambda(s_0))$ konstant ist. Die zwei Hälfte von K_l bezeichnen wir mit K_l', K_l'' $(l = 1, \ldots, \lambda(s_0))$. Dann sei

$$\Psi_s(x, y) = \begin{cases} \dfrac{1}{2}, & \text{für} \quad (x, y) \in K_l', \\ -\dfrac{1}{2}, & \text{für} \quad (x, y) \in K_l'' \end{cases}$$

$(l = 1, \ldots, \lambda(s_0))$. Es ist offensichtlich, daß die Treppenfunktionen $\Psi_{s_0}(x, y)$, $\Psi_{kl}(s_0, x, y)$ $((k, l) \in R(s_0))$ ein orthonormiertes System in F bilden, weiterhin die Bedingung (a) des Hilfssatzes 4 im Falle $s = s_0$ erfüllt ist.

Damit haben wir den Hilfssatz 4 bewiesen.

Beweis des Satzes 6. Wir wenden die Bezeichnungen des Hilfssatzes 4 an. Auf Grund von (36) gilt (19) für die Folge $\{a_{kl}\}_{k,l=1}^{\infty}$. Es sei

$$\overline{\Psi}_s(x, y) = \Psi_s\left(\frac{x}{2}, \frac{y}{2}\right),$$

$$\overline{\Psi}_{kl}(s; x, y) = \Psi_{kl}\left(s; \frac{x}{2}, \frac{y}{2}\right) \qquad ((k, l) \in R(s))$$

$(s = 1, 2, \ldots)$. Nach dem Hilfssatz 4 ist es offensichtlich, daß für jeden Index s die Treppenfunktionen $\overline{\Psi}_s(x, y)$, $\overline{\Psi}_{kl}(s; x, y)$, $((k, l) \in R(s))$ in dem Einheitsquadrat $E = (0,1) \times (0,1)$ ein orthonormiertes System bilden,

$$\text{(45)} \qquad \begin{aligned} &\iint_E \overline{\Psi}_s(x, y)\, dx\, dy = 0, \\ &\iint_E \overline{\Psi}_{kl}(s, x, y)\, dx\, dy = 0 \qquad ((k, l) \in R(s)) \end{aligned}$$

bestehen, und eine Anordnung $(k, l) \to (i(s, k, l), j(s, k, l))$ $(\in N_+^2)$ der Paare $(k, l) \in R(s)$ derart gibt, daß

$$\text{(46)} \qquad \begin{aligned} &\max\left(i(s, k, l), j(s, k, l) \colon (k, l) \in R(s)\right) < \\ &< \min\left(i(s, k, l), j(s, k, l) \colon (k, l) \in R(s+1)\right) \end{aligned}$$

gilt, und für jeden Punkt $(x, y) \in H = \left(0, \frac{1}{2}\right) \times \left(0, \frac{1}{2}\right)$ Indizes $m_1(s, x, y) \equiv$

$\leqq m_2(s, x, y)$, $n_1(s, x, y) \leqq n_2(s, x, y)$ mit

(47) $$\sum_{\substack{(k,l)\in R(s)\\ m_1(s,x,y)\leqq i(s,k,l)\leqq m_2(s,x,y)\\ n_1(s,x,y)\leqq j(s,k,l)\leqq n_2(s,x,y)}} a_{k,l}\Psi_{k,l}(s; x, y) \geqq C_{13} \qquad ((x,y)\in H)$$

existieren.

Wir definieren ein orthonormiertes System der Treppenfunktionen $\chi_s(x, y)$ $(s=1, 2, \dots)$, $\varphi_{kl}(x, y)$ $\left((k, l)\in \bigcup_{s=1}^{\infty} R(s)\right)$, eine Folge der einfachen Mengen H_s $(s=1, 2, \dots)$[1], und eine Anordnung $(k, l)\to(i(k, l), j(k, l))$ $(\in N_+^2)$ der Paare $(k, l)\in \bigcup_{s=1}^{\infty} R(s)$ derart, daß die Mengen H_s $(s=1, 2, \dots)$ stochastisch unabhängig sind, weiterhin für jeden Index s

(48) $$\text{mes } H_s = \frac{1}{4}$$

und

(49) $$\max_{\substack{N(s-1)<\varkappa_1\leqq\varkappa_2\leqq N(s)\\ N(s-1)<\lambda_1\leqq\lambda_2\leqq N(s)}} \left|\sum_{\substack{(k,l)\in R(s)\\ \varkappa_1\leqq i(k,l)\leqq\varkappa_2\\ \lambda_1\leqq j(k,l)\leqq\lambda_2}} a_{k,l}\varphi_{k,l}(x, y)\right| \geqq C_{13} \qquad ((x,y)\in H_s)$$

bestehen.

Es sei $H_1=H$, und $\chi_1(x, y)=\Psi_1(x, y)$, $\varphi_{kl}(x, y)=\Psi_{kl}(1; x, y)$ $((k, l)\in R(1))$, $i(k, l)=i(1, k, l)$, $j(k, l)=j(1, k, l)$ $((k, l)\in R(1))$. Auf Grund von (47) gilt (48) und (49) für $s=1$. Es sei s_0 eine positive ganze Zahl. Wir nehmen an, daß die Treppenfunktionen $\chi_s(x, y)$ $(s=1, \dots, s_0)$, $\varphi_{kl}(x, y)$ $\left((k, l)\in \bigcup_{s=1}^{s_0} R(s)\right)$, die einfachen Mengen $H_1, \dots, H_{s_0}$, und die Anordnung $(k, l)\to(i(k, l), j(k, l))$ der Paare $(k, l)\in \bigcup_{s=1}^{s_0} R(s)$ schon derart definiert sind, daß diese Funktionen ein orthonormiertes System in E bilden, die Mengen $H_1, \dots, H_{s_0}$ stochastisch unabhängig sind, die Anordnung $(k, l)\to$ $\to(i(k, l), j(k, l))$ $\left((k, l)\in \bigcup_{s=1}^{s_0} R(s)\right)$ umkehrbar eindeutig ist, und (48), (49) für $s=1, \dots, s_0$ erfüllt sind.

Für eine im E definierte Funktion $f(x, y)$ und für ein Rechteck $T=(a_1, b_1)\times$ $\times(a_2, b_2)(\subseteqq E)$ setzen wir

$$f(T, x, y) = \begin{cases} f\left(\dfrac{x-a_1}{b_1-a_1}, \dfrac{y-a_2}{b_2-a_2}\right) & ((x, y)\in T),\\ 0 & \text{sonst,}\end{cases}$$

weiterhin für eine Menge $G(\subseteqq E)$ bezeichnet $G(I)$ diejenige Menge, die mit der linearen Transformation $u=(b_1-a_1)x+a_1$, $v=(b_2-a_2)y+a_2$ aus der Menge G entsteht.

Da die Funktionen $\chi_s(x, y)$ $(s=1, \dots, s_0)$, $\varphi_{kl}(x, y)$ $\left((k, l)\in \bigcup_{s=1}^{s_0} R(s)\right)$ Treppenfunktionen sind, und die Mengen $H_1, \dots, H_{s_0}$ einfach sind, gibt es eine Einteilung

[1] Die Menge H wird einfach genannt, wenn sie die Vereinigung endlichvieler Rechtecke ist.

von E in endlich viele disjunkte Rechtecke $G_1, \ldots, G_\varrho$ derart, daß diese Funktionen in jedem Rechteck G_r konstant sind, und jede Menge $H_s (s=1, \ldots, s_0)$ die Vereinigung gewisser G_r ist. Dann setzen wir

$$\chi_{s+1}(x, y) = \sum_{r=1}^{\varrho} \Psi_{s_0+1}(G_r; x, y),$$

$$\varphi_{kl}(x, y) = \sum_{r=1}^{\varrho} \Psi_{kl}(G_r; s_0+1; x, y) \qquad ((k, l) \in R(s_0+1)),$$

$$H_{s_0+1} = \bigcup_{r=1}^{\varrho} H(G_r),$$

und sei $i(k, l) = i(s_0+1, k, l), j(k, l) = j(s_0+1, k, l)$, für $(k, l) \in R(s_0+1)$.

Aus (45) folgt, daß die Treppenfunktionen $\chi_s(x, y)$ $(s=1, \ldots, s_0+1)$, $\varphi_{kl}(x, y)$ $\left((k, l) \in \bigcup_{s=1}^{s_0+1} R(s)\right)$ im E ein orthonormiertes System bilden. Die Menge H_{s_0+1} ist offensichtlich einfach, (48) ist für $s = s_0+1$ erfüllt, und die Menge $H_1, \ldots, H_{s_0+1}$ sind stochastisch unabhängig. Aus (46) folgt, daß die Anordnung $(k, l) \to (i(k, l), j(k, l))$ $\left((k, l) \in \bigcup_{s=1}^{s_0+1} R(s)\right)$ umkehrbar eindeutig ist. Weiterhin, auf Grund von (47) gilt (49) für $s = s_0+1$. Das orthonormierte System der Treppenfunktionen $\chi_s(x, y)$ $(s=1, 2, \ldots)$, $\varphi_{kl}(x, y)$ $\left((k, l) \in \bigcup_{s=1}^{\infty} R(s)\right)$, die Mengenfolge $H_1, H_2, \ldots$ und die Anordnung $(k, l) \to (i(k, l), j(k, l)) (\in N_+^2)$ $\left((k, l) \in \bigcup_{s=1}^{\infty} R(s)\right)$ mit erforderten Eigenschaften bekommen wir durch Induktion.

Die Funktionen $\varphi_{kl}(x, y)$ $\left((k, l) \in N_+^2 - \bigcup_{s=1}^{\infty} R(s)\right)$ seien mit der Funktionen $\chi_s(x, y)$ $(s=1, 2, \ldots)$ gleich. Die Anordnung $(k, l) \to (i(k, l), j(k, l))$ $\left((k, l) \in N_+^2 - \bigcup_{s=1}^{\infty} R(s)\right)$ definieren wir derart, daß die Abbildung $(k, l) \to (i(k, l), j(k, l))$ $((k, l) \in N_+^2)$ eine umkehrbar eindeutige Abbildung von N_+^2 auf sich selbst sei.

Aus (49) bekommen wir, daß im Falle $(x, y) \in H_s$ von (x, y) abhängige Indizes $N(s-1) < m_1(s, x, y) \leqq m_2(s, x, y) \leqq N(s)$, $N(s-1) < n_1(s, x, y) \leqq n_2(s, x, y) \leqq N(s)$ mit

$$(50) \qquad \Big| \sum_{\substack{(k, l) \\ m_1(s, x, y) \leqq i(k, l) \leqq m_2(s, x, y) \\ n_1(s, x, y) \leqq j(k, l) \leqq n_2(s, x, y)}} a_{k,l} \varphi_{k,l}(x, y) \Big| \geqq C_{13} \qquad ((x, y) \in H_s)$$

existieren.

Es sei $H = \overline{\lim_{s \to \infty}} H_s$. Da die Mengen H_s stochastisch unabhängig sind, auf Grund von (48), durch Anwendung des zweiten Borel—Cantellischen Lemmas ergibt sich

$$\text{mes } H = 1.$$

Für $(x, y) \in H$ gilt aber (50) für unendlich viele s, woraus folgt, daß in diesem Punkt

$$\overline{\lim_{\substack{\varkappa_1, \varkappa_2 \to \infty \\ \lambda_1, \lambda_2 \to \infty}}} \Big| \sum_{\substack{(k, l) \\ \varkappa_1 \leqq i(k, l) \leqq \varkappa_2 \\ \lambda_1 \leqq j(k, l) \leqq \lambda_2}} a_{k,l} \varphi_{k,l}(x, y) \Big| \geqq C_{13}$$

gilt, d. h. divergiert die Anordnung

$$\sum_{\alpha=1}^{\infty}\sum_{\beta=1}^{\infty} a_{k_\alpha, l_\beta}\varphi_{k_\alpha, l_\beta}(x)$$

der Reihe

$$\sum_{k=1}^{\infty}\sum_{l=1}^{\infty} a_{kl}\varphi_{kl}(x, y)$$

in E fast überall, wobei k_α, l_β diejenige Indizes sind, für die $\alpha=i(k_\alpha, l_\beta)$, $\beta=j(k_\alpha, l_\beta)$ erfüllt sind. Damit haben wir den Satz 6 im Falle bewiesen, wenn der Raum das Einheitsquadrat mit dem gewöhnlichen Lebesgueschen Maß ist.

LITERATURVERZEICHNIS

[1] Kaczmarz, S., Notes on orthogonal series. II, *Studia Math.* **5** (1935), 103—106. *Zbl.* **13**, 109.

[2] Menchoff, D. E., Sur les séries de fonctions orthogonales (Première Partie), *Fund. Math.* **4** (1923), 82—105.

[3] Móricz, F., On the convergence in a restricted sense of multiple series, *Anal. Math.* **5** (1979), 135—147. *MR* **82i**: 40003.

[4] Orlicz, W., Zur Theorie der Orthogonalreihen, *Bull. Internat. Acad. Sci. Polonaise*, Cracovie, 1927, 81—115.

[5] Pandžakidze, S. P., Unconditional convergence of double orthogonal series, *Soobšč. Akad. Nauk Gruzin. SSR* **38** (1965), 521—526 (in Russian). *MR* **33** # 4573

[6] Tandori, K., Über die orthogonalen Funktionen. X. (Unbedingte Konvergenz), *Acta Sci. Math. (Szeged)* **23** (1962), 185—221. *MR* **26** # 1688.

(Received May 6, 1980)

JÓZSEF ATTILA TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR
BOLYAI INTÉZET
ARADI VÉRTANÚK TERE 1
H—6720 SZEGED
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 199—219

COMPACTIFICATIONS FOR SYNTOPOGENOUS SPACES

KÁLMÁN MATOLCSY

Introduction

The most important compactifications of a "classical" topological space are Alexandroff's one-point compactifications, Wallman-type compactifications and Smirnov's compactifications, if the space in question is completely regular.

Smirnov's theory of compactifications of proximity spaces was generalized by Á. Császár [3] by introducing the notion of double compactification of syntopogenous spaces and of symmetrizable compactifications of symmetrizable syntopological spaces.

In this paper *one-point compactifications* of a non-compact syntopogenous space (§ 4) and *simple compactifications* of an arbitrary syntopogenous space (§ 5) will be considered. The classical methods created by Alexandroff and Wallman, respectively, are recognizable in the construction of these compact extensions for the special cases.

In the study of these compactifications certain new separation axioms of syntopogenous spaces will play an important role; § 1—2 deals with these. As an introduction to § 4—5, some basic notions and results on syntopogenous extension theory will be recalled in § 3.

Speaking of syntopogenous structures, the terminology of [1] will be used throughout the paper.

I am very grateful to Professor Á. Császár for his valuable remarks and advices.

1. Separation axioms for syntopogenous spaces

It is a very special property for a syntopogenous space to be symmetrical, therefore in general one has to be satisfied with weaker conditions. In particular, for classical topological spaces, such ones are the separation axioms (S_1) and (S_2) of [2], (complete) regularity, and normality.

The properties mentioned here will be generalized in this paragraph to syntopogenous spaces. In this respect J. L. Sieber and W. J. Pervin obtained some results [9], but we shall go in another direction, therefore our theorems will be different from theirs. A comparative study of both theories would make up a supplementary paper; here we cannot deal with this problem.

A syntopogenous space $[E, \mathscr{S}]$ will be called an *S_1-space* if it satisfies the following axiom:

1980 *Mathematics Subject Classification.* Primary 54A15; Secondary 54D35.
Key words and phrases. Syntopogenous spaces, separation axioms, compactifications.

(S_1) *For every* $<\in\mathscr{S}$ *there exists* $<_1\in\mathscr{S}$ *such that*

$$x, y\in E, \quad x < E-y \quad \text{imply} \quad y <_1 E-x.$$

In such a case it will be also said that $\mathscr{S}$ is an S_1-*structure*; we shall use a similar terminology in connection with the further axioms, too.

LEMMA 1.1. *A syntopogenous space* $[E, \mathscr{S}]$ *is an* S_1-*space iff* $\mathscr{S}^b \sim \mathscr{S}^{bc}$. $\blacksquare$

For the sake of simpler formulation of our results, let us introduce two operators defined on order families. First of all, by the *composition* of two semi-topogenous orders $<_1$, $<_2$ on E we shall mean the order $<_1 \cdot <_2$ see ([5]) for which

$$A(<_1 \cdot <_2)B \Leftrightarrow A <_1 C <_2 B \quad \text{for some} \quad C\subset E.$$

If $\mathscr{A}$ is an order family on E, then obviously

$$\mathscr{A}^{\triangledown} = \{<^c \cdot <: \, <\in\mathscr{A}\}$$

and

$$\mathscr{A}^{\vartriangle} = \{< \cdot <^c: \, <\in\mathscr{A}\}$$

are symmetrical order families on E both of them coarser than $\mathscr{A}$ (cf. [5], (2.19)).

We shall call a syntopogenous space $[E, \mathscr{S}]$ an S_2-*space* if it has the following property:

(S_2) *For an arbitrary* $<\in\mathscr{S}$ *one can find an order* $<_1\in\mathscr{S}$ *such that* $x, y\in E$, $x<E-y$ *imply* $x<_1 C<_1^c E-y$ *with a suitable set* $C\subset E$.

LEMMA 1.2. $[E, \mathscr{S}]$ *is an* S_2-*space iff* $\mathscr{S}^b \sim \mathscr{S}^{\Delta b}$. $\blacksquare$

Every S_2-space is an S_1-space. It is obvious that if $\mathscr{S}$ is a topology, $[E, \mathscr{S}]$ is an S_i-space iff the associated classical topological space is S_i $(i=1, 2)$ in the sense of [2].

PROPOSITION 1.3. *Any separated* S_i-*structure is a* T_i-*structure* $(i=1, 2)$. *For topogenous structures the inverse statement is also true.*

(Cf. [2], (2.5.10), (2.5.17).) $\blacksquare$

EXAMPLE 1.4. A T_i-space does not necessarily satisfy condition (S_i) $(i=1, 2)$ even if it is perfect. This can be shown by the example of $\mathscr{S}=(\mathscr{I}^t \vee \mathscr{I}^c)^p$, where $\mathscr{I}$ is the natural syntopogenous structure on the real line. In fact $\mathscr{S}$ is a perfect T_2-structure, but it does not have property (S_1), because $\mathscr{S}^b=(\mathscr{I}^t \vee \mathscr{I}^c)^b$, $\mathscr{S}^{bc}=(\mathscr{I} \vee \mathscr{I}^{tc})^b$, therefore $\mathscr{S}^b \sim \mathscr{S}^{bc}$ does not hold. $\blacksquare$

LEMMA 1.5. $\mathscr{S}$ *is an* S_1- *or* S_2-*structure iff* $\mathscr{S}^p$ *has the same property.* $\blacksquare$

A syntopogenous space $[E, \mathscr{S}]$ will be called *regular* (or an S_3-*space*) if it satisfies the following condition stronger than (S_2):

(S_3) *For each* $<\in\mathscr{S}$ *there exists* $<_1\in\mathscr{S}$ *such that* $x\in E, x<V$ *imply* $x<_1 C<_1^c V$ *for a suitable set* $C\subset E$.

It is clear that a topology is regular iff its classical equivalent is regular in the traditional sense. $\blacksquare$

LEMMA 1.6. *A syntopogenous space $[E, \mathcal{S}]$ is regular if and only if $\mathcal{S}^p \sim \mathcal{S}^{\Delta p}$.* ∎

LEMMA 1.7. *If $\mathcal{S}$ is a regular syntopogenous structure, then $\mathcal{S}^p$ is also a regular syntopology.* ∎

PROPOSITION 1.8. *A subspace of an S_i-space is also an S_i-space* $(i=1, 2)$.

PROOF. If $<_1, <_2$ are semi-topogenous orders on a set E, and $E_0 \subset E$, then $(<_1 \cdot <_2)|E_0 \subset (<_1|E_0) \cdot (<_2|E_0)$ (see [5], (3.3)). ∎

In order to characterize regular syntopologies, let us generalize the concept of Efremovich's local proximity relation (see e.g. [10]) as follows:

An order family $\mathcal{S}$ is a *local syntopogenous structure* on the set E, if

(L_1) *for every $<_1, <_2 \in \mathcal{S}$ there exists $< \in \mathcal{S}$ such that $<_1 \cup <_2 \subset <$,*

and

(L_2) *for any $< \in \mathcal{S}$ an order $<_1 \in \mathcal{S}$ can be found such that $x \in E$, $x < B$ imply $x <_1 C <_1 B$ with a suitable set $C \subset E$.*

PROPOSITION 1.9. *If $\mathcal{S}$ is a local syntopogenous structure, then $\mathcal{S}^p$ is a syntopology.*

PROOF. Suppose $< \in \mathcal{S}$, and let $<_1$ be a member of $\mathcal{S}$ chosen in accordance with (L_2). If $A <^p B$, then, for any $x \in A$, there exists a set $C_x \subset E$ with $x <_1 C_x <_1 B$, so that we have $A <_1^p \bigcup_{x \in A} C_x <_1^p B$. This gives $\mathcal{S}^p < \mathcal{S}^{p2}$. ∎

THEOREM 1.10. *A syntopological space $[E, \mathcal{S}]$ is regular iff there exists a symmetrical local syntopogenous structure $\mathcal{S}_0$ on E such that $\mathcal{S} \sim \mathcal{S}_0^p$.*

PROOF. Assume $\mathcal{S}_0^p \sim \mathcal{S}$ for some symmetrical local syntopogenous structure $\mathcal{S}_0$ on E. Suppose $< \in \mathcal{S}$, $< \subset <_0^p$, $<_0 \in \mathcal{S}_0$, let $<_1 \in \mathcal{S}_0$ be chosen for $<_0$ in accordance with (L_2), finally put $<_1^p \subset <_2 \in \mathcal{S}$. If $x < B$, then $x <_0^p B$, i.e. $x <_0 B$, therefore $x <_1 C <_1 B$ for some $C \subset E$. Since $<_1$ is symmetrical, we have $E-B <_1 E-C$, consequently $E-B <_1^p E-C$. Thus $X <_1^p C <_1^{pc} B$, so that $x(<_2 \cdot <_2^c)B$. This means $\mathcal{S} < \mathcal{S}^{\Delta p}$, and owing to (1.6) $\mathcal{S}$ is regular.

Conversely, if $\mathcal{S}$ is regular, then $\mathcal{S}_0 = \mathcal{S}^{\Delta}$ is a symmetrical local syntopogenous structure for which $\mathcal{S} \sim \mathcal{S}_0^p$. In fact, it can be easily seen that $\mathcal{S}_0$ is a symmetrical order family satisfying (L_1). We need to show that $\mathcal{S}_0$ has property (L_2). Suppose that $< \in \mathcal{S}$, $<_1 \in \mathcal{S}$ are such that $< \subset <_1^s$, $<_2 \in \mathcal{S}$ is chosen for $<_1$ by (S_3). In this case $< \subset <_1 \subset <_2$. If $x < C <^c B$, then there exists $X \subset E$ such that $x <_1 X <_1 C$, and one can find a set $A \subset E$ for which $x <_2 A <_2^c X$. So that choosing the order $<_2 \cdot <_2^c$ for $< \cdot <^c$, $\mathcal{S}_0$ fulfils (L_2). Finally, in view of (1.6), $\mathcal{S} \sim \mathcal{S}_0^p$. ∎

In this paper we shall say that a syntopology $\mathcal{S}$ on E is *symmetrizable* if there exists a symmetrical syntopogenous structure $\mathcal{S}_0$ on E such that $\mathcal{S} \sim \mathcal{S}_0^p$. (This definition is essentially equivalent to the definition of a symmetrizable syntopology in [3].)

As a syntopogenous structure always satisfies (L_1)–(L_2), from (1.10) we can deduce the following result.

COROLLARY 1.11. *Any symmetrizable syntopology is regular.* ▮

According to the definition (2.3) in [5], a syntopogenous space $[E, \mathscr{S}]$ will be called *normal* if for every $<\in\mathscr{S}$ there exists $<_1\in\mathscr{S}$ for which $A<^c C<B$ implies $A<_1 C'<_1^c B$ with a suitable set $C'\subset E$. (We must call attention to the terminological difference between [5] and the present paper, there this notion was called "weakly normal space".)

LEMMA 1.12. *A syntopogenous structure $\mathscr{S}$ is normal iff $\mathscr{S}^\triangledown < \mathscr{S}^\vartriangle$ holds.* ▮

EXAMPLE 1.13. Let $\mathfrak{S}$ be a *lattice* of the set E in the following sense: $\emptyset$, $E\in\mathfrak{S}$, and $A, B\in\mathfrak{S}$ implies $A\cap B$, $A\cup B\in\mathfrak{S}$. Then the complements of the sets of $\mathfrak{S}$ form also a lattice denoted by $\mathfrak{S}^c$. If $<$ is the topogenous order generated by $\mathfrak{S}$, $\mathscr{T}=\{<\}$ is a normal topogenous structure iff $A, B\in\mathfrak{S}^c$, $A\cap B=\emptyset$ imply the existence of sets $C, D\in\mathfrak{S}$ such that $A\subset C$, $B\subset D$ and $C\cap D=\emptyset$. (Cf. [2], (6.1.53) (c).) ▮

The following theorem is taken in full from [5]:

THEOREM 1.14 ([5], (2.20)). *A syntopogenous structure $\mathscr{S}$ is normal iff $\mathscr{S}^\triangledown$ is a syntopogenous structure. In this case $\mathscr{S}^\triangledown$ is the finest of all symmetrical syntopogenous structures coarser than $\mathscr{S}$.* ▮

As it is shown by the example of the natural syntopogenous structure $\mathscr{I}$ of the real line $\mathbf{R}$, a normal syntopology is not necessarily symmetrizable (in fact, for any $\varepsilon>0$, $\emptyset\neq A<_\varepsilon^c C<_\varepsilon B$ implies $B=\mathbf{R}$, but $\mathscr{I}$ does not possess property (S_1)). The solution of the problem concerning the symmetrizability of a normal syntopology will issue from (1.17). First of all let us consider the following simple statement:

LEMMA 1.15. *Any regular syntopogenous structure $\mathscr{S}$ satisfies the condition $\mathscr{S}^p < \mathscr{S}^{cp}$, which is stronger than (S_1). This property is equivalent to (S_1) provided $\mathscr{S}$ is perfect.*

PROOF. If $\mathscr{S}$ is regular, then $\mathscr{S}^\vartriangle < \mathscr{S}^c$ implies $\mathscr{S}^p < \mathscr{S}^{\vartriangle p} < \mathscr{S}^{cp}$ (see (1.6)). Further suppose $\mathscr{S}^p < \mathscr{S}^{cp}$, from this $\mathscr{S}^b = \mathscr{S}^{pb} < \mathscr{S}^{cpb} = \mathscr{S}^{cb} = \mathscr{S}^{bc}$ follows, thus $\mathscr{S}$ is an S_1-structure. Conversely, if $\mathscr{S}$ is an S_1-syntopology, then $\mathscr{S}^p < \mathscr{S}^b \sim \mathscr{S}^{bc} = \mathscr{S}^{cb} = \mathscr{S}^{ccpcp} = \mathscr{S}^{pcp} = \mathscr{S}^{cp}$ (cf. [1], (5.33)). ▮

EXAMPLE 1.16. With the notations of (1.13) $\mathscr{T}^p < \mathscr{T}^{cp}$ holds iff $x\in A\in\mathfrak{S}$ implies $x\in B\subset A$ for some $B\in\mathfrak{S}^c$ (cf. [2], (6.1.52) (a)). ▮

THEOREM 1.17. *For a normal syntopogenous structure $\mathscr{S}$ the equivalence $\mathscr{S}^p \sim \mathscr{S}^{\triangledown p}$ is true iff $\mathscr{S}^p < \mathscr{S}^{cp}$.*

PROOF. $\mathscr{S}^\triangledown < \mathscr{S}^c$, hence $\mathscr{S}^p \sim \mathscr{S}^{\triangledown p}$ implies $\mathscr{S}^p < \mathscr{S}^{cp}$. Conversely, if $\mathscr{S}^p < \mathscr{S}^{cp}$, then assume $<\in\mathscr{S}$, $<_1\in\mathscr{S}$, $< \subset <_1^2$, $<_2\in\mathscr{S}$, $<_1^p \subset <_2^{cp}$, finally $<_3\in\mathscr{S}$, $<_1 \cup <_2 \subset <_3$. In this case $A<^p B$ implies $x<B$ for each $x\in A$, consequently, there is $C_x\subset E$ such that $x<_1 C_x <_1 B$. From this we infer the inequality $x<_3^{cp} C_x <_3 B$, which means $x<_3^c C_x <_3 B$. We got $x(<_3^c \cdot <_3) B$ for every $x\in A$, so that $A(<_3^c \cdot <_3)^p B$. Because of the arbitrary choice of $<$, we can write $\mathscr{S}^p < \mathscr{S}^{\triangledown p}$. By $\mathscr{S}^\triangledown < \mathscr{S}$, the converse inequality is clear. ▮

COROLLARY 1.18. *Let $\mathcal{S}$ be a normal syntopology. Then the following statements are equivalent:*
(1.18.1) $\mathcal{S}$ *is symmetrizable*;
(1.18.2) $\mathcal{S} \sim \mathcal{S}^{\nabla p}$;
(1.18.3) $\mathcal{S}$ *is an* S_1*-syntopology.*

PROOF. (1.18.1)⇔(1.18.2): If $\mathcal{S} \sim \mathcal{S}_0^p$ for a symmetrical syntopogenous structure $\mathcal{S}_0$, then by (1.14) $\mathcal{S}_0 < \mathcal{S}^{\nabla} < \mathcal{S}$, hence $\mathcal{S}^{\nabla p} \sim \mathcal{S}$. The converse statement is obvious. (1.18.2)⇔(1.18.3): In view of (1.15) this is a direct consequence of (1.17).

2. Separation axioms in compact spaces

A syntopogenous space $[E, \mathcal{S}]$ (or shortly $\mathcal{S}$) will be said to be *locally compact* if there exists an order $< \in \mathcal{S}$ such that for any $x \in E$ an $\mathcal{S}$-compact subset K of E can be found with $x < K$.

A topological space is locally compact iff it is associated with a locally compact classical topology (in the sense of [2]). Every compact syntopogenous space is locally compact.

PROPOSITION 2.1. *A syntopogenous structure $\mathcal{S}$ is locally compact iff so is $\mathcal{S}^p$.*

PROOF. For $x \in E$ and $< \in \mathcal{S}$, $x < K$ is equivalent to $x <^p K$, further K is compact in $\mathcal{S}$ iff it is compact in $\mathcal{S}^p$ (see [1], (15.78), (15.79)). ∎

In order to set the notion of local compactness in its proper light, let us consider the following examples.

EXAMPLE 2.2. As is usual, put $\mathcal{H} = \mathcal{I}^{sb}$. Then, for any $x \in \mathbf{R}$ and $\varepsilon > 0$, the interval $K = [x-\varepsilon, x+\varepsilon]$ is a compact set in $\mathcal{H}$ such that $x <_\varepsilon^{sb} K$, therefore $\mathcal{H}$ is locally compact. On the other hand $\mathcal{S} = \mathcal{H}|(0, 1)$ fails to have this property in spite of the fact that its topology, $\mathcal{S}^{tp}$ is obviously locally compact. Indeed, let us suppose indirectly that $< = <_\varepsilon^{sb}|(0, 1)$ is an order chosen in accordance with the definition, and let x be an element of $(0, 1)$ such that $1 - x < \varepsilon$. Then $x < V$ implies $x <_\varepsilon^{sb} W$ and $W \cap (0, 1) \subset V$ for some set $W \subset \mathbf{R}$. In this case $(x-\varepsilon, x+\varepsilon) \subset W$, therefore $[x, 1) \subset V$, consequently $\{[y, 1): x < y < 1\}$ is a compressed filter base in V which has no limit point in this set, that is, V cannot be compact in $\mathcal{S}$. ∎

As it is shown by (2.2), the subspaces of a syntopogenous space do not in general preserve its local compactness.

Yet there exists an important exception:

PROPOSITION 2.3. *A closed subspace of a locally compact syntopogenous space is also locally compact.*

PROOF. Denote by $[E, \mathcal{S}]$ the space in question and assume that $< \in \mathcal{S}$, $x \in E$ imply $x < K_x$ with a suitable compact set K_x in $\mathcal{S}$. If E_0 is a closed subset of E, then for any $\mathcal{S}$-compact set K, $K \cap E_0$ is also compact in $\mathcal{S}|E_0$, and for each $x \in E_0$ the inequality $x(<|E_0) K_x \cap E_0$ holds. ∎

One of the main results of this § is the verification of the fact that a locally compact S_2-syntopology is symmetrizable. This is based upon the following statement.

LEMMA 2.4. *Let $\mathcal{S}$ be a locally compact S_2-syntopology on the set E. If $<\in\mathcal{S}$, then there exists $<'\in\mathcal{S}$ such that for an arbitrary compact $K_0\subset E$,*

$$K_0 < G \quad \textit{implies} \quad K_0 <' K <'^c L <' G,$$

where K is also compact.

PROOF. Let $<$ be a member of $\mathcal{S}$ and let us choose the orders $<_1$, $<_2$, $<_3$, $<_0$, $<_1'$ and $<'$ of $\mathcal{S}$ as follows:

(a) $< \subset <_1^4$; (b) $<_1 \subset (<_2 \cdot <_2^c)^b$; (c) $<_2 \subset <_3^2$; (d) for $x\in E$ there exists a compact set K_x with $x<_0 K_x$; (e) $<_0 \subset <_1'^2$; (f) $<_1 \cup <_1' \cup <_3 \subset <'$.

Then let K_0 be a compact set for which $K_0<G$. There exist $F, J, L\subset E$ such that

(1) $$K_0 <_1 F <_1 J <_1 L <_1 G.$$

We shall prove the existence of a compact set $K\subset J$ for which $K_0<'K$. Indeed, assume $x<_1' H_x <_1' K_x$ for each $x\in K_0$. Then a finite number of these sets H_x covers K_0, thus denoting by H and K' the union of the corresponding sets H_x and K_x, we get that

(2) $$K' \text{ is compact, and } K_0\subset H<_1' K'.$$

If $K'\subset J$ then the statement is proved because in view of (f) $K_0<'K'$. Therefore suppose $K'\not\subset J$, and put $J'=K'-J$, $F'=K'-F$. If $x\in K_0$ is fixed, then owing to (b) and (1), $x<_2 A_{xy} <_3^c B_{xy} <_3^b E-y$ for $y\in F'$. In the compact topology $\mathcal{S}^{tp}|K'=\{<^0\}$ [1], (15.93) gives that a finite number of the sets $E-B_{xy}$ covers J' since $J'<^{0c}F'$ and $y<^0(E-B_{xy})\cap K'$ for $y\in F'$. Denoting by $E-B_x$ and $E-A_x$ the union of the covering sets $E-B_{xy}$ and $E-A_{xy}$, respectively, we have

$$x <_2 A_x <_3^c B_x \subset E-J' = (E-K')\cup J.$$

Putting $x<_3 C_x <_3 A_x$ for $x\in K_0$, in view of the compactness of K_0 we have sets C, A and B such that

(3) $$K_0\subset C<_2 A<_3^c B\subset (E-K')\cup J.$$

Let us denote by $\bar{A}$ the closure of A in $\mathcal{S}$. One can easily show that $K=\bar{A}\cap K'$ is compact, and from (3) $\bar{A}\subset B$ follows, thus $K\subset J$. Because of $A\subset\bar{A}$ and (2) we get $K_0<'K$.

Further if $k\in K$ and $y\in E-L$, then by (1) and (b) we have $k<_2 V_{ky} <_2^c E-y$. There exists a set V_y such that $K\subset V_y <_2^c E-y$, so that $E-L<_2 E-K$, since $<_2$ is perfect. Finally, in view of (f) and (c)

$$K_0 <' K <' L <' G. \quad \blacksquare$$

LEMMA 2.5. *Let $\mathcal{S}$ be a locally compact S_2-syntopology on E. Define, for each $<\in\mathcal{S}$, an order $<^+$ on E as follows:*

(2.5.1) *$A<^+B \Leftrightarrow B=E$, or there exists an $\mathcal{S}$-compact $K\subset E$ such that $A\subset K<^c X<B$ for some $X\subset E$.*

Then

(2.5.2) $<^+$ *is a semi-topogenous order for every* $<\in\mathscr{S}$.

(2.5.3) $<_1, <_2\in\mathscr{S}$, $<_1\subset<_2$ *imply* $<_1^+\subset<_2^+$.

(2.5.4) *For every* $<\in\mathscr{S}$ *there is an order* $<'\in\mathscr{S}$ *such that* $<^+\subset<'^{+2}$.

PROOF. (2.5.2) and (2.5.3) are clear.

(2.5.4): Suppose $<\in\mathscr{S}$, and choose $<_1\in\mathscr{S}$ for $<$ in accordance with (2.4), put $<_2\in\mathscr{S}$ for $<_1$ in the same manner, finally let $<'$ be an element of $\mathscr{S}$ with $<_1\cup<_2\subset<'$. Then $A<^+E$ implies $A<_2^+E<_2^+E$ automatically. If $A<^+B\neq E$, we have $A\subset K<^cX<B$ for some $X\subset E$ and for an $\mathscr{S}$-compact K. Now it is clear that $K<B$, therefore by (2.4) there exists a compact K_1 such that $K<_1K_1<_1^cY<_1B$, thus by (2.4) we get $A\subset K<_2^cL<_2K_1<_1^cY<_1B$. From the choice of $<'$ the inequality $A<'^{+2}B$ issues. ∎

The following theorem is a generalization of [2], (5.3.57).

THEOREM 2.6. *Any locally compact S_2-syntopology $\mathscr{S}$ is symmetrizable, namely $\mathscr{S}^+=\{<^{+s}: <\in\mathscr{S}\}$ is the coarsest of all symmetrical syntopogenous structures $\mathscr{S}_1$ such that $\mathscr{S}_1^p\sim\mathscr{S}$.*

PROOF. (2.5.2) implies that $\mathscr{S}^+$ consists of symmetrical topogenous orders. If $<_1, <_2\in\mathscr{S}$ and $<\in\mathscr{S}$, $<_1\cup<_2\subset<$, then from (2.5.3) and [1], (3.40) the relation $<_1^{+s}\cup<_2^{+s}\subset<^{+s}$ follows. Finally, by (2.5.4) and [1], (3.53) for any $<\in\mathscr{S}$ there exists $<'\in\mathscr{S}$ such that $<^{+s}\subset<'^{+s2}$. These give that $\mathscr{S}^+$ is a symmetrical syntopogenous structure.

Every point $x\in E$ is compact in $\mathscr{S}$, consequently if $<\in\mathscr{S}$, then an order $<'\in\mathscr{S}$ can be found such that $x<B$ implies $x<'^+B$ (see (2.4)). From this one can deduce $<\subset<'^{+p}\subset<'^{+sp}$, hence $\mathscr{S}<\mathscr{S}^{+p}$. On the other hand, $<^+\subset<$, $<^{+c}\subset<$ are clear for each $<\in\mathscr{S}$, thus $<^{+s}\subset<$, and $<^{+sp}\subset<$. This means that $\mathscr{S}^{+p}<\mathscr{S}$ is also valid. Summing up, we have $\mathscr{S}^{+p}\sim\mathscr{S}$.

Assume that $\mathscr{S}_1$ is a symmetrical syntopogenous structure such that $\mathscr{S}_1^p\sim\mathscr{S}$, $<\in\mathscr{S}$ is arbitrary, $<_1\in\mathscr{S}_1$, $<\subset<_1^p$, and for $<_2\in\mathscr{S}_1$: $<_1\subset<_2^2$. If $A<^+E$, then obviously $A<_2E$. Suppose $A<^+B\neq E$, then $A\subset K<^cX<B$ for some compact set K. Assume $x<_2C_x<_2B$ for $x\in K$, in this case from the compactness of K we get $K\subset C<_2B$, thus $A<_2B$. Consequently, $<^+\subset<_2$, and in view of the symmetricity of $<_2$, the inequality $<^{+s}\subset<_2$ holds. Thus we have $\mathscr{S}^+<\mathscr{S}_1$. ∎

Next we study the properties (S_1) and (S_2) in compact spaces.

PROPOSITION 2.7. *If $\mathscr{S}$ is a compact syntopogenous structure with property* (S_1), *then* $\mathscr{S}^p<\mathscr{S}^{cp}$ (cf. (1.15)).

PROOF. Put $<\in\mathscr{S}$, and choose the orders $<_1$, $<_2$, $<_3$ of $\mathscr{S}$ as follows: $<\subset<_1^2$, $<_1^b\subset<_2^{bc}$, $<_2\subset<_3^2$. If $x<B$, then $x<_1C<_1B$. $y<_2E-x$ holds for any $y\in E-C$ and thus $y<_3V_y<_3E-x$. Owing to theorem (15.93) of [1], a finite number of the sets V_y covers $E-B$, from this $E-B<_3E-x$, i.e. $x<_3^cB$. ∎

As a generalization of theorem (5.3.22) of [2] we can state:

THEOREM 2.8. *A compact syntopogenous space with property* (S_2) *is normal.*

PROOF. In a compact S_2-space $[E, \mathcal{S}]$, for an arbitrary $< \in \mathcal{S}$, let us choose the orders $<_1$, $<_2$, and $<_3$ of $\mathcal{S}$ as follows: $< \subset <_1^2$, $<_1^b \subset (<_2 \cdot <_2^c)^b$, $<_2 \subset <_3^2$. Assume $A <^c C < B$. Then $A <^c C <_1 X <_1 B$. If $x \in C$, $y \in E-X$, then $X <_2 C_{xy} <_2^c <_2^c E-y$ and from this

$$x <_3 A_{xy} <_3 C_{xy} <_3^c B_{xy} <_3^c E-y.$$

Put $\mathcal{S}^{tp} = \{<^0\}$. $A <^{0c} C$ and $x <^0 A_{xy}$ imply the existence of sets A_y, C_y, B_y such that

$$A \subset A_y <_3 C_y <_3^c B_y <_3^c E-y$$

(see [1], (15.93)). In view of $E-B <^{0c} E-X$, $y <^0 E-B_y$ $(y \in E-X)$ a finite number of the sets $E = B_y$ covers $E-B$. Denoting by $E-B'$, $E-A'$, $E-C'$ the union of the covering sets $E-B_y$, $E-A_y$, $E-C_y$, we have

$$A \subset A' <_3 C' <_3^c B' \subset B. \quad \blacksquare$$

COROLLARY 2.9. *For a compact syntopogenous space $[E, \mathcal{S}]$, $\mathcal{S}^p$ is symmetrizable iff $\mathcal{S}$ is an S_2-structure. In this case (up to equivalence) $\mathcal{S}^\triangledown$ is unique among the symmetrical syntopogenous structures $\mathcal{S}_0$ on E such that $\mathcal{S}_0^p \sim \mathcal{S}^p$.*

PROOF. If $\mathcal{S}^p$ is symmetrizable, then it is regular, consequently $\mathcal{S}^p$, and at the same time $\mathcal{S}$ has property (S_2) (see (1.11) and (1.5)). Conversely, if $\mathcal{S}$ is an S_2-structure, then by (2.8) it is normal, and in view of (2.7) $\mathcal{S}^p < \mathcal{S}^{cp}$. Applying (1.17) $\mathcal{S}^p \sim \mathcal{S}^{\triangledown p}$ can be deduced, which implies that $\mathcal{S}^p$ is symmetrizable. If this is true, then $\mathcal{S}^\triangledown$ is the unique symmetrical syntopogenous structure inducing $\mathcal{S}^p$ by Lemma 8 of [3]. $\blacksquare$

3. Some basic concepts of syntopogenous extension theory

The definitions and results presented here were discussed in [6] in detail (cf. also [4] and [8]).

In a syntopogenous space $[E, \mathcal{S}]$ a filter base $\mathfrak{r}$ is said to be *round* if $R \in \mathfrak{r}$ implies $R_1 < R$ for some $< \in \mathcal{S}$ and $R_1 \in \mathfrak{r}$. For any filter base $\mathfrak{r}$ in E

$$\mathcal{S}(\mathfrak{r}) = \{V \subset E\colon\ R < V \text{ for some } < \in \mathcal{S} \text{ and } R \in \mathfrak{r}\}$$

is a round filter in $[E, \mathcal{S}]$, in particular, denoting by $\mathcal{S}(x)$ the filter $\mathcal{S}(\{\{x\}\})$, we get the *neighbourhood filter* of the point $x \in E$.

For two systems $\mathfrak{A}$ and $\mathfrak{B}$ of subsets of E, another system can be constructed as follows:

$$\mathfrak{A}(\cap)\mathfrak{B} = \{A \cap B\colon\ A \in \mathfrak{A},\ B \in \mathfrak{B}\}.$$

A syntopogenous space $[E', \mathcal{S}']$ will be called an *extension* of the syntopogenous space $[E, \mathcal{S}]$ if E is a dense subset of $[E', \mathcal{S}']$ and $\mathcal{S} \sim \mathcal{S}'|E$. At the same time $\mathcal{S}'$ will be said to be an *extension* of $\mathcal{S}$ on E'. In this case the filters

$$\mathfrak{s}(x) = \mathcal{S}'(x)(\cap)\{E\} \qquad (x \in E')$$

(which are called the *trace filters* of this extension) are round filters in $[E, \mathcal{S}]$, in particular, $\mathfrak{s}(x) = \mathcal{S}(x)$ for each $x \in E$.

Conversely, if $[E, \mathscr{S}]$ is an arbitrary syntopogenous space, $E \subset E'$, and for any $x \in E'$ an $\mathscr{S}$-round filter $\mathfrak{s}(x)$ is given so that $\mathfrak{s}(x) = \mathscr{S}(x)$ for $x \in E$, then there exists an extension of $\mathscr{S}$ on E' the trace filters of which agree with the filters $\mathfrak{s}(x)$ for every $x \in E'$. In fact, suppose

$$s(A) = \{x \in E' \colon A \in \mathfrak{s}(x)\} \qquad (A \subset E).$$

If $<$ is a semi-topogenous order on E and $A', B' \subset E'$, define

$$A' <' B' \quad \text{iff} \quad A' \subset s(A),\ s(B) \subset B' \quad \text{for some} \quad A < B.$$

Then putting

$$s(<) = <'^{q},$$

the order family

$$s(\mathscr{S}) = \{s(<) \colon < \in \mathscr{S}\}$$

is a syntopogenous structure on E' satisfying the required conditions.

Iff $\mathscr{S} = \mathscr{T}$ is a topology on E, then the system

$$\{s(V) \colon V \text{ is open in } \mathscr{T}\}$$

forms a base for a topology $\mathscr{T}'$ on E', which is called the *strict extension* of $\mathscr{T}$ on E' corresponding to the filters $\mathfrak{s}(x)$ $(x \in E')$. In general, for an arbitrary $\mathscr{S}$, $s(\mathscr{S})^{tp}$ is a strict extension of $\mathscr{S}^{tp}$.

Let $[E', \mathscr{S}']$ and $[E'', \mathscr{S}'']$ be two extensions of the syntopogenous space $[E, \mathscr{S}]$. Similarly to the terminology of the symmetrizable compactifications (cf. [3], def. 4) we shall say that $[E'', \mathscr{S}'']$ is a *coarser extension* of $[E, \mathscr{S}]$ than $[E', \mathscr{S}']$ (or $[E', \mathscr{S}']$ is a *finer* one than $[E'', \mathscr{S}'']$) iff there exists an $(\mathscr{S}', \mathscr{S}'')$-continuous surjection h of E' onto E'' such that $x = h(x)$ for every $x \in E$. If h is an $(\mathscr{S}', \mathscr{S}'')$-isomorphism, then $[E', \mathscr{S}']$ will be said to be *equivalent* to $[E'', \mathscr{S}'']$.

4. One-point compactifications

It is well-known that P. Alexandroff was the first to compactify a locally compact Hausdorff space by adding one ideal point. A generalization of Alexandroff's one-point compactification can be found in [2], ch. 6.1. For an arbitrary non-compact topological space $[E, \mathscr{T}]$, this is defined as a strict extension of $\mathscr{T}$ on $E \cup \{p\}$ having the trace filter of the "imaginary" point p in the form

(4.1) $$\mathfrak{s}(p) = \{X \subset E \colon V \subset X,\ E - V \text{ is compact closed in } \mathscr{T}\}.$$

PROPOSITION 4.2. *Let $[E, \mathscr{T}]$ be a non-compact topological space, $E' = E \cup \{p\}$, $\mathscr{T}'$ be a topology on E' which is an extension of $\mathscr{T}$ with the trace filters $\mathfrak{s}(x)$ for $x \in E'$. In this case $\mathscr{T}'$ is compact iff $E - V$ is compact in $\mathscr{T}$ for any $\mathscr{T}$-open $V \in \mathfrak{s}(p)$.*

PROOF. Suppose that $\mathscr{T}'$ is compact and let $\{V_i \colon i \in I\}$ be a $\mathscr{T}$-open covering of the set $E - V$, where V is a $\mathscr{T}$-open member of $\mathfrak{s}(p)$. For any $i \in I$ there exists a $\mathscr{T}'$-open set V_i' in E' such that $V_i' \cap E = V_i$, and there is also a $\mathscr{T}'$-open set V' for which $p \in V'$ and $V' \cap E = V$. The system $\{V', V_i' \colon i \in I\}$ is a $\mathscr{T}'$-open covering of E', therefore a finite number of indices of I can be found such that $\{V', V_{i_1}', \ldots$ $\ldots, V_{i_n}'\}$ covers E'. Then $E - V \subset E' - V' \subset \bigcup_{k=1}^{n} V_{i_k}'$, consequently $E - V \subset \bigcup_{k=1}^{n} V_{i_k}$.

Conversely, suppose that $E-V$ is compact for each $\mathcal{T}$-open $V\in\mathfrak{s}(p)$, and let $\{V_i'\colon i\in I\}$ be a $\mathcal{T}'$-open covering of E'. If $p\in V_{i_0}'$, then $\{V_i'\cap E\colon i\in I\}$ is a $\mathcal{T}$-open covering of the set $E-(V_{i_0}'\cap E)$, which is compact in $\mathcal{T}$ since $V_{i_0}'\cap E$ is a $\mathcal{T}$-open member of $\mathfrak{s}(p)$. If a finite subsystem $\{V_{i_1}'\cap E, \ldots, V_{i_n}'\cap E\}$ covers $E-(V_{i_0}'\cap E)$, then $\{V_{i_0}', V_{i_1}', \ldots, V_{i_n}'\}$ is a finite covering of E'. ∎

In view of (4.2) we can prove Lemma (6.1.22) of [2] independently of whether $\mathcal{T}'$ is strict or not.

LEMMA 4.3. *If under the conditions of* (4.2) $\mathcal{T}'$ *is compact, then* $\mathfrak{s}(x)\subset\mathfrak{s}(p)$ *does not hold for any* $x\in E$.

PROOF. Let us assume that $x\in E$ is a point such that $\mathfrak{s}(x)\subset\mathfrak{s}(p)$, and $\{V_i\colon i\in I\}$ be an arbitrary $\mathcal{T}$-open covering of E. Then $x\in V_{i_0}$ for some $i_0\in I$, hence $V_{i_0}\in\mathfrak{s}(x)\subset\mathfrak{s}(p)$. By (4.2) $E-V_{i_0}$ is compact in $\mathcal{T}$, therefore a finite number of the sets V_i covers it. Adding V_{i_0} to this finite covering of $E-V_{i_0}$, we get a finite $\mathcal{T}$-open covering of E chosen from the original system. This gives that $\mathcal{T}$ is compact, which is impossible. ∎

PROPOSITION 4.4. *Under the conditions of* (4.2), *if* $\mathcal{T}'$ *is compact, then it is a strict extension of* $\mathcal{T}$ *on* E', *and* $\{p\}$ *is closed in* $\mathcal{T}'$.

PROOF. We shall show that if V is a $\mathcal{T}'$-neighbourhood of $x\in E'$, then there exists a $\mathcal{T}$-open set $V_0\in\mathfrak{s}(x)$ such that $s(V_0)\subset V$. First of all, consider a point x in E. Then $V\cap E\in\mathfrak{s}(x)$, and if $V_1\in\mathfrak{s}(x)$ is a $\mathcal{T}$-open set with the property $V_1\notin\mathfrak{s}(p)$ (see (4.3)), we have $s(V_0)\subset V$ for $V_0=V_1\cap V\cap E\in\mathfrak{s}(x)$. In fact, $V_0\in\mathfrak{s}(y)$, $y\in E'$ imply $y\neq p$ (i.e. $y\in E$), else $V_1\in\mathfrak{s}(p)$ would be true, therefore V_0 is a $\mathcal{T}$-neighbourhood of y, consequently $y\in V$.

On the other hand, suppose $x=p$, then $V_0=V\cap E$ is a member of $\mathfrak{s}(p)$ for which $s(V_0)\subset V$ is obvious.

$\{p\}$ is indeed closed in $\mathcal{T}'$, because for any $x\in E$ there is a set $V_x\in\mathfrak{s}(x)$ such that $V_x\notin\mathfrak{s}(p)$, so that $s(V_x)$ is a neighbourhood of x which does not contain p. ∎

LEMMA 4.5. *Let* $[E, \mathcal{T}]$ *be a non-compact topological space and* $E'=E\cup\{p\}$. *Suppose that the topologies* $\mathcal{T}_1'$ *and* $\mathcal{T}_2'$ *are extensions of* $\mathcal{T}$ *on* E' *with the trace filters* $\mathfrak{s}_1(x)$ *and* $\mathfrak{s}_2(x)$, *respectively, for* $x\in E'$, *further let* $\mathcal{T}_2'$ *be compact. Then* $\mathcal{T}_1'<\mathcal{T}_2'$ *iff* $\mathfrak{s}_1(p)\subset\mathfrak{s}_2(p)$.

PROOF. The condition is obviously necessary. In order to verify its sufficiency, suppose $\mathfrak{s}_1(p)\subset\mathfrak{s}_2(p)$, and let V be a $\mathcal{T}_1'$-open set in E'. Assume

$$s_i(A) = \{x\in E'\colon A\in\mathfrak{s}_i(x)\} \qquad (A\subset E,\ i=1, 2).$$

$V\cap E\in\mathfrak{s}_1(x)\subset\mathfrak{s}_2(x)$ for each $x\in V$, therefore if $p\in V$, then $s_2(V\cap E)\subset(V\cap E)\cup\{p\}=$ $=V$. On the other hand, if $V\subset E$, then for every $x\in V$ there exists $W_x\in\mathfrak{s}_2(x)$ such that $W_x\notin\mathfrak{s}_2(p)$ (see (4.3)), consequently $U_x=V\cap W_x\in\mathfrak{s}_2(x)$ and $s_2(U_x)\subset V$. These show that V is open in $\mathcal{T}_2'$, too. ∎

In a non-compact syntopogenous space $[E, \mathcal{S}]$ the complements of the compact subsets form a filter base, which generates a filter in E denoted by $\mathfrak{c}$. In view of (4.2)—(4.5) we can give a complete summary of the compact extensions of a topological space with one ideal point as follows:

THEOREM 4.6. *Let $[E, \mathscr{T}]$ be a non-compact topological space and $E' = E \cup \{p\}$. The compact extensions $\mathscr{T}' = \mathscr{T}''^{tp}$ of $\mathscr{T}$ on E' are exactly those strict extensions which have a trace filter for p contained in $\mathfrak{c}$. For two such extensions $\mathscr{T}_1'$ and $\mathscr{T}_2'$, $\mathscr{T}_1' < \mathscr{T}_2'$ iff $\mathfrak{s}_1(p) \subset \mathfrak{s}_2(p)$ holds for the corresponding trace filters.* ∎

Let us consider an arbitrary non-compact syntopogenous space $[E, \mathscr{S}]$. We shall say that $[E', \mathscr{S}']$ is a *one-point compactification* of $[E, \mathscr{S}]$ if $[E', \mathscr{S}']$ is a compact extension of $[E, \mathscr{S}]$ such that E' results from E by adding one ideal point p.

THEOREM 4.7. *Let $[E, \mathscr{S}]$ be a non-compact syntopogenous space, $E' = E \cup \{p\}$, and $\mathscr{S}'$ be an extension of $\mathscr{S}$ on E' with the trace filter $\mathfrak{s}(x)$ for $x \in E'$. Then $[E', \mathscr{S}']$ is a one-point compactification of $[E, \mathscr{S}]$ iff $\mathscr{S}'^{tp}$ is a strict extension of $\mathscr{S}^{tp}$ and $\mathfrak{s}(p) \subset \subset \mathscr{S}(\mathfrak{c})$.*

PROOF. $\mathscr{S}'$ is compact iff so is $\mathscr{S}'^{tp}$ ([1], (15.78), (15.79)) and it is proved that this holds if and only if $\mathscr{S}'^{tp}$ is a strict extension of $\mathscr{S}'^{tp}|E = \mathscr{S}^{tp}$ and $\mathfrak{s}(p) \subset \mathfrak{c}$. But this latter condition is equivalent to $\mathfrak{s}(p) \subset \mathscr{S}(\mathfrak{c})$, provided $\mathscr{S}'$ is an extension of $\mathscr{S}$. In fact, $\mathscr{S}(\mathfrak{c}) \subset \mathfrak{c}$, and on the other hand, because of the $\mathscr{S}$-roundness of $\mathfrak{s}(p)$, the inclusion $\mathfrak{s}(p) \subset \mathfrak{c}$ implies $\mathfrak{s}(p) \subset \mathscr{S}(\mathfrak{c})$. ∎

THEOREM 4.8. *Any non-compact syntopogenous space has a finest one-point compactification. This can be topogenous or perfect, provided that so is the original space.*

PROOF. Let $[E, \mathscr{S}]$ be the space in question, $E' = E \cup \{p\}$, $\mathfrak{s}(p) = \mathscr{S}(\mathfrak{c})$ and $\mathfrak{s}(x) = \mathscr{S}(x)$ for $x \in E$. Suppose that $\mathscr{T}'$ is the strict extension of $\mathscr{S}^{tp}$ on E' corresponding to the filters $\mathfrak{s}(x)$ $(x \in E')$. Denoting by $\mathscr{S}'$ the syntopogenous structure on E' constructed in Theorem 2.1 of [6] for $\mathscr{S}$ and $\mathscr{T}'$, we have $\mathscr{S}'|E \sim \mathscr{S}$ and $\mathscr{S}'^{tp} = \mathscr{T}'$. Thus $[E', \mathscr{S}']$ is a one-point compactification of $\mathscr{S}$ by (4.7). If $[E'', \mathscr{S}'']$ is another one-point compactification of $[E, \mathscr{S}]$ with the trace filter $\mathfrak{s}'(p')$ for $\{p'\} = E'' - E$, then define the mapping $h\colon E' \to E''$ so that $h(x) = x$ for $x \in E$ and $h(p) = p'$. It can be easily seen that $[E', h^{-1}(\mathscr{S}'')]$ is a one-point compactification of $[E, \mathscr{S}]$ with the trace filter $\mathfrak{s}'(p')$ for p. Then by (4.7) $\mathfrak{s}'(p') \subset \mathscr{S}(\mathfrak{c}) = \mathfrak{s}(p)$, therefore $h^{-1}(\mathscr{S}'')^{tp} < \mathscr{T}'$ (see (4.6)). In view of Theorem 2.2 of [6] we get $h^{-1}(\mathscr{S}'') < \mathscr{S}'$, consequently $[E', \mathscr{S}']$ is finer than $[E'', \mathscr{S}'']$. Finally, from [6], 2.1 it follows that $\mathscr{S}'$ is topogenous or perfect, provided so is $\mathscr{S}$. ∎

For the study of the existence of symmetrical one-point compactifications let us mention the following simple lemma, which can be deduced immediately from [1], (15.47) and [2], (6.3.10) (or [2], (6.3.14)):

LEMMA 4.9. *If $\mathscr{S}$ is a symmetrical (and biperfect) syntopogenous structure, then $\mathscr{S}(\mathfrak{r})$ is a compressed (Cauchy) filter for any compressed (Cauchy) filter base $\mathfrak{r}$ in $\mathscr{S}$.* ∎

THEOREM 4.10. *A non-compact symmetrical syntopogenous space $[E, \mathscr{S}]$ has a symmetrical one-point compactification iff $< \in \mathscr{S}$, $A < B$ imply either $A \subset K$ or $E - B \subset K$ for some $\mathscr{S}$-compact set K.*

PROOF. If there exists a symmetrical one-point compactification for $\mathscr{S}$ on a set $E' = E \cup \{p\}$, then by [6], 4.1 we get that $\mathfrak{s}(p)$ is compressed, so that because of $\mathfrak{s}(p) \subset \mathscr{S}(\mathfrak{c}) \subset \mathfrak{c}$ the filter $\mathfrak{c}$ is also compressed (see (4.7) and [1], (15.48)). Conversely, if $\mathfrak{c}$ is compressed, then $\mathscr{S}(\mathfrak{c})$ is also of this kind by (4.9). Putting $\mathfrak{s}(p) = \mathscr{S}(\mathfrak{c})$, we get $s(\mathscr{S})^s$ as a symmetrical one-point compactification of $\mathscr{S}$ on E', because $s(\mathscr{S})^{stp}$ is a strict extension of $\mathscr{S}^{tp}$ with the trace filter $\mathfrak{s}(p)$ contained in $\mathscr{S}(\mathfrak{c})$ (cf. [6], 4.2 and (4.7)). Thus we saw that the existence of a symmetrical one-point compactification is equivalent to the fact that $\mathfrak{c}$ is compressed in $\mathscr{S}$. This latter condition is satisfied iff $A < B$, $< \in \mathscr{S}$ imply $A \cap (E-K) = \emptyset$ or $(E-B) \cap (E-K) = \emptyset$ for some compact set K, that is, either A or $E-B$ is contained in an $\mathscr{S}$-compact subset of E. ∎

Analogously, we can state:

THEOREM 4.11. *A non-compact symmetrical syntopological space $[E, \mathscr{S}]$ has a symmetrical biperfect one-point compactification iff for any $< \in \mathscr{S}$ there exists an $\mathscr{S}$-compact set K such that $A < B$ implies either $A \subset K$ or $E-B \subset K$.*

PROOF. We have to write the term "Cauchy filter" instead of "compressed filter" and to use Lemma 5.2 of [6]. ∎

If $[E', \mathscr{S}']$ is a one-point compactification of the syntopogenous space $[E, \mathscr{S}]$, then $\mathscr{S}'$ is relatively separated with respect to E by (4.4), consequently, two symmetrical one-point compactifications of $[E, \mathscr{S}]$ are always equivalent extensions, because these are double compactifications of the space in question (cf. [3]). However, as the following example shows, a symmetrical (and biperfect) space in general has several one-point compactifications even if it possesses the property described in (4.10) (or 4.11)).

EXAMPLE 4.12. Suppose $E = [0, 1)$, $\mathscr{S} = \mathscr{I}^{sb}|E$. With the notations $E' = [0, 1]$ and $\mathscr{S}' = \mathscr{I}^{sb}|E'$, it is obvious that $\mathscr{S}'$ is a symmetrical biperfect one-point compactification of $\mathscr{S}$, namely $p = 1$. Let us denote by $\mathscr{S}''$ the finest syntopogenous structure on E' compatible with the pair $(\mathscr{S}, \mathscr{S}''^{tp})$ (see [6], Theorem 2.1, 2.2), which is also a one-point compactification of $\mathscr{S}$ on E' by (4.7). It is easy to verify that $\mathscr{S}'$ and $\mathscr{S}''$ are not equivalent. In fact, for any $<' \in \mathscr{S}'$, $E <' E$ cannot hold, at the same time E is an $\mathscr{S}''^{tp}$-open set and $E < E$ for every $< \in \mathscr{S}$, hence $E <'' E$ is fulfilled by each $<'' \in \mathscr{S}''$ (cf. [6], (2.1.1)). ∎

Further we shall study the one-point compactifications of locally compact S_2-syntopologies.

In a syntopological space $[E, \mathscr{S}]$ a set E_0 will be called *strongly open* if $E_0 < E_0$ for some $< \in \mathscr{S}$. It is clear that every strongly open set is open.

PROPOSITION 4.13. *In a locally compact syntopological space with property (S_2) any strongly open subspace is also locally compact.*

PROOF. Let $\mathscr{S}$ be a locally compact S_2-syntopology on E, and suppose $E_0 < E_0 \subset E$ for some $< \in \mathscr{S}$. If the order $<_1 \in \mathscr{S}$ is chosen in accordance with (2.4), then for any $x \in E_0$, we have $x <_1 K_x \subset E_0$, where K_x is $\mathscr{S}$-compact. K_x is also $(\mathscr{S}|E_0)$-compact, and $x(<_1|E_0)K_x$, thus $\mathscr{S}|E_0$ is locally compact, too. ∎

This implies that if a syntopological space $[E, \mathscr{S}]$ has a symmetrizable compactification in which E is strongly open, then $\mathscr{S}$ is necessarily a locally compact S_2-syntopology. In order to verify that a locally compact S_2-syntopology always has such a compactification, and what is more, this can be a one-point compactification, let us introduce the following notion.

A syntopological $[E', \mathscr{S}']$ will be called an *Alexandroff-type compactification* of the non-compact syntopological space $[E, \mathscr{S}]$ iff $E' = E \cup \{p\}$ and $\mathscr{S}' \sim s(\mathscr{S})^p$, where $\mathfrak{s}(p) = \mathscr{S}(\mathfrak{c})$ and $\mathfrak{s}(x) = \mathscr{S}(x)$ for $x \in E$. Since in this case $\mathscr{S}'^{tp} = s(\mathscr{S})^{ptp} = s(\mathscr{S})^{tp}$ is a strict extension of $\mathscr{S}^{tp}$ corresponding to the filters $\mathfrak{s}(x)$ $(x \in E')$, such an extension is in fact a one-point compactification of $[E, \mathscr{S}]$ (cf. § 3 and (4.7)). It is also easy to see that two Alexandroff-type compactifications of a syntopological space are always its equivalent extensions.

LEMMA 4.14. *Let $[E, \mathscr{S}]$ be a non-compact locally compact syntopological space with property (S_2), and let $[E', \mathscr{S}']$ denote an Alexandroff-type compactification of this space. Then there exists an order $<'_0 \in \mathscr{S}'$ such that for every $x \in E$ the inequality $p(<'_0 \cdot <'^c_0)\, E' - x$ is valid.*

PROOF. Assume $x < K_x$ $(< \in \mathscr{S}$ is fixed) for any $x \in E$ with a compact set K_x, and let $<_1$ be an order of $\mathscr{S}$ satisfying the condition of (2.4). Suppose $<_2 \in \mathscr{S}$, $<_1 \subset <_2^2$ and $s(<_2)^p \subset <'_0 \in \mathscr{S}'$. Because of the compactness of any point $x \in E$ there exist compact sets K'_x such that $x <_1 K'_x <_1^c K_x$, and from this

$$E - K_x <_2 C_x <_2 E - K'_x <_2^c V_x <_2^c E - x.$$

Then $p \in s(C_x) <'_0 s(E - K'_x) \subset E' - s(K'_x) <'^c_0 E' - s(E - V_x) \subset E' - x$. This proves the lemma. ▮

In possession of (4.14) we can complete (4.13) as follows.

PROPOSITION 4.15. *Any Alexandroff-type compactification of a non-compact locally compact syntopological S_2-space $[E, \mathscr{S}]$ is also an S_2-space in which the set E is strongly open.*

PROOF. Under the notations of (4.14), the inequality $x <'_0 E' - p = E$ holds for each $x \in E$, therefore E is strongly open. Trivially, for any $<' \in \mathscr{S}'$ there exists $<'_1 \in \mathscr{S}'$ such that $x, y \in E$, $x <' E' - y$ imply $x(<'_1 \cdot <'^c_1) E' - y$, as seen from property (S_2) of $\mathscr{S}$. If $<'_2 \in \mathscr{S}'$ for which $<'_0 \cup <'_1 \subset <'_2$, then by (4.14) $<' \subset (<'_2 \cdot <'^c_2)^b$, hence $\mathscr{S}'$ is an S_2-syntopology. ▮

THEOREM 4.16. *An Alexandroff-type compactification of a non-compact locally compact syntopological S_2-space is one of the coarsest symmetrizable compactifications of this space.*

PROOF. Let $[E', \mathscr{S}']$ denote an Alexandroff-type compactification of the non-compact locally compact syntopological S_2-space $[E, \mathscr{S}]$. This compactification is obviously symmetrizable (cf. (2.9) and (1.5)). By Theorem 14 of [3] it will be sufficient to verify that denoting by $\mathscr{S}'_1$ the unique symmetrical structure on E' such that $\mathscr{S}'^p_1 \sim \mathscr{S}'$, the restriction $\mathscr{S}'_1 | E = \mathscr{S}_1$ is equivalent to $\mathscr{S}^+$ (see (2.6)). Owing to (2.6) and $\mathscr{S}^p_1 \sim \mathscr{S}$, the inequality $\mathscr{S}^+ < \mathscr{S}_1$ is clear. Conversely, suppose $<_1 \in \mathscr{S}_1$, $<_2 \in \mathscr{S}_1$, $<_1 \subset <_2^5$, finally assume, for a suitable $< \in \mathscr{S}$, $<_2^p \subset <$. Then we have $<_2 \subset <$, and from the symmetry of $<_2$ the relation $<_2 \subset <^c$ also follows. It can be easily seen that $\mathscr{S}'_1$ is a symmetrical one-point compactification of $\mathscr{S}_1$, therefore it fulfils the condition given in (4.10). Suppose $A <_1 B$, then

$$A <_2 C <_2 D <_2 F <_2 G <_2 B$$

for suitable $C, D, F, G \subset E$. On the basis of (4.10) $D \subset K$ or $E-F \subset K$ for some $\mathscr{S}_1$-compact set K (which is at the same time $\mathscr{S}$-compact). Denoting by $\bar{X}$ the $\mathscr{S}$-closure of the set $X \subset E$, in the first case $A \subset \bar{A} \subset C <^c D < F \subset B$, and because of $\bar{A} \subset K$, the set $\bar{A}$ is $\mathscr{S}$-compact. In the second one $E-B \subset \overline{E-B} \subset E-G <^c E-F < < E-D \subset E-A$, and from $\overline{E-B} \subset K$ the compactness of $\overline{E-B}$ can be deduced. We got that $A <_1 B$ implies $A <^+ B$ or $A <^{+c} B$, so that $<_1 \mathbf{\subset} <^+ s$, consequently $\mathscr{S}_1 \sim \mathscr{S}^+$, which was to be proved. ▮

On the other hand, we have the following remarkable result.

THEOREM 4.17. *An Alexandroff-type compactification of a non-compact locally compact syntopological S_2-space is one of the finest one-point compactifications of the space in question.*

PROOF. Suppose that $[E', \mathscr{S}']$ is an Alexandroff-type compactification and $[E_1', \mathscr{S}_1']$ is an arbitrary one-point compactification of the non-compact locally compact syntopological S_2-space $[E, \mathscr{S}]$. If $E' = E \cup \{p\}$ and $E_1' = E \cup \{p_1\}$, then let the mapping $h\colon E' \to E_1'$ be defined so that $h(x) = x$ for $x \in E$ and $h(p) = p_1$. It is obvious that $\mathscr{S}'' = h^{-1}(\mathscr{S}_1')$ is another one-point compactification of $\mathscr{S}$ on E'. We have to show the $(\mathscr{S}', \mathscr{S}_1')$-continuity of h, i.e. $\mathscr{S}'' \mathbf{<} \mathscr{S}'$. Suppose $<' \in \mathscr{S}''$, $<' \mathbf{\subset} <_1'^3$, $<_1' \in \mathscr{S}''$ $< \in \mathscr{S}$, $<_1' | E \mathbf{\subset} <$. Let us choose the order $<_1 \in \mathscr{S}$ for $<$ in accordance with (2.4), finally assume $<_2 \in \mathscr{S}$, $<_1 \mathbf{\subset} <_2^2$, $<_3 \in \mathscr{S}$, $< \mathbf{\cup} <_2 \mathbf{\subset} <_3$. The inequality $<' \mathbf{\subset} s(<_3)^p$ will be verified. In fact, if $p <' B$ for some $B \subset E'$, then there exist $C, D \subset E'$ such that $p <_1' C <_1' D <_1' B$, and from this $C \cap E < D \cap E$ follows. Denoting by $\mathfrak{s}'(p)$ the trace of the neighbourhood filter of p in $\mathscr{S}''$, we get $C \cap E \in \mathfrak{s}'(p) \subset \mathscr{S}(\mathfrak{c}) = \mathfrak{s}(p)$ (see (4.7) and the definition of Alexandroff-type compactification), thus $p \in s(C \cap E)$. Trivially, $s(D \cap E) \cap E \subset B \cap E$, therefore $s(D \cap E) \subset \subset (B \cap E) \cup \{p\} = B$, hence $ps(<_3)B$. Now suppose that $x \in E$ and $x <' B$. Then $x <_1' B$ is also true, therefore $x < B \cap E$. One can find an $\mathscr{S}$-compact set K such that $x <_1 K \subset B \cap E$, because $\{x\}$ is compact in $\mathscr{S}$, too (see (2.4)). If $x <_2 V <_2 K$, then

$$x \in s(V)\, s(<_3)\, s(K) \subset B \cap E \subset B.$$

Indeed, $p \in s(K)$ cannot be valid because in that case $E - K_1 \subset K$, i.e. $E = K \cup K_1$ for some compact set K_1, which would contradict the fact that $[E, \mathscr{S}]$ is not compact.

In possession of the above result we can state $\mathscr{S}'' \mathbf{<} s(\mathscr{S})^p$, hence h is an $(\mathscr{S}', \mathscr{S}_1')$-continuous surjection, that is, $[E', \mathscr{S}']$ is a finer extension of $[E, \mathscr{S}]$ than $[E_1', \mathscr{S}_1']$. ▮

As is well-known (Theorem 14 of [3]), two symmetrizable compactifications of a symmetrizable syntopological space are equivalent iff each of them is finer than the other one. In view of this, from (4.16) and (4.17) the following result issues:

COROLLARY 4.18. *Up to equivalence of extensions, a non-compact locally compact syntopological S_2-space has a unique symmetrizable one-point compactification, which is an Alexandroff-type compactification.* ▮

5. Simple compactifications

O. Frink [7] introduced the notion of a Wallman-type compactification of a "classical" topological space, which is a generalization of Wallman's original method for the construction of a compact T_1-extension of T_1-spaces. The separation properties of Wallman-type compactifications were studied by several authors, their results are summarized and generalized in Chapter 6.1.e of [2]. In [4] Á. Császár pointed out the close connection existing between the Wallman compactification and the double compactification of topological spaces.

A *double compactification* of a syntopogenous space $[E, \mathscr{S}]$ is a doubly compact syntopogenous space $[E^*, \mathscr{S}^*]$ such that E is $\mathscr{S}^{*s}$-dense in E^*, $\mathscr{S}^*$ is relatively separated with respect to E and $\mathscr{S}^*|E \sim \mathscr{S}$ (see [3]). It is known ([3] and [8], § 4) that an arbitrary subspace $[E', \mathscr{S}']$ of $[E^*, \mathscr{S}^*]$ containing E can be constructed as follows:

Assume $\mathfrak{f}(x) = \mathscr{S}^s(x)$ for $x \in E$, and let $x \leftrightarrow \mathfrak{f}(x)$ be a one-to-one correspondence between the points $x \in E' - E$ and a family of non-convergent round compressed filters in $\mathscr{S}^s$ (in the case of $E' = E^*$ this family consists of all such filters). Supposing $\mathfrak{s}(x) = \mathscr{S}(\mathfrak{f}(x))$ for $x \in E'$, we have $\mathscr{S}' \sim s(\mathscr{S})$.

First of all we prove two lemmas.

Lemma 5.1 (cf. [8], (4.6.4)). *Under the conditions mentioned above*

(5.1.1) $$< \in \mathscr{S}, \quad A_i < B_i \qquad (1 \leq i \leq n)$$

imply $s\left(\bigcup_{i=1}^{n} A_i\right) \subset \bigcup_{i=1}^{n} s(B_i)$ *for any natural number* n.

(5.1.2) *If* $<_1, <_2 \in \mathscr{S}$ *and* $A <_1^c C <_2 B$, *then* $E' - s(E-A) \subset s(B)$.

Proof. (5.1.1): Suppose $x \in s\left(\bigcup_{i=1}^{n} A_i\right)$. From $\bigcup_{i=1}^{n} A_i \in \mathfrak{s}(x)$, $\mathfrak{s}(x) \subset \mathfrak{f}(x)$ we can conclude $\emptyset \notin \{A_{i_0}\}(\cap)\mathfrak{f}(x)$ for some index i_0. In fact, $\emptyset \in \{A_i\}(\cap)\mathfrak{f}(x)$ $(1 \leq i \leq n)$ means $E - A_i \in \mathfrak{f}(x)$ $(1 \leq i \leq n)$ and $E - \bigcup_{i=1}^{n} A_i = \bigcap_{i=1}^{n} (E - A_i) \in \mathfrak{f}(x)$, which contradicts $\bigcup_{i=1}^{n} A_i \in \mathfrak{f}(x)$. $\mathfrak{f}(x)$ is compressed, therefore if $<_1 \in \mathscr{S}$, $< \subset <_1^2$ and $A_{i_0} <_1 <_1 C <_1 B_{i_0}$, then $C \in \mathfrak{f}(x)$, and from this $B_{i_0} \in \mathfrak{f}(x)$.

(5.1.2): Put $<_3 \in \mathscr{S}$, $<_2 \subset <_3^2$ and $C <_3 D <_3 B$. Suppose $x \in E' - s(B)$. Then $D \notin \mathfrak{f}(x)$. This implies $\emptyset \notin \{E - D\}(\cap)\mathfrak{f}(x)$, therefore in view of $E - C \in \mathfrak{f}(x)$, we have $E - A \in \mathfrak{s}(x)$, that is $x \in s(E-A)$. ∎

Lemma 5.2. *In a syntopogenous space* $[E, \mathscr{S}]$

(5.2.1) *the maximal round filters coincide with the round compressed filters;*

(5.2.2) *for two* $\mathscr{S}^s$*-round compressed filters* $\mathfrak{f}_1$ *and* $\mathfrak{f}_2$

$$\mathscr{S}(\mathfrak{f}_1) = \mathscr{S}(\mathfrak{f}_2) \Leftrightarrow \mathfrak{f}_1 = \mathfrak{f}_2;$$

(5.2.3) *if under the conditions of* (5.2.2) $\mathfrak{f}_1$ *is round in* $\mathscr{S}^c$, *then* $\emptyset \notin \mathfrak{f}_1(\cap)\mathscr{S}(\mathfrak{f}_2) \Leftrightarrow \mathfrak{f}_1 = \mathfrak{f}_2$.

PROOF. (5.2.1): Suppose that $\mathfrak{f}$ is a round compressed filter in $\mathscr{S}$, and $\mathfrak{f}_1$ is a round filter such that $\mathfrak{f}\subset\mathfrak{f}_1$. Then $F\in\mathfrak{f}_1$ implies $F_1<F$ for some $F_1\in\mathfrak{f}_1$ and $<\in\mathscr{S}$. $\emptyset\notin\{F_1\}(\cap)\mathfrak{f}$, consequently $F\in\mathfrak{f}$. This shows $f=f_1$, that is, $\mathfrak{f}$ is a maximal round filter. Conversely, let us assume that $\mathfrak{f}$ is a maximal round filter and $\emptyset\notin\{A\}(\cap)\mathfrak{f}$, where $A<B$, $<\in\mathscr{S}$. We show that $B\in\mathfrak{f}$. Indeed, $\mathfrak{f}$ can be included into the round filter $\mathscr{S}(A)(\cap)\mathfrak{f}$, hence the condition of the maximality of $\mathfrak{f}$ implies

$$B\in\mathscr{S}(A)\subset\mathscr{S}(A)(\cap)\mathfrak{f}=\mathfrak{f}.$$

(5.2.2): Suppose $\mathscr{S}(\mathfrak{f}_1)=\mathscr{S}(\mathfrak{f}_2)$ and $B\in\mathfrak{f}_1$. There exist $<\in\mathscr{S}$ and $A\in\mathfrak{f}_1$ for which $A<^s B$. Then by [1], (3.44) $A=\bigcup_{i=1}^{m}(A_i\cap A_i')$ and $B=\bigcup_{i=1}^{m}(B_i\cap B_i')$, where m is a suitable natural number, further $A_i<B_i$ and $A_i'<^c B_i'$ $(1\leqq i\leqq m)$. Put $<_1\in\mathscr{S}$, $<\subset<_1^2$ and $A_i<_1C_i<_1B_i$, $A_i'<_1^cC_i'<_1^cB_i'$ $(1\leqq i\leqq m)$. By $\emptyset\notin\{A\}(\cap)\mathfrak{f}_1$ we get $\emptyset\notin\{A_i\cap A_i'\}(\cap)\mathfrak{f}_1$ for an index i. $\mathfrak{f}_1$ is compressed, therefore $C_i\cap C_i'\in\mathfrak{f}_1$. We have $B_i\in\mathscr{S}(\mathfrak{f}_1)=\mathscr{S}(\mathfrak{f}_2)\subset\mathfrak{f}_2$. On the other hand, $B_i'\in\mathfrak{f}_2$ is also true. In fact, $\emptyset\notin\{C_i'\}(\cap)\mathfrak{f}_2$, else $E-C_i'\in\mathfrak{f}_2$ would imply $E-A_i'\in\mathscr{S}(\mathfrak{f}_2)=\mathscr{S}(\mathfrak{f}_1)\subset\mathfrak{f}_1$. Thus $B_i\cap B_i'\in\mathfrak{f}_2$, consequently, $B\in\mathfrak{f}_2$. This gives $\mathfrak{f}_1\subset\mathfrak{f}_2$, and from (5.2.1) $\mathfrak{f}_1=\mathfrak{f}_2$ issues. The inverse implication is clear.

(5.2.3): By the compressedness of $\mathfrak{f}_1$, $\emptyset\notin\mathfrak{f}_1(\cap)\mathscr{S}(\mathfrak{f}_2)$ implies $\mathscr{S}(\mathfrak{f}_2)\subset\mathfrak{f}_1$, so that $\mathscr{S}(\mathfrak{f}_2)\subset\mathscr{S}(\mathfrak{f}_1)$. Conversely, let us suppose $V\in\mathscr{S}(\mathfrak{f}_1)$. Then there exist $<\in\mathscr{S}$ and $B\in\mathfrak{f}_1$ such that $B<V$. For $V_1\in\mathscr{S}$, $<\subset<_1^2$, we can find a set C with $B<_1C<_1<_1V$. Choose $<_2\in\mathscr{S}$ and $A\in\mathfrak{f}_1$ for which $A<_2^cB$. In view of our condition $E-B\notin\notin\mathfrak{f}_2$, else $E-A\in\mathscr{S}(\mathfrak{f}_2)$ would be true. Thus $\emptyset\notin\{B\}(\cap)\mathfrak{f}_2$, so that because of the compressedness of $\mathfrak{f}_2$ we get $C\in\mathfrak{f}_2$ and $V\in\mathscr{S}(\mathfrak{f}_2)$. These show $\mathscr{S}(\mathfrak{f}_1)=\mathscr{S}(\mathfrak{f}_2)$, hence by (5.2.2) $\mathfrak{f}_1=\mathfrak{f}_2$. The converse implication is obvious. ∎

Further on we shall study the subset

$$E_1^*=\{x\in E^*-E:\ \mathfrak{f}(x)\text{ is }\mathscr{S}^c\text{-round}\}\cup E$$

of E^* containing E, denoting by $\mathscr{S}_1^*$ the syntopogenous structure $\mathscr{S}^*|E_1^*$.

LEMMA 5.3. *For* $E\subset E'\subset E^*$, $\mathscr{S}'=\mathscr{S}^*|E'$ *is compact if and only if* $E_1^*\subset E'$.

PROOF. First of all we prove that if $E_1^*\subset E'$, then $\mathscr{S}'$ is compact. Suppose that $\mathfrak{f}'$ is an arbitrary filter in E'. Then $\mathscr{S}'^c(\mathfrak{f}')$ is round in $\mathscr{S}'^c$, and since E is $\mathscr{S}'^c$-dense, the filter $\mathfrak{f}=\{E\}(\cap)\mathscr{S}'^c(\mathfrak{f}')$ is round in $\mathscr{S}^c$. $\mathfrak{f}$ can be included into a maximal (i.e. compressed by (5.2.1)) $\mathscr{S}^c$-round filter, say $\mathfrak{f}(x)$, for some $x\in E_1^*\subset E'$. (If this maximal filter $\mathfrak{f}_0$ is non-convergent in $\mathscr{S}^s$, then $x\in E_1^*-E$. If $\mathfrak{f}_0\to x\in E$ in $\mathscr{S}^s$, then $\mathfrak{f}(x)=\mathscr{S}^s(x)\subset\mathfrak{f}_0$, the $\mathscr{S}^s$-roundness of $\mathfrak{f}_0$ and the maximality of $\mathfrak{f}(x)$ imply $\mathfrak{f}(x)=\mathfrak{f}_0$.) x is a cluster point of $\mathfrak{f}'$, otherwise there exists $<'\in\mathscr{S}'$ such that $x<'<'E'-F$ for some $F\in\mathfrak{f}'$. If $<_1'\in\mathscr{S}'$, $<'\subset<_1'^2$, then $x<_1'V<_1'E'-F$ can be written, so that $E\cap V\in\mathfrak{s}(x)\subset\mathfrak{f}(x)$. On the other hand, $E'-V\in\mathscr{S}'^c(\mathfrak{f}')$, hence $E-(E\cap V)=E\cap(E'-V)\in\mathfrak{f}\subset\mathfrak{f}(x)$, which is a contradiction.

Conversely, if $\mathscr{S}'$ is compact, then, for any $x\in E_1^*-E$, $\mathfrak{f}(x)$ has a cluster point $y\in E'$. One can easily show that in this case $\emptyset\notin f(x)(\cap)\mathfrak{s}(y)=\mathfrak{f}(x)(\cap)\mathscr{S}(\mathfrak{f}(y))$, hence (5.2.3) $x=y$. Thus $E_1^*\subset E'$. ∎

LEMMA 5.4. *Suppose* $E \subset E' \subset E^*$ *and* $x \in E' - E$. *Then* $\{x\}$ *is closed in* $\mathscr{S}'$ *iff* $x \in E_1^*$.

PROOF. Put $x \in E_1^*$. $y \in E'$ lies in the $\mathscr{S}'$-closure of $\{x\}$ if $\mathfrak{s}(y) \subset \mathfrak{s}(x)$. Then $\emptyset \notin \mathfrak{f}(x)(\cap)\mathfrak{s}(y) = \mathfrak{f}(x)(\cap)\mathscr{S}(\mathfrak{f}(y))$ implies $x = y$, so that $\{x\}$ is $\mathscr{S}'$-closed.

Conversely, if $\{x\}$ is closed in $\mathscr{S}'$, then for any $y \in E'$ $(x \neq y)$, $\mathfrak{s}(y) \not\subset \mathfrak{s}(x)$. Consider the system

$$\mathfrak{f} = \{E - X\colon X < Y,\ Y \notin \mathfrak{s}(x),\ < \in \mathscr{S}\},$$

which is an $\mathscr{S}^c$-round filter. In fact, $E - X' \supset E - X \in \mathfrak{f}$ implies $X' \subset X < Y \notin \mathfrak{s}(x)$ for some $< \in \mathscr{S}$, therefore $E - X' \in \mathfrak{f}$. If $E - X_1,\ E - X_2 \in \mathfrak{f}$, then $X_1 <_1 Y_1 \notin \mathfrak{s}(x)$, $X_2 <_2 Y_2 \notin \mathfrak{s}(x)$ for suitable $<_1, <_2 \in \mathscr{S}$ and $Y_1, Y_2 \subset E$. Assume $<_1 \subset <_1'^2$, $<_2 \subset <_2'^2$, where $<_1', <_2' \in \mathscr{S}$. Then, for $<_1' \cup <_2' \subset <' \in \mathscr{S}$, there are sets $Z_1, Z_2 \subset E$ such that $X_i <' Z_i <' Y_i$ $(i = 1, 2)$, and from this $X_1 \cup X_2 <' Z_1 \cup Z_2 \notin \mathfrak{s}(x)$ (see (5.1.1)), thus $(E - X_1) \cap (E - X_2) = E - (X_1 \cup X_2) \in \mathfrak{f}$. If $E - X \in \mathfrak{f}$, then $X < Y \notin \mathfrak{s}(x)$ $(< \in \mathscr{S},\ Y \subset E)$ implies the existence of an order $<_1 \in \mathscr{S}$ and a set Z such that $X <_1 Z <_1 Y$. Since $E - Z <_1^c E - X$ and $E - Z \in \mathfrak{f}$, $\mathfrak{f}$ is an $\mathscr{S}^c$-round filter indeed. $\mathfrak{f}$ can be included into an $\mathscr{S}^c$-round compressed filter, say $\mathfrak{f}(y)$, where $y \in E_1^*$ ($y \in E$ or $y \in E_1^* - E$ depends on the convergence of this compressed filter in $\mathscr{S}^s$). We show that $x = y$. In fact, if $x \neq y$, then there exists $Y \in \mathfrak{s}(y)$ such that $Y \notin \mathfrak{s}(x)$. One can find a set $X \in \mathfrak{f}(y)$ so that $X < Y$ for a suitable $< \in \mathscr{S}$, hence $E - X \in \mathfrak{f} \subset \mathfrak{f}(y)$, which is a contradiction. ∎

A compact extension $[E', \mathscr{S}']$ will be called a *simple compactification* of the syntopogenous space $[E, \mathscr{S}]$ iff E is dense in $[E', \mathscr{S}'^s]$ and $\{x\}$ is closed in $\mathscr{S}'$ for $x \in E' - E$.

THEOREM 5.5. *Any syntopogenous space* $[E, \mathscr{S}]$ *has simple compactifications. If* $[E', \mathscr{S}']$ *is a simple compactification and* $[E^*, \mathscr{S}^*]$ *is a double compactification of* $[E, \mathscr{S}]$, *then there exists a unique isomorphism of* $[E', \mathscr{S}']$ *onto* $[E_1^*, \mathscr{S}_1^*]$, *which coincides with the identity mapping on E, consequently, two simple compactifications of* $[E, \mathscr{S}]$ *are always equivalent.*

PROOF. In view of (5.3) and (5.4) $[E_1^*, \mathscr{S}_1^*]$ is a simple compactification of $[E, \mathscr{S}]$ for every double compactification $[E^*, \mathscr{S}^*]$ of $[E, \mathscr{S}]$. If $[E', \mathscr{S}']$ is a simple compactification of $[E, \mathscr{S}]$, then by the $\mathscr{S}'$-closedness of the points of $E' - E$, $\mathscr{S}'$ is relatively separated with respect to E, thus the existence of the isomorphism described in the theorem is a direct consequence of [1], (16.45) and Lemmas 5.3, 5.4 of the present paper. ∎

Let us consider an arbitrary syntopological space $[E, \mathscr{S}_0]$, and let $\mathscr{S}$ denote a syntopogenous structure on E such that $\mathscr{S}_0 \sim \mathscr{S}^p$. It is obvious that if we choose a simple compactification $[E', \mathscr{S}']$ of $[E, \mathscr{S}]$, then $[E', \mathscr{S}'^p]$ is a compact extension of $[E, \mathscr{S}_0]$. As it will be shown by (5.6), in this way we get a generalization of the notion of a Wallman-type compactification of a topological space.

EXAMPLE 5.6. Let $\mathfrak{S}$ be a lattice-base for the topology $\mathscr{T}_0$ on E, and let us denote by $\mathscr{T}$ the topogenous structure generated by $\mathfrak{S}$. Then $\mathscr{T}_0 = \mathscr{T}^p$, and if $[E', \mathscr{T}']$ is a simple compactification of $[E, \mathscr{T}]$, $[E', \mathscr{T}'^p]$ is a Wallman-type compactification of $[E, \mathscr{T}_0]$ belonging to $\mathfrak{S}$ (in the sense of 6.1.e of [2]). In fact, with the notation of (1.13) the $\mathfrak{S}^c$-ultrafilters are identical with the $\mathscr{T}^c$-round compressed filters in E. ∎

Now we shall discuss the separation axioms (S_1) and (S_2) in the simple compactifications. First of all an easy lemma will be verified.

LEMMA 5.7. *Let us consider the following two pairs of conditions for a syntopogenous space* $[E, \mathcal{S}]$:

(i) $\mathcal{S}^p < \mathcal{S}^{cp}$;

(ii) *if* $<\in\mathcal{S}$, *there exists* $<_1\in\mathcal{S}$ *such that for each* $\mathcal{S}^c$*-round compressed non-convergent filter* $\mathfrak{f}$ *in* $\mathcal{S}$, (5.7.1) $A<B$ *and* $A\in\mathfrak{f}$ *imply* $C<_1^c B$ *with some* $C\in\mathfrak{f}$. *Further*

(i′) *For every* $x\in E$, $\mathcal{S}^s(x)$ *is round in* $\mathcal{S}^c$;

(ii′) *for* $<\in\mathcal{S}$ *there exists* $<_1\in\mathcal{S}$ *such that any* $\mathcal{S}^c$*-round compressed filter* $\mathfrak{f}$ *has property* (5.7.1). *Then the statement* (i) & (ii) *is equivalent to* (i′) & (ii′).

PROOF. (i) & (ii)⇒(i′) & (ii′): Let us observe that if $x\in E$, then

$$\mathcal{S}^s(x) = \{V_1\cap V_2\colon\ x<V_1,\ x<^c V_2,\ <\in\mathcal{S}\},$$

therefore (i) implies (i′). In fact, suppose $V\in\mathcal{S}^s(x)$. Then $V=V_1\cap V_2$, where $x<V_1$, $x<^c V_2$ for some $<\in\mathcal{S}$. If $<_0\in\mathcal{S}$, $<^p\subset<_0^{cp}$, $<'\in\mathcal{S}$, $<_0\cup<\subset<'$, finally $<_1\in\mathcal{S}$, $<'\subset<_1^2$, then there exist sets $W_1, W_2\subset E$ such that $x<_1^c W_1<_1^c V_1$ and $x<_1^c W_2<_1^c V_2$, consequently, $W_1\cap W_2\in\mathcal{S}^c(x)\subset\mathcal{S}^s(x)$ and $W_1\cap W_2<_1^c V$.

For the sake of the verification of (ii′) let $<$ be a member of $\mathcal{S}$ and choose $<'\in\mathcal{S}$ in accordance with (ii). Suppose that $<^p\subset<_0^{cp}$, $<_0\in\mathcal{S}$ (see (i)), $<_0'\in\mathcal{S}$, $<_0\subset<_0'^2$ and $<'\cup<_0'\subset<_1\in\mathcal{S}$. Let $\mathfrak{f}$ be an $\mathcal{S}^c$-round compressed filter. If it does not converge in $\mathcal{S}$, then by (ii) $A<B$, $A\in\mathfrak{f}$ imply $C<_1^c B$ for some $C\in\mathfrak{f}$. On the other hand, assume that $x\in E$ is a limit point of $\mathfrak{f}$ in $\mathcal{S}$. Then $x\in A$ for every $A\in\mathfrak{f}$, else, for a suitable $<_2\in\mathcal{S}$ and $X\in\mathfrak{f}$, $x\in E-A<_2 E-X$ would be true, consequently x could not be a cluster point of $\mathfrak{f}$. Therefore $A<B$, $A\in\mathfrak{f}$ imply $x<B$, thus $x<_0^c B$, and $x<_0'^c C<_0'^c B$ for some $C\subset E$. Then $C<_1^c B$ and by the compressedness of $\mathfrak{f}$ we have $C\in\mathfrak{f}$.

(i′) & (ii′)⇒(i) & (ii): Condition (ii′) is obviously stronger than (ii). In order to prove the validity of (i), put $<\in\mathcal{S}$, $<_0\in\mathcal{S}$, $<\subset<_0^2$ and choose $<_1\in\mathcal{S}$ for $<_0$ in accordance with (ii′). Then $x<B$ implies $x<_0 A<_0 B$, hence by the compressedness of $\mathcal{S}^s(x)$ and $A\in\mathcal{S}^s(x)$, from (ii′) we obtain $C<_1^c B$ for some $C\in\mathcal{S}^s(x)$. $x\in C$ yields $x<_1^c B$, therefore $<^p\subset<_1^{cp}$. ∎

As in a topogenous space (5.7) (ii) and (ii′) are always true, we get the following generalization of Theorem (6.1.52) of [2] (cf. also (1.16) of the present paper).

THEOREM 5.8. *A simple compactification of the syntopogenous space* $[E, \mathcal{S}]$ *is an* S_1*-space iff* $[E, \mathcal{S}]$ *satisfies conditions* (i)—(ii) (*or equivalently* (i′)—(ii′)) *of* (5.7).

PROOF. By (5.5) it is sufficient to verify the assertion for the case of the space $[E_1^*, \mathcal{S}_1^*]$, where $[E^*, \mathcal{S}^*]$ is the double compactification of $[E, \mathcal{S}]$. Suppose that $[E_1^*, \mathcal{S}_1^*]$ has the property (S_1). Then in view of (2.7) $\mathcal{S}_1^{*p}<\mathcal{S}_1^{*cp}$, thus the validity of (i) is clear. Further assume $<\in\mathcal{S}$, $<_0\in\mathcal{S}$, $<\subset<_0^2$, $<'=s(<_0)$, $<_1'\in\mathcal{S}_1^*$, $<'^p\subset$ $\subset<_1'^{cp}$, $<_2'\in\mathcal{S}_1^*$, $<_1'\subset<_2'^2$, finally $<_2'|E\subset<_1\in\mathcal{S}$. If $\mathfrak{f}$ is an $\mathcal{S}^c$-round compressed non-convergent filter in $\mathcal{S}$, there exists $x\in E_1^*-E$ such that $\mathfrak{f}=\mathfrak{f}(x)$. Putting $A<B$ and $A\in\mathfrak{f}$, we have $A<_0 X<_0 B$ for some $X\subset E$. Then $x\in s(X)$, therefore

$x<'s(B)$. Consequently, $x<_2'^c C'<_2'^c s(B)$ for a set $C'\subset E_1^*$, and this implies $C=C'\cap E\in\mathfrak{f}$ and $C<_1^c B$.

In fact, if $<_2\in\mathcal{S}$, $<_2'=s(<_2)$, then $E_1^*-C'\,s(<_2)\,E_1^*-x$, that is,

$$E_1^*-C'\subset\bigcup_{i=1}^{m}\bigcap_{j=1}^{n_i}s(V_{ij}),\qquad \bigcup_{i=1}^{m}\bigcap_{j=1}^{n_i}s(W_{ij})\subset E_1^*-x,$$

where m and n_i are suitable natural numbers and $V_{ij}<_2W_{ij}$ $(1\leqq i\leqq m,\ 1\leqq j\leqq n_i)$. For $\bigcap_{j=1}^{n_i}V_{ij}=V_i$ and $\bigcap_{j=1}^{n_i}W_{ij}=W_i$, we have

$$s(V_i)=\bigcap_{j=1}^{n_i}s(V_{ij})\quad\text{and}\quad s(W_i)=\bigcap_{j=1}^{n_i}s(W_{ij}),$$

therefore

$$E_1^*-C'\subset\bigcup_{i=1}^{m}s(V_i)\quad\text{and}\quad\bigcup_{i=1}^{m}s(W_i)\subset E_1^*-x.$$

Suppose $<_3\in\mathcal{S}$, $<_2\boldsymbol{\subset}<_3^2$. In view of the topogenity of $<_2$, there exists $Z_i\subset E$ for any $1\leqq i\leqq m$ such that $V_i<_3Z_i<_3W_i$. Putting $V=\bigcup_{i=1}^{m}V_i$ and $Z=\bigcup_{i=1}^{m}Z_i$, we obtain $V<_3Z$, $E_1^*-C'\subset s(V)$ and $s(Z)\subset\bigcup_{i=1}^{m}s(W_i)\subset E_1^*-x$ (cf. (5.1.1)). Further assume $<_4\in\mathcal{S}$, $<_3\boldsymbol{\subset}<_4^2$ and $V<_4U<_4Z$. Then $x\notin s(Z)\Rightarrow U\notin\mathfrak{f}(x),\Rightarrow$ $\Rightarrow E-V\in\mathfrak{f}(x)\Rightarrow C\in\mathfrak{f}(x)$, because $E-V\subset E\cap(E_1^*-s(V))\subset C$. Finally, from the choice of $<_1$ we get $C=C'\cap E<_1^c s(B)\cap E\subset B$, that is $C<_1^c B$.

Conversely, under (i′) and (ii′) put $<'\in\mathcal{S}_1^*$, $<\in\mathcal{S}$, $<'\boldsymbol{\subset}s(<)$, $<_1\in\mathcal{S}$, $<\boldsymbol{\subset}<_1^2$. Let $<_2\in\mathcal{S}$ be an order determined by (ii′) for $<_1$, finally, suppose $<_2\boldsymbol{\subset}<_3^2$ for $<_3\in\mathcal{S}$ and $s(<_3)\boldsymbol{\subset}<_1'\in\mathcal{S}_1^*$. By (i′) $\mathfrak{f}(x)$ is an $\mathcal{S}^c$-round compressed filter for any $x\in E_1^*$. Then $x,y\in E_1^*$, $x<'E_1^*-y$ imply the existence of sets $A,B\subset E$ such that $A<B$, $x\in s(A)$ and $s(B)\subset E_1^*-y$. Assume $A<_1X<_1B$, then because of $A\in\mathfrak{f}(x)$ we have $C<_2^c X$, where $C\in\mathfrak{f}(x)$. Consequently, $C<_3^c Y<_3^c X$ for some $Y\subset E$. On the basis of (5.1.2) we obtain from $Y<_3^c X<_1B$ the relations

$$y\in E_1^*-s(B)\subset s(E-Y)<_1's(E-C)\subset E_1^*-x,$$

so that $x<_1'^c E_1^*-y$. ∎

In general there is no connection between the normality of a syntopogenous space and that of its subspaces. In this respect we have a remarkable exception.

Proposition 5.9. *Let $[E',\mathcal{S}']$ be an arbitrary subspace of the double compactification of the syntopogenous space $[E,\mathcal{S}]$ and suppose $E\subset E'$. Then $\mathcal{S}'$ is normal iff so is $\mathcal{S}$. In this case we have $\mathcal{S}'^{\triangledown}|E\sim\mathcal{S}^{\triangledown}$.*

Proof. Assume that $\mathcal{S}'$ is normal, and let $<\in\mathcal{S}$, $<_1\in\mathcal{S}$, $<\boldsymbol{\subset}<_1^2$, $<_1'=s(<_1)$, $<_2'\in\mathcal{S}'$, $<_1'^c\cdot<_1'\boldsymbol{\subset}<_2'\cdot<_2'^c$, $<_2\in\mathcal{S}$ such that $<_2'|E\boldsymbol{\subset}<_2$. Put $A<^cC<B$, further $A<_1^cC_1<_1^cC<_1B_1<_1B$. Then by (5.1.2)

(1) $$A'=E'-s(E-A)<_1'^c E'-s(E-C_1)\subset s(B_1)<_1's(B)=B',$$

hence $A' <_2' C' <_2'^c B'$ for a set $C' \subset E'$, thus

$$A \subset A' \cap E <_2 C' \cap E <_2^c B' \cap E \subset B. \tag{2}$$

Conversely, suppose that $\mathscr{S}$ is normal, and let $<' \in \mathscr{S}'$, $< \in \mathscr{S}$, $<' \subset s(<)$, $<_0 \in \mathscr{S}$, $< \subset <_0^2$, $<_1 \in \mathscr{S}$, $<_0 \subset <_1^2$, $<_2 \in \mathscr{S}$, $<_1^c \cdot <_1 \subset <_2 \cdot <_2^c$, $<_3 \in \mathscr{S}$, $<_2 \subset <_3^2$, finally $<_1' \in \mathscr{S}'$, $s(<_3) \subset <_1'$. We shall prove $<'^c \cdot <' \subset <_1' \cdot <_1'^c$. In fact, put $X < <'^c Z <' Y$. With the help of (5.1.1) one can verify that there exist subsets A, B, C, D of E such that $A <_0 B$, $C <_0 D$ and

$$E' - Z \subset s(A), \quad s(B) \subset E' - X, \quad Z \subset s(C), \quad s(D) \subset Y.$$

From this $E - A \subset C$ follows, hence $E - B <_0^c E - A \subset C <_0 D$. If $E - B <_1^c B_1 <_1^c <_1^c C <_1 D_1 <_1 D$, then $B_1 <_2 F <_2^c D_1$ for a suitable $F \subset E$. Put $B_1 <_3 F_1 <_3 F <_3^c F_2 <_3^c D_1$. In view of (5.1.2) $X \subset E' - s(B) \subset s(F_1) <_1' s(F)$ and $E' - Y \subset E' - s(D) \subset s(E - F_2) <_1' <_1' s(E - F) \subset E' - s(F)$, consequently, $X <_1' s(F) <_1'^c Y$.

It can be easily seen that $\mathscr{S}'^\triangledown | E < \mathscr{S}^\triangledown$, and the relation $\mathscr{S}^\triangledown < \mathscr{S}'^\triangledown | E$ can be read from formulas (1), (2) of the first part of the proof. ▮

Thus Theorem (6.1.53) of [2] can be generalized as follows (cf. also (1.13) of this paper).

THEOREM 5.10. *A simple compactification of the syntopogenous space* $[E, \mathscr{S}]$ *is an* S_2*-space iff* $[E, \mathscr{S}]$ *satisfies conditions* (i)—(ii) (*or equivalently* (i′)—(ii′)) *of* (5.7) *and it is normal.*

PROOF. These conditions are sufficient, because in this case the simple compactification is a normal S_1-space, hence it is also an S_2-space (see (5.8), (5.9), (2.7), (1.17), (1.11), (1.5)). In view of (2.8), (5.8) and (5.9), the conditions are necessary, too. ▮

Let us observe that, under the conditions of (5.10), denoting by $[E', \mathscr{S}']$ the simple S_2-compactification in question, we get $[E', \mathscr{S}'^p]$ as a symmetrizable compactification of $[E, \mathscr{S}^p]$ belonging to the symmetrical syntopogenous structure $\mathscr{S}^\triangledown$ on E by (2.9) and (5.9). The connection existing between Wallman compactification and Čech—Stone compactification of a completely regular topological space, can be extended as follows (cf. [2], (6.4.25) and [3]):

THEOREM 5.11. *Let* $[E', \mathscr{S}']$ *be a simple compactification of the symmetrizable syntopological space* $[E, \mathscr{S}]$. *Then* $[E', \mathscr{S}'^p]$ *is symmetrizable iff* $[E, \mathscr{S}]$ *is normal and has property* (ii) *of* (5.7). *In this case* $[E', \mathscr{S}'^p]$ *is the finest symmetrizable compactification of* $[E, \mathscr{S}]$.

PROOF. The conditions are necessary and sufficient, indeed, because the symmetrizable syntopology $\mathscr{S}$ satisfies (5.7) (i) automatically (see (1.15)). Since this compactification belongs to $\mathscr{S}^\triangledown$ (in the sense of [3]), by (1.14) we get that it is the finest one.

REFERENCES

[1] Császár, Á., *Foundations of general topology*, A Pergamon Press Book, The Macmillan Co., New York, 1963. *MR* **28** # 575.

[2] Császár, Á., *General topology*, Adam Hilger, Ltd., Bristol, 1978. *MR* **57** # 13812

[3] Császár, Á., Double compactification d'espaces syntopogènes, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* 7 (1964), 3—11. *MR* **30** # 3443.

[4] Császár, Á., Erweiterung, Kompaktifizierung und Vervollständigung syntopogener Räume, *Contributions to Extension Theory of Topological Structures*, Berlin, 1969, 51—54.

[5] Császár, Á., Normality is productive, *Topology*, Vol. I (Proc. Fourth Colloq., Budapest, 1978), Colloq. Math. Soc. J. Bolyai, 23, North-Holland Publ. Co., Amsterdam, 1980, 323—341. *MR* **82i**: 54005.

[6] Császár, Á. and Matolcsy, K., Syntopogenous extensions for prescribed topologies, *Acta Math. Acad. Sci. Hungar.* **37** (1981), 59—75. *MR* **82i**: 54006.

[7] Frink Jr., O., Compactifications and semi-normal spaces, *Amer. J. Math.* **86** (1964), 602—607. *MR* **29** # 4028.

[8] Matolcsy, K., On extensions of syntopogenous spaces, *Publ. Math. Debrecen* **28** (1981), 103—119. *MR* **82i**: 54007.

[9] Sieber, J. L. and Pervin, W. J., Separation axioms for syntopogenous spaces, *Nederl. Akad. Wetensch. Proc. Ser. A.* **66** = *Indag. Math.* **25** (1963), 755—760. *MR* **28** # 3399.

[10] Singal, M. K. and Lal, Sunder, Local proximity ordered spaces, *Math. Student* **41** (1973), 269—275. *MR* **52** # 9187.

(Received July 15, 1980)

SZABÓ ISTVÁN ALTÁBORNAGY TÉR 8
XIV. EMELET 112
H—4032 DEBRECEN
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 221—228

ПРИМЕНЕНИЕ ЧИСЛЕННО-АНАЛИТИЧЕСКОГО МЕТОДА А. М. САМОЙЛЕНКО К ИССЛЕДОВАНИЮ ПЕРИОДИЧЕСКИХ ЛИНЕЙНЫХ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ С МАКСИМУМАМИ

Г. Х. САРАФОВА и Д. Д. БАЙНОВ

Рассмотрим скалярное дифференциальное уравнение

$$\dot{y}(t)+y(t)+q\max_{\tau\in[t-h,\,t]}y(\tau)=f_0(t), \tag{1}$$

где q — положительная постоянная, $f_0(t)-T$-периодическая, непрерывная в интервале $(-\infty, +\infty)$ функция, h — малая положительная постоянная. Поставим себе задачу найти тех значений h, для которых уравнение (1) будет иметь периодическое решение. При решении этой задачи будем применять численно-аналитический метод А. М. Самойленко [1]—[3] и некоторые его обобщения [4].

Положим в (1) $h=0$. Получим линейное дифференциальное уравнение

$$\dot{y}_0(t)+(1+q)y_0(t)=f_0(t). \tag{2}$$

Как известно, это уравнение имеет единственное периодическое решение, которое записывается формулой

$$y_0(t)=\int_{-\infty}^{t}e^{-(1+q)(t-\tau)}f_0(\tau)\,d\tau. \tag{3}$$

Положим

$$y(t)=y_0(t)+x(t). \tag{4}$$

Тогда функция $x(t)$ будет удовлетворять уравнению

$$\dot{x}(t)+(1+q)x(t)=-q\Big[\max_{\tau\in[t-h,\,t]}\big(y_0(\tau)+x(\tau)\big)-\big(y_0(t)+x(t)\big)\Big]. \tag{5}$$

Введем обозначения. Если дано линейное дифференциальное уравнение

$$\dot{x}(t)+\lambda x(t)=f(t),$$

то через $\widehat{f(t)}$ будем обозначать выражение

$$\widehat{f(t)}=\frac{\lambda}{1-e^{-\lambda T}}\int_{0}^{T}e^{-\lambda(T-\tau)}f(\tau)\,d\tau. \tag{6}$$

Периодическое решение уравнения (5) будем искать численно-аналитическим методом А. М. Самойленко.

1980 *Mathematics Subject Classification.* Primary 34C25.
Key words and phrases. Numerical-analytic method, linear differential equations with maximums.

Рассмотрим последовательность периодических по t периода T функций

(7)
$$x_{n+1}(t, x_0) = x_0 - q \int_0^t e^{-(1+q)(t-\tau)}\{[\max_{s\in[\tau-h,\tau]}(y_0(s)+x_n(s)) - (y_0(\tau)+x_n(\tau))] - \widetilde{[\max_{s\in[\tau-h,\tau]}(y_0(s)+x_n(s))-(y_0(\tau)+x_n(\tau))]}\}\, d\tau.$$

Если предположить, что последовательность $\{x_n(t, x_0)\}$ равномерно сходится к $x_\infty(t, x_0)$, то функция $x_\infty(t, x_0)$ будет периодическим решением (5), проходящее при $t=0$ через точку x_0, если x_0 является решением уравнения

(8)
$$(1+q)x_0 + q\widetilde{[\max_{\tau\in[t-h,t]}(y_0(\tau)+x_\infty(\tau, x_0))-(y_0(t)+x_\infty(t, x_0))]} = 0.$$

Итак, вопрос существования и отыскания T-периодического решения уравнения (5) сводится к нахождению условий равномерной сходимости последовательности (7) и разрешимости (8).

Обозначим через $t_1\in[t-h, t]$ точку, в которой достигается

$$\max_{s\in[t-h,t]}(y_0(s)+x_0), \quad \text{т.е.,} \quad \max_{s\in[t-h,t]}(y_0(s)+x_0)=y_0(t_1)+x_0.$$

Тогда

$$\max_{s\in[t-h,t]}(y_0(s)+x_0)-(y_0(t)+x_0) = y_0(t_1)-y_0(t).$$

Оценим $|y_0(t_1)-y_0(t)|$. Имеем

$$|y_0(t_1)-y_0(t)| \leqq h|\dot{y}_0(\tilde{t})|, \qquad \tilde{t}\in(t-h, t).$$

С другой стороны, существует постоянная $\beta=\mathrm{const}$, такая, что $\sup_t|\dot{y}_0(t)|\leqq\beta$. Следовательно, в интервале $(-\infty, +\infty)$ будем иметь

(9)
$$|y_0(t_1)-y_0(t)| \leqq h\beta.$$

Введем обозначение $\lambda=-(1+q)$. Оценим модуль разности $x_1(t, x_0)-x_0$. Имеем

(10)
$$|x_1(t, x_0)-x_0| = |(1+\lambda)|\left|\int_0^t e^{\lambda(t-\tau)}\{[\max_{s\in[\tau-h,\tau]}(y_0(s)+x_0) - (y_0(\tau)+x_0)] - \widetilde{[\max_{s\in[\tau-h,\tau]}(y_0(s)+x_0)-(y_0(\tau)+x_0)]}\}\, d\tau\right| \leqq$$
$$\leqq |1+\lambda|h\beta\left(\frac{e^{\lambda T}-e^{\lambda t}}{e^{\lambda T}-1}\int_0^t e^{-\lambda\tau}\,d\tau + \frac{e^{\lambda t}-1}{e^{\lambda T}-1}\int_t^T e^{\lambda(T-\tau)}\,d\tau\right) =$$
$$= |1+\lambda|h\beta\alpha(\lambda, t) \leqq |1+\lambda|h\beta\alpha\left(\lambda, \frac{T}{2}\right) = qh\beta\alpha\left(\lambda, \frac{T}{2}\right),$$

где

(11)
$$\alpha(\lambda, t) = \frac{2(e^{\lambda(T-t)}-1)(e^{\lambda t}-1)}{\lambda(e^{\lambda T}-1)}.$$

Введем обозначение $d_\lambda = h\beta|1+\lambda|\alpha\left(\lambda, \frac{T}{2}\right)$. Из (10) видно, что при $t\in(-\infty, +\infty)$, $x_0\in[a+d_\lambda, b-d_\lambda]$ функция $x_1(t, x_0)\in[a, b]$. Методом математической индукции можно показать, что для всех $n=0, 1, 2, \ldots$ и при $t\in(-\infty, +\infty)$, $x_0\in[a+d_\lambda, b-d_\lambda]$, функции $x_n(t, x_0)$ принадлежат интервалу $[a, b]$.

Оценим $|\dot{x}_1(t, x_0)|$. Имеем

$$\dot{x}_1(t, x_0) = -q\Big\{\Big[\max_{s\in[t-h, t]}(y_0(s)+x_0)-(y_0(t)+x_0)\Big]-$$

$$-\widehat{\Big[\max_{s\in[t-h, t]}(y_0(s)+x_0)-(y_0(t)+x_0)\Big]}\Big\}-$$

$$-q\lambda\int_0^t e^{\lambda(t-\tau)}\Big\{\Big[\max_{s\in[\tau-h, \tau]}(y_0(s)+x_0)-(y_0(\tau)+x_0)\Big]-$$

$$-\widehat{\Big[\max_{s\in[\tau-h, \tau]}(y_0(s)+x_0)-(y_0(\tau)+x_0)\Big]}\Big\}d\tau+\lambda x_0-\lambda x_0=$$

$$=\lambda[x_1(t, x_0)-x_0]-q\Big\{\Big[\max_{s\in[t-h, t]}(y_0(s)+x_0)-(y_0(t)+x_0)\Big]-$$

$$-\widehat{\Big[\max_{s\in[t-h, t]}(y_0(s)+x_0)-(y_0(t)+x_0)\Big]}\Big\}.$$

Тогда, используя оценки (9) и (10), находим

$$|\dot{x}_1(t, x_0)| \leqq |\lambda|\, h\beta|1+\lambda|\alpha(\lambda, t)+2h\beta|1+\lambda| =$$

(12)
$$= qh\beta[(1+q)\alpha(\lambda, t)+2] \leqq qh\beta\left[(1+q)\,\alpha\left(\lambda, \frac{T}{2}\right)+2\right].$$

Оценим теперь выражение $|x_2(t, x_0)-x_1(t, x_0)|$. Имеем

$$|x_2(t, x_0)-x_1(t, x_0)| \leqq q\frac{e^{\lambda(T-t)}-1}{e^{\lambda T}-1}\int_0^t e^{\lambda(t-\tau)}\times$$

$$\times\Big|\Big[\max_{s\in[\tau-h, \tau]}(y_0(s)+x_1(s))-(y_0(\tau)+x_1(\tau))\Big]-$$

$$-\Big[\max_{s\in[\tau-h, \tau]}(y_0(s)+x_0)-(y_0(\tau)+x_0)\Big]\Big|\,d\tau+$$

$$+q\frac{e^{\lambda t}-1}{e^{\lambda T}-1}\int_t^T e^{\lambda(T-\tau)}\Big|\Big[\max_{s\in[\tau-h, \tau]}(y_0(s)+x_1(s))-(y_0(\tau)+x_1(\tau))\Big]-$$

$$-\Big[\max_{s\in[\tau-h, \tau]}(y_0(s)+x_0)-(y_0(\tau)+x_0)\Big]\Big|\,d\tau.$$

Для дальнейших оценок нам понадобится помощная оценка. Рассмотрим выражение

$$F(y(t),\ x(t), t, h) = \max_{\tau\in[t-h, t]}[y(\tau)+x(\tau)]-\max_{\tau\in[t-h, t]}y(\tau)-x(t),$$

где $x(t)$ и $y(t)$ — произвольные непрерывные функции. Пусть t_1 и t_2 — точки интервала $[t-h, t]$, в которых достигаются максимумы функций $y(t)$ и $y(t)+x(t)$

соответственно. Тогда справедливо неравенство [5]

$$\left|F\big(y(t),\ x(t),\ t,\ h\big)\right| \leqq |x(t)-x(t_1)|+|x(t_2)-x(t_1)|. \tag{13}$$

Возьмем в (13) в качестве $x(t)$ — функцию $x_k(t)-x_{k-1}(t)$, $k=1, 2, 3, \ldots$, а в качестве $y(t)$ — функцию $y_0(t)+x_{k-1}(t)$. Тогда неравенство (13) перепишется в виде

$$\begin{aligned} &\Big|\max_{\tau\in[t-h,\,t]}[x_k(\tau)+y_0(\tau)]-\max_{\tau\in[t-h,\,t]}[x_{k-1}(\tau)+y_0(\tau)]-x_k(t)+x_{k-1}(t)\Big| \leqq \\ &\leqq \Big|\big(x_k(t)-x_{k-1}(t)\big)-\big(x_k(t_1^{(k)})-x_{k-1}(t_1^{(k)})\big)\Big|+ \\ &+\Big|\big(x_k(t_2^{(k)})-x_{k-1}(t_2^{(k)})\big)-\big(x_k(t_1^{(k)})-x_{k-1}(t_1^{(k)})\big)\Big|. \end{aligned} \tag{14}$$

Заметим еще, что при $t'\in[t-h,\ t]$ существует точка $t''\in(t-h,\ t)$, такая что

$$\begin{aligned} &\Big|\big(x_k(t)-x_{k-1}(t)\big)-\big(x_k(t')-x_{k-1}(t')\big)\Big| \leqq h\left|\frac{d}{dt}\big(x_k(t'')-x_{k-1}(t'')\big)\right| \leqq \\ &\leqq h\sup_t\left|\frac{d}{dt}\big(x_k(t)-x_{k-1}(t)\big)\right|, \qquad k=1, 2, 3, \ldots. \end{aligned} \tag{15}$$

Имея ввиду (15), получаем

$$\begin{aligned} &\Big|\max_{\tau\in[t-h,\,t]}[x_k(\tau)+y_0(\tau)]-\max_{\tau\in[t-h,\,t]}[x_{k-1}(\tau)+y_0(\tau)]-x_k(t)+x_{k-1}(t)\Big| \leqq \\ &\leqq 2h\sup_t\left|\frac{d}{dt}\big(x_k(t)-x_{k-1}(t)\big)\right|. \end{aligned} \tag{16}$$

Тогда из (16) и (12) следует оценка

$$\begin{aligned} &|x_2(t, x_0)-x_1(t, x_0)| \leqq 2hq\cdot qh\beta\left[(1+q)\,\alpha\left(\lambda, \frac{T}{2}\right)+2\right]\times \\ &\times\left\{\frac{e^{\lambda(T-t)}-1}{e^{\lambda T}-1}\int_0^t e^{\lambda(t-\tau)}\,d\tau+\frac{e^{\lambda t}-1}{e^{\lambda T}-1}\int_t^T e^{\lambda(T-\tau)}\,d\tau\right\} \leqq \\ &\leqq 2\beta h^2q^2\left[(1+q)\,\alpha\left(\lambda, \frac{T}{2}\right)+2\right]\alpha\left(\lambda, \frac{T}{2}\right). \end{aligned} \tag{17}$$

Найдем оценку для разности $\dot{x}_2(t, x_0)-\dot{x}_1(t, x_0)$. Имеем

$$\begin{aligned} &|\dot{x}_2(t, x_0)-\dot{x}_1(t, x_0)| \leqq |\lambda|\,|x_2(t, x_0)-x_1(t, x_0)|+ \\ &+q\Big|\Big[\max_{s\in[t-h,\,t]}\big(y_0(s)+x_1(s)\big)-\max_{s\in[t-h,\,t]}\big(y_0(s)+x_0\big)-\big(x_1(t)-x_0\big)\Big]- \\ &-\Big[\max_{s\in[t-h,\,t]}\big(y_0(s)+x_1(s)\big)-\max_{s\in[t-h,\,t]}\big(y_0(s)+x_0\big)-\big(x_1(t)-x_0\big)\Big]\Big| \leqq \\ &\leqq (1+q)\,2\beta h^2q^2\left[(1+q)\,\alpha\left(\lambda, \frac{T}{2}\right)+2\right]\alpha\left(\lambda, \frac{T}{2}\right)+4qh\sup_t\left|\frac{d}{dt}x_1(t)\right| \leqq \\ &\leqq 2\beta q^2h^2\left[(1+q)\,\alpha\left(\lambda, \frac{T}{2}\right)+2\right]^2. \end{aligned}$$

Итак,

(18) $$|\dot{x}_2(t, x_0)-\dot{x}_1(t, x_0)| \leqq 2\beta q^2 h^2 \left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]^2.$$

Допустим, что для некоторого $n \geqq 1$ выполняются неравенства

(19) $$|x_n(t, x_0)-x_{n-1}(t, x_0)| \leqq \frac{\beta}{2}(2hq)^n \left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]^{n-1} \alpha\left(\lambda, \frac{T}{2}\right),$$

(20) $$|\dot{x}_n(t, x_0)-\dot{x}_{n-1}(t, x_0)| \leqq \frac{\beta}{2}(2hq)^n \left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]^n.$$

Покажем, что тогда выполняются и неравенства

(21) $$|x_{n+1}(t, x_0)-x_n(t, x_0)| \leqq \frac{\beta}{2}(2hq)^{n+1} \left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]^n \alpha\left(\lambda, \frac{T}{2}\right),$$

(22) $$|\dot{x}_{n+1}(t, x_0)-\dot{x}_n(t, x_0)| \leqq \frac{\beta}{2}(2hq)^{n+1} \left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]^{n+1}.$$

Имеем

$$|x_{n+1}(t, x_0)-x_n(t, x_0)| \leqq 2hq\frac{\beta}{2}(2hq)^n \left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]^n \alpha\left(\lambda, \frac{T}{2}\right) =$$

$$= \frac{\beta}{2}(2hq)^{n+1} \left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]^n \alpha\left(\lambda, \frac{T}{2}\right),$$

т.е., неравенство (21) выполнено.

Для выражения $|\dot{x}_{n+1}(t, x_0)-\dot{x}_n(t, x_0)|$ находим

$$|\dot{x}_{n+1}(t, x_0)-\dot{x}_n(t, x_0)| \leqq (1+q)\frac{\beta}{2}(2hq)^{n+1}\times$$

$$\times\left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]^n \alpha\left(\lambda, \frac{T}{2}\right)+4hq\frac{\beta}{2}(2hq)^n\left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]^n =$$

$$= \frac{\beta}{2}(2hq)^{n+1}\left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]^n\left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right] =$$

$$= \frac{\beta}{2}(2hq)^{n+1}\left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]^{n+1}.$$

Таким образом, методом математической индукции доказано, что для всех $n \geqq 0$ справедливы оценки (21) и (22).

Для доказательства равномерной сходимости последовательностей $\{x_n(t, x_0)\}$ и $\{\dot{x}_n(t, x_0)\}$ оценим $|x_{n+k}(t, x_0)-x_n(t, x_0)|$ и $|\dot{x}_{n+k}(t, x_0)-\dot{x}_n(t, x_0)|$, где k — произвольное натуральное число, больше единицы. Находим

$$|x_{n+k}(t, x_0)-x_n(t, x_0)| \leqq \frac{\beta}{2}(2hq)^{n+1}\left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]^n \times$$

(23)
$$\times\alpha\left(\lambda, \frac{T}{2}\right)\frac{1-(2hq)^k\left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]^k}{1-2hq\left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]},$$

$$|\dot{x}_{n+k}(t, x_0)-\dot{x}_n(t, x_0)| \leqq \frac{\beta}{2}(2hq)^{n+1}\left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]^{n+1} \times$$

(24)
$$\times\frac{1-(2hq)^k\left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]^k}{1-2hq\left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]}.$$

Из (23) и (24) видно, что для равномерной сходимости последовательностей $\{x_n(t, x_0)\}$ и $\{\dot{x}_n(t, x_0)\}$ достаточно выполнение неравенства

(25)
$$2hq\left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right] < 1.$$

Переходя к пределу в (23) и (24) при $k\to\infty$ и при выполнение условия (25), находим оценки

(26)
$$|x_\infty(t, x_0)-x_n(t, x_0)| \leqq \frac{\frac{\beta}{2}(2hq)^{n+1}\left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]^n\alpha\left(\lambda, \frac{T}{2}\right)}{1-2hq\left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]},$$

(27)
$$|\dot{x}_\infty(t, x_0)-\dot{x}_n(t, x_0)| \leqq \frac{\frac{\beta}{2}(2hq)^{n+1}\left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]^{n+1}}{1-2hq\left[(1+q)\alpha\left(\lambda, \frac{T}{2}\right)+2\right]}.$$

Перейдем к вопросу существования периодического решения. Как было отмечено выше, нужно найти условия разрешимости (8). Обозначим левую часть (8) через $\Delta(x_0)$, а через $\Delta_n(x_0)$ — выражение

$$\Delta_n(x_0) = (1+q)x_0+q\Big[\overline{\max_{s\in[t-h,t]}\big(y_0(s)+x_n(s, x_0)\big)-\big(y_0(t)+x_n(t, x_0)\big)}\Big].$$

Разрешить уравнение $\Delta(x_0)=0$ в общем случае невозможно, поскольку не всегда можно найти предельную функцию $x_\infty(t, x_0)$. Поэтому мы поставим себе задачу как зная о существовании нулей функции $\Delta_n(x_0)$ судить о существовании нулей функции $\Delta(x_0)$. Найдем оценку для разности $\Delta(x_0)-\Delta_n(x_0)$. Имеем

$$|\Delta(x_0)-\Delta_n(x_0)| = q\left|\frac{\lambda}{e^{\lambda T}-1}\int_0^T e^{\lambda(T-\tau)}\Big[\max_{s\in[\tau-h,\tau]}\big(y_0(s)+\right.$$

$$+x_\infty(s,x_0)\big)-\big(y_0(\tau)+x_\infty(\tau,x_0)\big)\Big]\,d\tau-$$

$$\left.-\frac{\lambda}{e^{\lambda T}-1}\int_0^T e^{\lambda(T-\tau)}\Big[\max_{s\in[\tau-h,\tau]}\big(y_0(s)+x_n(s,x_0)\big)-\big(y_0(\tau)+x_n(\tau,x_0)\big)\Big]\,d\tau\right| \leqq$$

$$\leqq \frac{\beta}{2}\,\frac{(2hq)^{n+2}\left[(1+q)\,\alpha\left(\lambda,\frac{T}{2}\right)+2\right]^{n+1}}{1-2hq\left[(1+q)\,\alpha\left(\lambda,\frac{T}{2}\right)+2\right]}.$$

Из последней оценки следует, что $\Delta_n(x_0)\to\Delta(x_0)$ при $n\to\infty$.

Теорема 1. *Пусть выполнены следующие условия:*

(1) *На отрезке* $[a,b]$ *задано уравнение* (5).

(2) *Отрезок* $[a,b]$ *такой, что* $\dfrac{b-a}{2}>d_\lambda$.

(3) *Для некоторого целого числа* $n\geqq 1$ *функция* $\Delta_n(x_0)$ *удовлетворяет неравенствам*

$$\min_{a+d_\lambda\leqq x_0\leqq b-d_\lambda}\Delta_n(x_0)\leqq-\frac{\beta}{2}\,\frac{(2hq)^{n+2}\left[(1+q)\,\alpha\left(\lambda,\frac{T}{2}\right)+2\right]^{n+1}}{1-2hq\left[(1+q)\,\alpha\left(\lambda,\frac{T}{2}\right)+2\right]},$$

$$\max_{a+d_\lambda\leqq x_0\leqq b-d_\lambda}\Delta_n(x_0)\geqq\frac{\beta}{2}\,\frac{(2hq)^{n+2}\left[(1+q)\,\alpha\left(\lambda,\frac{T}{2}\right)+2\right]^{n+1}}{1-2hq\left[(1+q)\,\alpha\left(\lambda,\frac{T}{2}\right)+2\right]}.$$

(4) *Число* h *настолько мало, что выполняется неравенство* (25).

Тогда уравнение (5) *имеет периодическое решение* $x=x(t)$, *для которого* $x_0=x(0)\in[a+d_\lambda,\,b-d_\lambda]$, *являющееся пределом последовательности периодических функций* (7).

Доказательство. Пусть x_1 и x_2 точки отрезка $[a+d_\lambda,\,b-d_\lambda]$ такие, что $\Delta_n(x_1)=\min\Delta_n(x)$, $\Delta_n(x_2)=\max\Delta_n(x)$, $x\in[a+d_\lambda,\,b-d_\lambda]$.

Тогда

$$\Delta(x_1)=\Delta_n(x_1)+\big(\Delta(x_1)-\Delta_n(x_1)\big)\leqq 0,$$
$$\Delta(x_2)=\Delta_n(x_2)+\big(\Delta(x_2)-\Delta_n(x_2)\big)\geqq 0. \tag{28}$$

Из (28) в силу непрерывности функции $\Delta(x)$ следует существование точки $x^0\in[x_1,x_2]$, такой, что $\Delta(x^0)=0$, что и доказывает теорему 1.

Следствие. При выполнении условий теоремы, если функция $f_0(t)$ непрерывна и периодическая, то уравнение (1) имеет периодическое решение $y=y_0(t)+x_\infty(t, x_0)$, проходящее при $t=0$ через точку $x^0\in[a+d_\lambda, b-d_\lambda]$.

ЛИТЕРАТУРА

[1] Самойленко, А. М., Численно-аналитический метод исследования периодических систем обыкновенных дифференциальных уравнений, I, *Ukrain. Mat. Ž.* **17** (1965), no. 4, 82—93 (Samoĭlenko, A. M., A numerical-analytic method for investigation of periodic systems of ordinary differential equations, 1, in Russian). *MR* **33** # 3469a.

[2] Самойленко, А. М., Численно-аналитический метод исследования периодических систем обыкновенных дифференциальных уравнений, II, *Ukrain. Mat. Ž.* **18** (1966), no. 2, 50—59 (Title as above, 2, in Russian). *MR* **33** # 3469b.

[3] Самойленко, А. М и Ронто, Н. И., Численно-аналитические методы исследования периодических решений, Издат. объединение «Вища школа», Изд. при Киевском государственном университете, Киев, 1976 (Samoilenko, A. M. and Ronto, N. J., Numerical-analytic methods for the study of periodic solutions, in Russian; Izdat. Obe″d. Višča Škola" Izdat. pri Kiev. Gosudarstv. Univ., Kiev). *RŽ Mat.* **1976** # 11B1109K; *MR* **58** # 22817.

[4] Перестюк, Н. А., О периодических решениях некоторых систем дифференциальных уравнений, *Асимптотические и качественные методы в теории нелинейных колебаний*, Изд. Института математики АН УССР, Киев, 1971, 136—146 (Perest'juk, M. O., The periodic solutions of certain systems of differential equations, in Russian; Asymptotic and qualitative methods in the theory of nonlinear oscillations, Izdanie Inst. Mat. Akad. Nauk Ukrain. SSR, Kiev). *MR* **51** # 1022.

[5] Рябов, Ю Я. и Магомедов, А. Р., О периодических решениях линейных дифференциальных уравнений с максимумами (Rjabov, Ju. Ja. and Magomedov, A. R., Periodic solutions of linear differential equations with maxima, in Russian), *Mat. Fiz. Vyp.* **23** (1978), 3—9, 113. *MR* **58** # 17424.

(Поступила 29 августа 1980)

КАФЕДРА МАТЕМАТИЧЕСКОГО АНАЛИЗА
ПЛОВДИВСКИЙ УНИВЕРСИТЕТ
ИМЕНИ П. ХИЛЕНДАРСКОГО
ЦАР АСЕН 24
ПЛОВДИВ
BULGARIA

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 229—234

ON THE $\mathscr{P}_\Phi$-SPACES AND THE GENERALIZATION OF HERZ'S AND FEFFERMAN'S INEQUALITIES I

S. ISHAK and J. MOGYORÓDI

1. Introduction

The purpose of the present note is to generalize the notion of the $\mathscr{P}_p$-spaces. These new spaces are interesting in themselves but at the same time they will help us to extend the validity of the Herz and Fefferman—Garsia inequalities. This will be the subject of the forthcoming papers II and III.

2. Basic notions and definitions

Let $X \in L_1$ be a random variable defined on the probability space $(\Omega, \mathscr{A}, \mathsf{P})$ and consider the regular martingale

$$X_n = \mathsf{E}(X|\mathscr{F}_n), \quad n \geqq 0,$$

where $\{\mathscr{F}_n\}$, $n \geqq 0$, is an increasing sequence of σ-fields of events, such that

$$\sigma\left(\bigcup_{n=0}^{\infty} \mathscr{F}_n\right) = \mathscr{A}.$$

For the sake of the commodity we suppose that

$$X_0 = 0 \quad \text{a.e.}$$

Let $\Phi(x)$ be an arbitrary Young function, i.e. let

$$\Phi(x) = \int_0^x \varphi(t)\,dt, \quad x \geqq 0,$$

where $\varphi(t)$ is a right-continuous, nondecreasing function such that $\varphi(0)=0$ and $\lim_{t\to+\infty} \varphi(t) = +\infty$. We suppose that the power

$$p = \sup_{x>0} \frac{x\varphi(x)}{\Phi(x)}$$

1980 *Mathematics Subject Classification.* Primary 60G42; Secondary 60G46, 42B30, 42B99.

Key words and phrases. Martingale transforms, L_p-predictable martingales and the generalization of this notion, decomposition of martingales, Herz's inequality, Fefferman's inequality, Hardy spaces, the $\mathscr{K}$-spaces of Garsia and their generalizations.

of Φ is finite. This implies that the inverse function Φ^{-1} of Φ exists and has the form

$$\Phi^{-1}(x) = \int_0^x m(t)\,dt.$$

Here $m(t)$ is a decreasing function and we can easily see that

$$m(t) = 1/\varphi(\Phi^{-1}(t)), \quad t > 0.$$

To show these we remark that if p is finite then neither $\Phi(x)$ nor $\varphi(x)$ do vanish for $x>0$. Consequently, $\Phi(x)$ strictly increases. This and the continuity of Φ imply that inverse function Φ^{-1} exists, it is concave and is of the above integral form.

Let Φ be a Young function. We say that the random variable X belongs to the space $L^\Phi = L^\Phi(\Omega, \mathscr{A}, \mathsf{P})$ if there exists an $a>0$ such that the inequality $\mathsf{E}(\Phi(a^{-1}|X|)) \leqq 1$ is satisfied. In this case we put

$$\|X\|_\Phi = \inf\big(a > 0\colon\ \mathsf{E}(\Phi(a^{-1}|X|)) \leqq 1\big).$$

$\|\cdot\|_\Phi$ is a norm on L^Φ. The normed vector space L^Φ is complete.

For these definitions we refer to [2].

Now we shall introduce the notion of the $\mathscr{P}_\Phi$-spaces defined for arbitrary Young function Φ. These are the direct generalizations of the $\mathscr{P}_p$-spaces, were $p>1$ is a power (for the definition of $\mathscr{P}_p$, $p \geqq 1$, cf. e.g. Garsia [1]).

DEFINITION 2.1. A random variable $X \in L_1$ is L^Φ-predictable if there is a sequence $\{\lambda_n\}$ of random variables such that

(a) $|X_n| \leqq \lambda_{n-1}$ a.e., $n \geqq 1$,

(b) $\lambda_0 \leqq \lambda_1 \leqq \lambda_2 \leqq \dots$ a.e.,

(c) λ_n is $\mathscr{F}_n$-measurable, $n \geqq 0$,

(d) $\lim\limits_{n\to+\infty} \lambda_n = \lambda_\infty \in L^\Phi$.

We define $\mathscr{P}_\Phi$ as the class of L^Φ-predictable functions X with $X_0=0$ a.e. We take in this case

$$\|X\|_{\mathscr{P}_\Phi} = \inf_{\{\lambda_n\}} \|\lambda_\infty\|_\Phi,$$

where "inf" is taken over all the sequences $\{\lambda_n\}$ satisfying the preceding conditions.

The sequence $\{\lambda_n\}$ will be called an L^Φ-predicting sequence of $X \in \mathscr{P}_\Phi$.

It is easily seen that $\|\cdot\|_{\mathscr{P}_\Phi}$ is a norm on the space $\mathscr{P}_\Phi$.

REMARK 2.1. Suppose that $X \in \mathscr{P}_\Phi$ with

$$\|X\|_{\mathscr{P}_\Phi} = \inf_{\{\lambda_n\}} \|\lambda_\infty\|_\Phi,$$

where the sequences $\{\lambda_n\}$ satisfy the properties (a), (b), (c) and (d). Now it is clear that the "inf" is achieved in this definition. This is because if each $\{\lambda_n^{(k)}\}$ satisfies the conditions of Definition 2.1 and $\|\lambda_\infty^{(k)}\|_\Phi \to \|X\|_{\mathscr{P}_\Phi}$ as $k\to+\infty$, then setting

$$\lambda_n = \inf_k \lambda_n^{(k)}, \quad n \geqq 0,$$

$\{\lambda_n\}$ will not only satisfy the condition in that definition but also

$$\|X\|_{\mathscr{P}_\Phi} = \|\lambda_\infty\|_\Phi.$$

Such a $\{\lambda_n\}$ will be called an optimal L^Φ-predicting sequence of X.

3. Relations amongst the $\mathscr{P}_\Phi$- and the $\mathscr{P}_1$-spaces

The martingale transform techniques can be used to show that given any Young function Φ with finite power p the corresponding $\mathscr{P}_\Phi$-space is the martingale transform of $\mathscr{P}_1$ and conversely. More precisely, we can transform the martingale $(X_n, \mathscr{F}_n)$ corresponding to $X \in \mathscr{P}_1$ in such a way that the limit $X' = \lim_{n\to+\infty} X'_n$ belongs to $\mathscr{P}_\Phi$. Here $(X'_n, \mathscr{F}_n)$ is the martingale transform of $(X_n, \mathscr{F}_n)$. The converse result is trivially true.

THEOREM 3.1. *Let Φ be a Young function having finite power p. Suppose that $X \in \mathscr{P}_1$. Then the martingale transform*

$$X'_0 = 0 \quad \text{a.e.}, \quad X'_n = \sum_{i=1}^{n} d_i m(\lambda_{i-1}), \quad n \geqq 1,$$

converges a.e. to a limit X'. Here $\{d_i\}$ denotes the martingale difference sequence corresponding to X. Moreover, $X' \in \mathscr{P}_\Phi$ and

$$\|X\|_{\mathscr{P}_\Phi} \leqq \|\Phi^{-1}(\lambda_\infty)\|_\Phi,$$

where $\{\lambda_n\}$ is an optimal L_1-predicting sequence of X.

PROOF. The above martingale transform can also be written in the form

$$X'_n = \sum_{i=1}^{n} (X_i - X_{i-1}) m(\lambda_{i-1}) = X_n m(\lambda_{n-1}) + \sum_{i=1}^{n-1} X_i\big(m(\lambda_{i-1}) - m(\lambda_i)\big).$$

Consequently, with $\lambda_{-1} = 0$

$$|X'_n| \leqq \lambda_{n-1} m(\lambda_{n-1}) + \sum_{i=1}^{n-1} \lambda_{i-1}\big(m(\lambda_{i-1}) - m(\lambda_i)\big) = \sum_{i=1}^{n} m(\lambda_{i-1})(\lambda_{i-1} - \lambda_{i-2}).$$

It follows that

$$|X'_n| \leqq \int_0^{\lambda_{n-1}} m(\lambda)\, d\lambda \quad \text{a.e.},$$

or in other words

$$|X'_n| \leqq \Phi^{-1}(\lambda_{n-1}) \quad \text{a.e.}$$

Now the limit $X' = \lim_{n\to+\infty} X'_n$ exists. To show this it is enough to prove that $(X'_n, \mathscr{F}_n)$ is uniformly integrable. This follows from the inequality

$$|X'_n| \leqq \Phi^{-1}(\lambda_{n-1}) \leqq \Phi^{-1}(\lambda_\infty)$$

and from the fact that $\Phi^{-1}(\lambda_\infty)$ has finite expectation since Φ^{-1} is concave and $\mathsf{E}(\lambda_\infty) < +\infty$. It also follows that X' belongs to $\mathscr{P}_\Phi$. In fact, $\{\Phi^{-1}(\lambda_n)\}$, $n \geqq 0$, is an increasing sequence adapted to $\{\mathscr{F}_n\}$, and for $\Phi^{-1}(\lambda_\infty)$ we have $\mathsf{E}\big(\Phi(\Phi^{-1}(\lambda_\infty))\big) =$

$=\mathsf{E}(\lambda_\infty)$. Moreover, it also follows that

$$\|X'\|_{\mathcal{P}_\Phi} \leqq \|\Phi^{-1}(\lambda_\infty)\|_\Phi .$$

This proves the assertion.

REMARK 3.1. Especially, when $\Phi(x)=x^p$ with $p>1$ then $\Phi^{-1}(x)=x^{1/p}$ and

$$\|\Phi^{-1}(\lambda_\infty)\|_\Phi = \big(\mathsf{E}((\lambda_\infty^{1/p})^p)\big)^{1/p} = \big(\mathsf{E}(\lambda_\infty)\big)^{1/p} = (\|X\|_{\mathcal{P}_1})^{1/p}.$$

Consequently,

$$\|X'\|_{\mathcal{P}_p} \leqq \|X\|_{\mathcal{P}_1}^{1/p}.$$

So we obtained as a special case of our assertion the inequality of Garsia [1] (Theorem IV.4.1).

REMARK 3.2. We have

$$\|\Phi^{-1}(\lambda_\infty)\|_\Phi \leqq \max(1, \|X\|_{\mathcal{P}_1}).$$

In fact,

$$\mathsf{E}\left(\Phi\left(\frac{\Phi^{-1}(\lambda_\infty)}{\max(1, \|X\|_{\mathcal{P}_1})}\right)\right) \leqq \frac{1}{\max(1, \|X\|_{\mathcal{P}_1})} \mathsf{E}\big(\Phi(\Phi^{-1}(\lambda_\infty)\big) = \frac{\|X\|_{\mathcal{P}_1}}{\max(1, \|X\|_{\mathcal{P}_1})} \leqq 1.$$

REMARK 3.3. In Section 2 and 3 we have supposed that the Young function Φ has finite power p. This is only made for the sake of commodity, namely, to ensure that Φ^{-1} exist.

In the following assertion a lower estimate for $\|X'\|_{\mathcal{P}_\Phi}$ will be given.

THEOREM 3.2. *Let X and X' be as in Theorem 3.1. Then*

$$\frac{1}{4\|\varphi(\Phi^{-1}(\lambda_\infty)\|_\Psi} \|X\|_{\mathcal{P}_1} \leqq \|X'\|_{\mathcal{P}_\Phi},$$

where

$$\Psi(x) = \int_0^x \psi(t)\,dt$$

is the conjugate Young function of Φ, i.e. $\psi(t)=\varphi^{-1}(t)$.

PROOF. Denote by $\{\lambda_n\}$ and $\{\lambda_n'\}$ the corresponding optimal predicting sequences for X and X', respectively. From the representation of X_n' figuring in the proof of the preceding theorem we have

$$X_n - X_{n-1} = \frac{X_n' - X_{n-1}'}{m(\lambda_{n-1})}, \quad n \geqq 1.$$

(If $m(\lambda_{n-1})=0$ then we can add an $\varepsilon>0$ to each λ_n and at the end let $\varepsilon\to 0$.) Therefore,

$$X_n = \sum_{i=1}^n \frac{X_i' - X_{i-1}'}{m(\lambda_{i-1})} = \sum_{i=1}^n (X_i' - X_{i-1}')\varphi(\Phi^{-1}(\lambda_{i-1})) =$$

$$= X_n'\varphi(\Phi^{-1}(\lambda_{n-1})) - \sum_{i=1}^{n-1} X_i'\big(\varphi(\Phi^{-1}(\lambda_i)) - \varphi(\Phi^{-1}(\lambda_{i-1}))\big).$$

On the basis of this relation using the fact that $|X_i'| \leqq \lambda_{i-1}'$ we get

$$|X_n| \leqq 2\lambda_{n-1}' \varphi(\Phi^{-1}(\lambda_{n-1})).$$

Consequently, applying the generalized Hölder inequality[1] to the right-hand side we obtain

$$\|X\|_{\mathscr{P}_1} \leqq 4\|X'\|_{\mathscr{P}_\Phi} \|\varphi(\Phi^{-1}(\lambda_\infty))\|_\Psi,$$

which proves the assertion.

REMARK 3.4. Let $\Phi(x)=\dfrac{x^p}{p}$ with $1<p<+\infty$. Then $\varphi(x)=x^{p-1}$ and $\Psi(x)= =\dfrac{x^q}{q}$, $q=\dfrac{p}{p-1}$, while $\Phi^{-1}(x)=(px)^{1/p}$. Consequently,

$$\varphi(\Phi^{-1}(\lambda_\infty)) = (p\lambda_\infty)^{1/q}.$$

The L^Ψ-norm of this is

$$\left(\frac{p}{q}\right)^{1/q} (\mathsf{E}(\lambda_\infty))^{1/q} = \left(\frac{p}{q}\right)^{1/q} \|X\|_{\mathscr{P}_1}^{1/q}$$

and from the inequality of the preceding theorem we get

$$\frac{1}{4}\left(\frac{q}{p}\right)^{1/q} \|X\|_{\mathscr{P}_1}^{1/p} \leqq \frac{1}{p^{1/p}} \|X'\|_{\mathscr{P}_p}.$$

This is the inequality of Garsia (cf. [1], Theorem IV; 4.1).

REMARK 3.5. If p is the power of Φ then

$$\|\varphi(\Phi^{-1}(\lambda_\infty))\|_\Psi \leqq \max(1, (p-1)\|X\|_{\mathscr{P}_1}).$$

In fact,

$$\mathsf{E}\left(\Psi\left(\frac{\varphi(\Phi^{-1}(\lambda_\infty))}{\max(1, (p-1)\|X\|_{\mathscr{P}_1})}\right)\right) \leqq \frac{p-1}{\max(1, (p-1)\|X\|_{\mathscr{P}_1})} \mathsf{E}(\Phi(\Phi^{-1}(\lambda_\infty))) =$$

$$= \frac{(p-1)\|X\|_{\mathscr{P}_1}}{\max(1, (p-1)\|X\|_{\mathscr{P}_1})} \leqq 1.$$

Theorem 3.1 and Theorem 3.2 have the following

COROLLARY 3.1. *Let $X\in\mathscr{P}_1$ and let Φ be any Young function having finite power p. Then there is a random variable $X'\in\mathscr{P}_\Phi$ such that X_n is the martingale transform of X_n'. Namely, we have*

$$X_n = \sum_{i=1}^{n} (X_i' - X_{i-1}')T_{i-1}, \quad n \geqq 1, \quad X_0 = 0 \quad \text{a.e.},$$

where $T_i=\varphi(\Phi^{-1}(\lambda_i))$ and $\{\lambda_n\}$ is an optimal L_1-predicting sequence of X. We have

$$\|T_\infty\|_\Psi \leqq \max(1, (p-1)\|X\|_{\mathscr{P}_1})$$

[1] If $X\in L^\Phi$ and $Y\in L^\Psi$ then $XY\in L_1$ and $\mathsf{E}(|XY|)\leqq 2\|X\|_\Phi \|Y\|_\Psi$.

and

$$\|X'\|_{\mathscr{P}_\Phi} \leqq \|\Phi^{-1}(\lambda_\infty)\|_\Phi.$$

This assertion needs not to be proved.

We notice that the converse statement to this corollary is quite obvious.

REMARK 3.6. Corollary 3.1 points out that "the elements of the space $\mathscr{P}_\Phi$ are none other than the martingale transform ot those of $\mathscr{P}_1$ and viceversa".

REFERENCES

[1] GARSIA, A. M., *Martingale inequalities: Seminar notes on recent progress*, Mathematics Lecture Notes Series, W. A. Benjamin, Inc., Reading, Mass.—London—Amsterdam, 1973. *MR* 56 # 6844.

[2] NEVEU, J., *Discrete-parameter martingales*, North-Holland Mathematical Library, Vol. 10, North-Holland Publishing Co., Amsterdam—Oxford; American Elsevier Publishing Co., Inc., New York, 1975. *MR* 53 # 6729.

(Received October 27, 1980)

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

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 235—241

APPROXIMATION AND SHORT TIME PREDICTION OF ECONOMIC TIME SERIES BY SPLINE FUNCTIONS

MARGIT LÉNÁRD

Abstract

A spline algorithm for time series approximation is given. It is proved that the procedure yields the best polynomial approximation. The results are applied to interpolate and extrapolate economic processes observed in relatively short time units. The method gives a possibility for the computation of approximative intermediate values of the time series which increases the number of data.

Introduction

In this paper we construct a new method for interpolation and extrapolation of time series ([4]).

The essence of the method is that we fit a twice or once continuously differentiable spline function to the empirical data. This function takes the empirical values at the mesh of points and approximates the unknown function (further its derivatives approximate those of the unknown function). This approximation is best in the sense of polynomial approximation. Our basic idea is the same as that of Fawzy ([2]), but his procedure needs the values of the derivatives of the unknown function at the mesh of points.

The coefficients of the spline function can easily be computed by means of the differences of the given function values.

A very important area of applications is the short time extrapolation of short economic time series and the increase of the number of data by interpolation in order to apply other theoretical methods of time series and to construct econometrical models which demand a large number of data.

An advantage of our method is its simplicity and consistency in the economical sense, that is the linear relations between the time series are preserved.

The results are illustrated on numerical examples.

The interpolational method

Let $[a, b]\subset \mathbf{R}$ be an interval, $m>0$ an integer and let $x\colon [a, b]\to \mathbf{R}^m$ be a twice continuously differentiable function ($\mathbf{R}$ denotes the set of real numbers). Let further $a=t_0<\ldots<t_{N+1}=b$ be an equidistant subdivision of $[a, b]$, $t_{k+1}-t_k=h$ and let $x_k=x(t_k)$ $(k=0, 1, \ldots, N+1)$. Denote by Δ the difference operator: $\Delta x_k=x_{k+1}-x_k$. We are to find a function S with the properties:

1980 *Mathematics Subject Classification.* Primary 41A15; Secondary 90A20.
Key words and phrases. Spline functions, time series.

(i) it is a polynomial of minimal degree on each subinterval $[t_k, t_{k+1}]$ $(k=0, \ldots, N)$;

(ii) it is twice continuously differentiable at the points t_k $(k=1, \ldots, N)$;

(iii) its first and second derivatives approximate the first and second derivatives of x at the points t_k $(k=1, \ldots, N)$, that is,

$$S'(t_k) = \frac{1}{2h}(x_{k+1}-x_{k-1}),$$

$$S''(t_k) = \frac{1}{h^2}\Delta^2 x_{k-1},$$

further

$$S(t_k) = x_k \qquad (k = 0, 1, \ldots, N+1).$$

It is obvious that S is of degree at least five on each subinterval, hence for $t\in[t_{k-1}, t_k]$ let

$$S(t) = S_k(t) = x_k + \frac{1}{2h}(x_{k+1}-x_{k-1})(t-t_k) + \tag{1}$$
$$+\frac{1}{2h^2}\Delta^2 x_{k-1}(t-t_k)^2 + \sum_{i=3}^{5} a_i^{(k)}(t-t_k)^i.$$

The function S is twice continuously differentiable if and only if the functions S_k satisfy the conditions

$$S_k^{(j)}(t_k) = S_{k+1}^{(j)}(t_k)$$

$(j=0, 1, 2;\ k=2, \ldots, N-1)$. This system of equations in $a_i^{(k)}$ $(i=3, 4, 5)$ can be solved and the solution is unique:

$$a_3^{(k)} = -\frac{3}{2h^3}\Delta^3 x_{k-2},$$

$$a_4^{(k)} = -\frac{5}{2h^4}\Delta^3 x_{k-2}, \tag{2}$$

$$a_5^{(k)} = -\frac{1}{h^5}\Delta^3 x_{k-2}.$$

THEOREM. *If $x\colon [a, b]\to \mathbf{R}^m$ is a twice continuously differentiable function, then for the spline function S of the form* (1) *with the coefficients* (2) *the following relations hold:*

$$\|x(t) - S(t)\| \leqq 6.5\, h^2\omega_2(h; x),$$

$$\|x'(t) - S'(t)\| \leqq 23\, h\omega_2(h; x),$$

$$\|x''(t) - S''(t)\| \leqq 68\,\omega_2(h; x),$$

$(t_1 \leqq t \leqq t_N)$. *Here $\omega_2(h; x)$ denotes the modulus of continuity of x''.*

PROOF. Let $t_{k-1} \leqq t \leqq t_k$ $(k=1, \ldots, N)$, then

$$x(t) = x_k + x_k'(t-t_k) + \frac{1}{2} x''(\tau_k)(t-t_k)^2 \tag{3}$$

where $t < \tau_k < t_k$, further

$$x_k^*(t) = x_k + x_k'(t-t_k) + \frac{1}{2} x_k''(t-t_k)^2$$

where $x_k' = x'(t_k)$, $x_k'' = x''(t_k)$. Then we have

$$\|x(t) - S(t)\| \leqq \|x(t) - x_k^*(t)\| + \|x_k^*(t) - S_k(t)\| \leqq$$

$$\leqq \frac{1}{2} \|x''(\tau_k) - x_k''\| h^2 + \left\| x_k' - \frac{1}{2h}(x_{k+1} - x_{k-1}) \right\| h +$$

$$+ \frac{1}{2} \left\| x_k'' - \frac{1}{h^2} \Delta^2 x_{k-1} \right\| h^2 + 5 \|\Delta^3 x_{k-2}\|.$$

Substituting t_{k-1} and t_{k+1} into (3) one can easily see that

$$\left\| x_k' - \frac{1}{2h}(x_{k+1} - x_{k-1}) \right\| \leqq \frac{h}{2} \omega_2(h; x),$$

$$\left\| x_k'' - \frac{1}{h^2} \Delta^2 x_{k-1} \right\| \leqq \omega_2(h; x).$$

Further by the Lagrange theorem we have

$$\|\Delta^3 x_{k-2}\| \leqq h^2 \omega_2(h; x).$$

Using these estimations we obtain

$$\|x(t) - S(t)\| \leqq 6.5 h^2 \omega_2(h; x).$$

If $t_{k-1} \leqq t \leqq t_k$ $(k=1, \ldots, N)$, we apply the Taylor formula for x':

$$x'(t) = x_k' + x''(\tau_k)(t-t_k)$$

where $t < \tau_k < t_k$. Similarly to the above we have

$$\|x'(t) - S'(t)\| \leqq \|x''(\tau_k) - x_k''\| h + \left\| x_k' - \frac{1}{2h}(x_{k+1} - x_{k-1}) \right\| +$$

$$+ \left\| x_k'' - \frac{1}{h^2} \Delta^2 x_{k-1} \right\| h + \frac{41}{2h} \|\Delta^3 x_{k-2}\| \leqq 23 h \omega_2(h; x).$$

Finally, for $t_{k-1} \leqq t \leqq t_k$,

$$\|x''(t) - S''(t)\| \leqq \|x''(t) - x_k''\| + \|x_k'' - S_k''(t)\| \leqq 68 \omega_2(h; x).$$

This interpolational method is stable, as shown by the next theorem.

THEOREM. *Let* $x_k, \bar{x}_k \in \mathbf{R}^m$ *be given and* $\|x_k - \bar{x}_k\| \leqq \varepsilon$ $(k=0, 1, \ldots, N+1)$. *Let* S *and* $\bar{S}$ *denote the spline functions of the form* (2) *with the coefficients* (1) *and elements* x_k, $\bar{x}_k$, *respectively. Then for any* $t \in [t_1, t_N]$ *we have*

$$\|S(t) - \bar{S}(t)\| \leqq 44\varepsilon.$$

This theorem can be verified by easy computation.

Let now $x\colon [a, b] \to \mathbf{R}^m$ be a continuously differentiable function and define the function S as follows: for $t_{k-1} \leqq t \leqq t_k$ $(k=2, \ldots, N+1)$

$$S(t) = S_k(t) = x_k + \frac{1}{h} \Delta x_{k-1}(t-t_k) - \frac{1}{h^2} \Delta^2 x_{k-2}(t-t_k)^2 - \frac{1}{h^3} \Delta^2 x_{k-2}(t-t_k)^3. \tag{4}$$

It is obvious that S is continuously differentiable.

THEOREM. *Let* $x\colon [a, b] \to \mathbf{R}^m$ *be a continuously differentiable function. For the spline function* S *of the form* (4) *we have*

$$\|x(t) - S(t)\| \leqq 4h\, \omega_1(h; x),$$

$$\|x'(t) - S'(t)\| \leqq 7\omega_1(h; x)$$

$(t_1 \leqq t \leqq t_{N+1})$, *where* $\omega_1(h; x)$ *denotes the modulus of continuity of* x'.

PROOF. As the function x is continuously differentiable, for $t_{k-1} \leqq t \leqq t_k$ $(k=1, \ldots, N)$ we have

$$x(t) = x_k + x'(\tau_k)(t-t_k)$$

where $t < \tau_k < t_k$, further put

$$x_k^*(t) = x_k + x_k'(t-t_k)$$

where $x_k' = x'(t_k)$. Then we have

$$\|x(t) - S(t)\| \leqq \|x(t) - x_k^*(t)\| + \|x_k^*(t) - S_k(t)\| \leqq$$

$$\leqq \|x'(\tau_k) - x_k'\| h + \left\| x_k' - \frac{1}{h} \Delta x_{k-1} \right\| h + 2\|\Delta^2 x_{k-2}\|.$$

By the Lagrange theorem we have

$$\left\| x_k' - \frac{1}{h} \Delta x_{k-1} \right\| = \|x_k' - x'(\bar{\tau}_k)\| \leqq \omega_1(h; x),$$

$$\|\Delta^2 x_{k-2}\| = h \|\Delta x'(\theta_{k-1})\| \leqq h\omega_1(h; x),$$

where $t_{k-1} < \bar{\tau}_k < t_k$, $t_{k-2} < \theta_{k-1} < t_{k-1}$.

Using these estimations we obtain

$$\|x(t) - S(t)\| \leqq 4h\omega_1(h; x).$$

Finally, for $t_{k-1} \leqq t \leqq t_k$,

$$\|x'(t) - S'(t)\| \leqq \|x'(t) - x_k'\| + \|x_k' - S_k'(t)\| \leqq 7\omega_1(h; x).$$

This interpolational method is also stable for the following theorem can be verified by easy computation.

THEOREM. *Let* $x_k, \bar{x}_k \in \mathbf{R}^m$ *be given and* $\|x_k - \bar{x}_k\| \leqq \varepsilon$ $(k=0, 1, \ldots, N+1)$. *Let* S *and* $\bar{S}$ *denote the spline functions of the form* (4) *with the elements* x_k, $\bar{x}_k$, *respectively. Then for any* $t \in [t_1, t_{N+1}]$ *we have*

$$\|S(t) - \bar{S}(t)\| \leqq 11\varepsilon.$$

The consistency of the method in the economical sense follows from the linearity of the construction and of the Δ operator. This means the preservation of the linear relations between the original time series.

Now we apply our method to the interpolation of the time series describing the national income of the Hungarian national economy. We used the data of the Hungarian Central Statistical Office for 1970—1975 (all data are given in 10^6 Ft units, 20 Ft = 1 US \$). We approximate by interpolating data of 3, 6 and 9 months, respectively. The interpolation of degree five was performed from 1971 until 1974 by (1), (2), and the interpolation of degree three was done from 1971 until 1975 by (4).

Year	Interpolation of degree 5	Annual data	Interpolation of degree 3
1970		303 258	
1971		321 010	
	325 766 330 600 335 628		325 689 330 712 335 873
1972		340 965	
	346 671 352 834 359 052		346 391 352 442 358 743
1973		364 919	
	370 489 375 773 381 050		370 659 376 043 381 284
1974		386 597	
			392 239 398 200 404 287
1975		410 311	

The extrapolational method

The polynomial S_k of the form (4) with the values x_{k-2}, x_{k-1}, x_k can be used for the extrapolation of the value at the next quarter year $t_k+\frac{h}{4}$:

$$x\left(t_k+\frac{h}{4}\right) \approx S_k\left(t_k+\frac{h}{4}\right).$$

This value and the previous two — interpolated — values enable us to extrapolate the value in the next half year in the following way: instead of the points t_{k-2}, t_{k-1}, t_k we use the points $t_{k-2}+\frac{h}{4}$, $t_{k-1}+\frac{h}{4}$ and $t_k+\frac{h}{4}$ and the approximated (interpolated or extrapolated) function values at these points. The spline function S of the form (4) with these values approximates x at $t_k+\frac{h}{2}$. Continuing this process we can extrapolate the function x for about two years. The estimation of the error will show that it is not worth to extrapolate for more than two years.

If $x\colon [a, b] \to \mathbf{R}^m$ is continuously differentiable, $x_i = x(t_i)$ $(i=k-2, k-1, k)$, then the error of extrapolation in the case $h=1$ is

$$\left\| x\left(t_k+\frac{1}{4}\right) - S_k\left(t_k+\frac{1}{4}\right) \right\| \leq \frac{37}{64}\,\omega_1(1; x).$$

If we use the values $\bar{x}_i$ $(i=k-2, k-1, k)$ instead of x_i, where $\|x_i-\bar{x}_i\| \leq \varepsilon_i$, then

$$\left\| S_k\left(t_k+\frac{1}{4}\right) - \bar{S}_k\left(t_k+\frac{1}{4}\right) \right\| \leq \frac{75}{64}\,\varepsilon_k + \frac{3}{32}\,\varepsilon_{k-1} + \frac{5}{64}\,\varepsilon_{k-2}.$$

This estimation and the error of interpolation enable us to compute the error of extrapolation. The error is of the form $c\omega_1(1; x)$, where c is a constant (error coefficient).

In the next table $h=1$, and the formula (4) was used for interpolation:

Year	Interpolation	Extrapolation	Error coeff.
1973	364 919		
	370 659	370 595	0.578
	376 043	376 372	1.544
	381 284	382 183	2.592
1974	386 597	387 999	3.710
	392 239	393 837	4.926
	398 200	399 723	6.540
	404 287	405 605	8.476
1975	410 311	411 462	10.797

REFERENCES

[1] AHLBERG, J. H., NILSON, E. N. and WALSH, J. L., *The theory of splines and their applications*, Academic Press, New York—London, 1967. *MR* **39** # 684.
[2] FAWZY, T., Spline functions and the Cauchy problems, III., *Ann. Univ. Sci. Budapest. Eötvös Sect. Computatorica* 1 (1978), 35—45. *MR* **83f**: 65107b
[3] LÉNÁRD, M., Kétváltozós spline-interpoláció (Spline interpolation in two variables), Doctoral dissertation, Debrecen, 1979 (in Hungarian).
[4] Népgazdasági mutatók éves adataiból negyedéves adatok becslése és két éves extrapolációja, Report made by the Mathematical Institute of the Lajos Kossuth University for the National Planning Office (directed by Z. Daróczy) (in Hungarian).

(Received November 26, 1980)

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

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 243—250

THE KUROŠ—ORE THEOREM, FINITE AND INFINITE DECOMPOSITIONS

GERD RICHTER

1. Introduction

As we can see in the papers Richter [9], Richter—Stern [10], [11] and Stern [12], [13], [14], the Kuroš—Ore theorem (Kuroš [7], Ore [8]) holds not only in modular lattices. However, the proofs are similar in all these cases. In the present paper we define a class of lattices having a property (V) in which the Kuroš—Ore theorem holds and in which the proof of this theorem is similar to the modular case. In fact, this property (V) was made use of in all the above mentioned proofs. Moreover, we show that in this class of lattices the same replacement property holds as in modular lattices (Dilworth [5]). Further we exhibit that some results by Crawley [2], Crawley—Dilworth [3], Dilworth [4], [5] and Dilworth—Crawley [6] can be proved under weaker conditions. If we mention results of Crawley or Dilworth, we shall refer to the well-known book [3].

2. Basic notions

Let L be a lattice. If L contains a largest or a least element, these elements will be denoted by 1 or 0, respectively. For two elements $a, b \in L$ we define $b/a := \{x : a \leqq x \leqq b\}$. An element $q \in b/a$ is called *inaccessible in b/a* iff $q = \bigvee T$ with $T \subseteqq b/a$ implies $q = \bigvee T'$ with $T' \subseteqq T$ and $|T'| < \infty$ (cf. Birkhoff—Frink [1]).

$Q(b/a)$ is defined to be the set of all inaccessible elements of b/a. An element $v \in b/a$ is called *join-irreducible in b/a* iff, for all $x, y \in b/a$, $v = x \vee y$ implies $x = v$ or $y = v$.

An element $v \in b/a$ is called *completely join-irreducible in b/a* iff, for all $T \subseteqq b/a$, $v = \bigvee T$ implies $v \in T$.

If $V(b/a)$ denotes the set of all join-irreducible elements of b/a then $V_1(b/a)$ denotes the set of all completely join-irreducible elements of b/a. Notice that $V_1(b/a) = V(b/a) \cap Q(b/a)$.

Further, we define $Q := Q(L)$, $V := V(L)$, $V_1 := V_1(L)$.

A lattice in which every element is a join of elements of Q or of V_1, respectively, is called a *Q-lattice* or a *V_1-lattice*, respectively. We say that L satisfies the property (V) iff $v \in V$, $b \in L$ and $b = u_1 \vee \ldots \vee u_n$ with $u_i \in V$ for $i = 1, \ldots, n$ imply that $b \vee v \in V(b \vee v/b)$. L satisfies the property (V_1) iff $v \in V_1$ and $b \in L$ imply $b \vee v \in V_1(b \vee v/b)$. Note a modular lattice satisfies (V) as well as (V_1). If a is an element of the lattice L, then a representation $a = \bigvee T$ with $T \subseteqq V$ in Section 3 and $T \subseteqq V_1$ in Section 4 is called a *(join) decomposition of a*.

1980 *Mathematics Subject Classifications.* Primary 06B05.

Key words and phrases. Irredundant decompositions, replacement properties, Kuroš—Ore property, join-irreducible and completely join-irreducible elements, lower continuous lattices.

A decomposition $a = \bigwedge R$ of a in the dual of L is called a *meet decomposition of a*.

A decomposition $a = \bigvee T$ is called *irredundant* iff, for each $t \in T$, $a > \bigvee (T \setminus \{t\})$ holds.

A lattice L is said to have the property (R) iff, whenever an $a \in L$ has two decompositions $a = v_1 \vee \ldots \vee v_n = u_1 \vee \ldots \vee u_m$ then for each v_i there is a u_j such that

$$a = v_1 \vee \ldots \vee v_{i-1} \vee u_j \vee v_{i+1} \vee \ldots \vee v_n.$$

If $a = \bigvee T$ is a decomposition of a, we define

$$\overline{t_{i,j,\ldots,n}} := \bigvee (T \setminus \{t_i, t_j, \ldots, t_n\})$$

for each subset $\{t_i, t_j, \ldots, t_n\}$ of T.

We say that a complete lattice L has *replaceable irredundant decompositions* iff each element of L has at least one irredundant decomposition and whenever $a = \bigvee T = \bigvee R$ are two irredundant decompositions of an element $a \in L$, for each $t_0 \in T$ there exists an $r_0 \in R$ such that $a = \overline{t_0} \vee r_0$, and this resulting decomposition is irredundant.

A finite lattice L has the *Kuroš—Ore property* iff, whenever $a \in L$ has an irredundant decomposition $a = v_1 \vee \ldots \vee v_n$, then for every irredundant decomposition $a = \bigvee T$ always $|T| = n$ holds.

3. Finite decompositions

Crawley—Dilworth ([3], p. 39) mentioned that if a lattice L satisfies the ascending chain condition then every element of L has an irredundant finite meet decomposition. Therefore, every element of L has an irredundant finite (join) decomposition if L satisfies the descending chain condition. But a simple example shows that the converse does not hold.

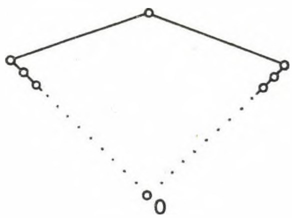


Fig. 1

Our first theorem is a generalization of [3] (Theorem 5.2) and [15] (Hilfssatz p. 112).

THEOREM 1. *Every lattice satisfying* (V) *has the property* (R).

The proof is the same as in [3] or [15], since in presence of (V) we do not need modularity.

Repeated application of (R) yields the Kuroš—Ore theorem in lattices having (R) and therefore by Theorem 1 also in lattices having (V):

If L has property (R), *then the number of join-irreducible elements is the same in every irredundant finite decomposition of an element $a \in L$.*

COROLLARY 2. *If L has* (R), $a=u_1\vee\ldots\vee u_n$ *is a decomposition and* $a=v_1\vee\ldots\vee v_n$ *is an irredundant decomposition of a, then* $a=u_1\vee\ldots\vee u_n$ *is an irredundant decomposition.*

The proof is obvious.

As in the modular case, we can ask if it is true that each v_i can replace some u_j in the decomposition $a=v_1\vee\ldots\vee v_n=u_1\vee\ldots\vee u_n$. Moreover, we can ask whether to each v_i there exists an element u_j with $a=u_j\vee\overline{v_i}=\overline{u_j}\vee v_i$. We are not able to prove this in a similar way as in [3] (Theorem 5.3).

Theorem 4 shows that modularity can be dropped if the lattice L has property (R). Before proving this, let us answer the first question as follows.

COROLLARY 3. *If L has property* (R) *and $a\in L$ has two irredundant decompositions* $a=v_1\vee\ldots\vee v_n=u_1\vee\ldots\vee u_n$, *then to each v_i there exists an element u_j with* $a=\overline{u_j}\vee v_i$.

PROOF. We choose an arbitrary v_i. Applying (R) we replace the elements $v_1, \ldots$ $\ldots, v_{i-1}, v_{i+1}, \ldots, v_n$ in $n-1$ steps by elements $u_{j_1}, \ldots, u_{j_{i-1}}, u_{j_{i+1}}, \ldots, u_{j_n}$ and get $a=\overline{u_{j_i}}\vee v_i$.

THEOREM 4 (cf. [3], Theorem 5.3). *If the lattice L has property* (R) *and if* $a=v_1\vee\ldots\vee v_n=u_1\vee\ldots\vee u_n$ *are two irredundant finite decompositions of* $a\in L$, *then*

(1) *for each v_i there is a u_j such that* $a=\overline{u_j}\vee v_i=u_j\vee\overline{v_i}$;

(2) *there is a permutation π of the integers* $1, \ldots, n$ *such that, for each* $i=1, \ldots, n$, $a=\overline{v_i}\vee u_{\pi(i)}$.

PROOF. We choose an arbitrary v_i. Suppose the u_j's are rearranged in such a way that $a=\overline{u_j}\vee v_i$ for $j=1, \ldots, m$ and

(a) $$a > \overline{u_j}\vee v_i \quad \text{for} \quad j = m+1, \ldots, n.$$

Corollary 3 yields that $m\geqq 1$. By applying (R) we can replace $u_{m+1}, \ldots, u_n$ by $v_{j_{m+1}}, \ldots, v_{j_n}$ and we get

$$a = u_1\vee\ldots\vee u_m\vee v_{j_{m+1}}\vee\ldots\vee v_{j_n} = u_1\vee\ldots\vee u_n.$$

If $v_i=v_{j_r}$ for some r with $m+1\leqq r\leqq n$, we apply (R) and replace the v_{j_t}'s $(t\neq r)$ by the elements $u_{i_{m+1}}, \ldots, u_{i_{r-1}}, u_{i_{r+1}}, \ldots, u_{i_n}$ and get $a=\overline{u_{i_r}}\vee v_i$ with $m<i_r$, contrary to (a). Therefore $v_i\neq v_{j_r}$ for $r=m+1, \ldots, n$, i.e.,

(b) $$a > \overline{v_i} = \overline{v_i}\vee v_{j_r}.$$

Now we apply (R) to the two decompositions

$$a = v_1\vee\ldots\vee v_n = u_1\vee\ldots\vee u_m\vee v_{j_{m+1}}\vee\ldots\vee v_{j_n}$$

and to the element v_i. By (b) we get the existence of an element u_j with $1\leqq j\leqq m$ and $a=u_j\vee\overline{v_i}$.

As in the modular case, (2) is an easy consequence of (1).

As we can see in Theorem 1 and in the Kuroš—Ore theorem, condition (V) is sufficient for a finite lattice to have the Kuroš—Ore property. In the following theorem we shall show that this condition is necessary and sufficient for a lattice of finite length to have replaceable irredundant decompositions. We can get this theorem also as a corollary of [3] (Theorem 7.5).

THEOREM 5. *A lattice L of finite length has replaceable irredundant decompositions iff it satisfies* (V).

PROOF. Since a lattice of finite length satisfies the descending chain condition, every element of L has a finite irredundant decomposition. If L satisfies (V), Theorem 1 and Corollary 2 yield that L has replaceable irredundant decompositions. Suppose now that L is a lattice of finite length which has replaceable irredundant decompositions. If L does not satisfy (V), then there exists an element $a \in L$ and an element $v \in V$ with $b = a \vee v \notin V(b/a)$. Therefore, there exist two elements $b_1, b_2 \in b/a$ with $b_1, b_2 < b = b_1 \vee b_2$. Since a, b_1, b_2 have finite decompositions $a = u_1 \vee \ldots \vee u_m$, $b_1 = w_1 \vee \ldots \vee w_r$ and $b_2 = w_{r+1} \vee \ldots \vee w_n$, we get

$$b = v \vee u_1 \vee \ldots \vee u_m = w_1 \vee \ldots \vee w_n.$$

By deleting all superfluous u_i's and all superfluous w_j's and by rearranging, we get the two irredundant decompositions

$$b = v \vee u_1 \vee \ldots \vee u_k = w_1 \vee \ldots \vee w_h.$$

Since L has replaceable irredundant decompositions, there is an element w_j such that $b = v \vee u_1 \vee \ldots \vee u_k = w_j \vee u_1 \vee \ldots \vee u_k$. But $w_j \vee u_1 \vee \ldots \vee u_k \leqq w_j \vee a \leqq b_i < b$ with $1 \leqq i \leqq 2$. Therefore L has property (V).

The following theorem describes a class of finite lattices which have the Kuroš—Ore property. (V) will not be used.

THEOREM 6. *If a finite lattice L has a planar diagram then it has the Kuroš—Ore property, and every irredundant decomposition of an element of L contains at most two elements.*

PROOF. Let L have a planar diagram and let a be an arbitrary element of L with $a = v_1 \vee \ldots \vee v_n$ such that v_i and v_j are incomparable for $i \neq j$. Now we can assume that in the planar diagram v_1 is on the left-hand side of all v_i and v_n is on the right-hand side of all v_i. If $v_1 \vee v_n < a$, then there is an element v_i in the middle with $v_1 \vee v_n < v_1 \vee v_n \vee v_i$. But then the diagram is not planar. Therefore, if a is not join-irreducible, every irredundant decomposition of a contains exactly two elements.

Notice that a finite lattice in which every irredundant decomposition of any element contains at most two elements, need not have a planar diagram, as shown by

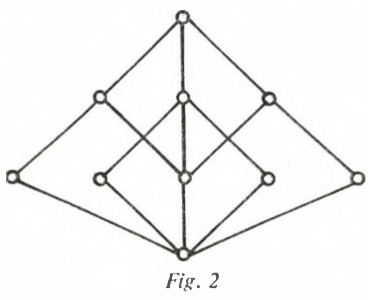


Fig. 2

4. Infinite decompositions

For investigations of infinite irredundant decompositions, property (V) does not yield anything. Therefore we shall use property (V_1). Note that in a V_1-lattice property (V_1) is equivalent to the following property, defined in [3] (p. 53):

($*$) for all $x, y \in L$, if the sublattice $x/x \wedge y$ has exactly one dual atom, then the sublattice $x \vee y/y$ has exactly one dual atom.

Remark further that in [3] (Theorem 6.1) it is shown that the dual of an algebraic lattice is always a V_1-lattice. An algebraic lattice is always *upper continuous*, i.e., L is complete and for every $a \in L$ and for every chain $C \subseteq L$, $a \wedge \bigvee C = \bigvee (a \wedge c\colon c \in C)$.

The dual of an upper continuous lattice is called *lower continuous*. Most of the investigations in this section will concern lower continuous V_1-lattices.

Our first theorem in this section is closely connected with Theorem 1.

THEOREM 7 (cf. [3], Theorem 7.2). *If an element a in a complete lattice L satisfying* (V_1) *has two decompositions* $a = \vee T = \vee R$, *then for each* $t_0 \in T$ *there exists an* $r_0 \in R$ *such that* $a = r_0 \vee \overline{t_0}$. *Moreover, the resulting decomposition is irredundant if the decomposition* $a = \vee T$ *is such.*

PROOF. Let t_0 be an arbitrary element of T. We set $S := \{s\colon s = \overline{t_0} \vee r,\ r \in R\}$. By ($V_1$) $a = t_0 \vee \overline{t_0}$ is completely join-irreducible in $a/\overline{t_0}$. Consequently, from $a = \vee R = \overline{t_0} \vee \bigvee R = \vee S$ and $S \subseteq a/\overline{t_0}$ we get $a \in S$, i.e., $a = s_0 = \overline{t_0} \vee r_0$ with $r_0 \in R$. It remains to prove that the irredundance of $a = \vee T$ implies the irredundance of $a = t_0 \vee r_0$. If $a = \vee (T \setminus \{t_0\}) \vee r_0$ is not an irredundant decomposition, then there is an element $t_1 \in T \setminus \{t_0\}$ with $a = \overline{t_{0,1}} \vee r_0$. Since $a = \vee T = \overline{t_{0,1}} \vee r_0$ are two decompositions of a, we get by the first part of the proof the existence of an element $t_i \in T$ such that $a = \bigvee T = \overline{t_{0,1}} \vee r_0 = \overline{t_{0,1}} \vee t_i$ with $(T \setminus \{t_0, t_1\}) \cup \{t_i\} \subset T$, contrary to the irredundance of $a = \vee T$. Therefore the proof is complete.

For our following investigations we need the concept of a strongly dually atomic lattice. We say that a lattice L is *strongly dually atomic* iff $a, b \in L$ and $a < b$ imply the existence of an element $p \in b/a$ covered by b. A lattice is strongly dually atomic iff its dual is strongly atomic. The next theorem is a generalization of [3] (Theorem 6.3).

THEOREM 8. *If every element of a lower continuous V_1-lattice L satisfying* (V_1) *has an irredundant decomposition, then L is strongly dually atomic.*

PROOF. Let a be an arbitrary element of L. Then, by the assumption, a has an irredundant decomposition $a = \vee T$. If $0 < a$ then $\overline{t_0} < a$ for any $t_0 \in T$. By (V_1) a is a completely join-irreducible element in $a/\overline{t_0}$. Since L is complete, there exists exactly one element $p \in a/\overline{t_0}$ covered by a, i.e., for each $a \in L$ with $a > 0$ there exists a $p \in a/0$ covered by a. Suppose now $b < a$ and let U be the set of all $u \in L$ with $b \vee u = a$. U is nonempty, since $a \in U$. Let C be a chain in U. Then lower continuity yields $b \vee \bigwedge C = \bigwedge (b \vee c\colon c \in C) = a$, i.e., every chain of U has a minimal element. An application of Zorn's lemma yields the existence of an element $u_0 \in U$ which is minimal in U. Since there exists an element r covered by u_0, we get $b \vee r < a = b \vee u_0$, in particular, $u_0 \not\leqq b \vee r$. By (V_1) $a = b \vee r \vee u_0 = b \vee r \vee v$ with $r \vee u_0 = r \vee v$ and $v \in V_1(u_0/0)$ is completely join irreducible in $a/b \vee r$. Consequently, there exists an element $p \in a/b$ with $b \vee r \leqq p \prec a$. Thus L is strongly dually atomic and the proof is complete.

In the proof of [3] (Theorem 6.4) it was not used that L is an algebraic lattice, only that L is upper continuous. Therefore, Lemma 9 and Theorem 10 below can be proved analogously, and their proofs will be omitted.

LEMMA 9 (cf. [3], p. 45). *Every element of a lower continuous V_1-lattice L has an irredundant decomposition iff, for each element $a\in L$ distinct from 0, there exists a $v\in V_1$ and an element $x\prec a$ such that $a=x\vee v$.*

THEOREM 10 (cf. [3], Theorem 6.4). *If a lower continuous V_1-lattice L is strongly dually atomic, then every element of L has an irredundant decomposition.*

The following result is a slight generalization of [3] (Theorem 7.5). We can get Theorem 5 as a corollary of this result.

THEOREM 11. *A strongly dually atomic lower continuous V_1-lattice L has replaceable irredundant decompositions iff it satisfies* (V_1).

PROOF. Since L is a strongly dually atomic lower continuous V_1-lattice, every element of L has irredundant decompositions by Theorem 10. These decompositions are replaceable irredundant decompositions by Theorem 7, since L satisfies (V_1).

Let us assume that L has replaceable irredundant decompositions but does not satisfy (V_1). Then there is an element $a\in L$ and an element $v\in V_1$ such that $b=a\vee v$ is not completely join-irreducible in b/a, i.e., there are elements p_1, $p_2\in b/a$ which are covered by b and it holds $a\vee v=(p_1\wedge p_2)\vee v=p_1\vee p_2=b$, since L is strongly dually atomic. Lower continuity yields that there is a minimal element $w\leqq p_1\wedge p_2$ with respect to $b=w\vee v$. w has an irredundant decomposition $w=\bigvee R$.

Then $v\vee\bigvee R=b$ is an irredundant decomposition of b, since $w=\bigvee R$ is minimal with respect to $b=w\vee v$. From $p_1, p_2\prec b=p_1\vee p_2$ we get the existence of an element $v_1\in V_1(p_1/0)$ with $p_2\vee v_1=b$. Again by lower continuity, there exists an element $w_1\leqq p_2$ which is minimal with respect to $v_1\vee w_1=b$. Let $w_1=\bigvee R_1$ be an irredundant decomposition of w_1. Then $b=v_1\vee\bigvee R_1$ is an irredundant decomposition of b. Since $b=v\vee\bigvee R=v_1\vee\bigvee R_1$ are two replaceable irredundant decompositions of b, we can replace v by some $r\in R_1\cup\{v_1\}$. Thus $b=v\vee\bigvee R=r\vee\bigvee R=r\vee w\leqq r\vee(p_1\wedge p_2)\leqq p_i\prec b$, with $i=1$ if $r=v_1\leqq p_1$ or $i=2$ if $r\in R_1\subseteq p_2/0$. This contradiction shows that there are no elements p_1, p_2 with $p_1, p_2\prec p_1\vee p_2=a\vee v$. Therefore, $b=a\vee v$ is completely join-irreducible in b/a.

The last result raises the question whether the dual of a lower continuous strongly dually atomic V_1-lattice is algebraic. By [1] (Theorem 2) we have only to ask if the dual of such a lattice is a Q-lattice. For instance, the dual of the lattice L in the example of [3], (p. 16) is a lower continuous V_1-lattice and L is not a Q-lattice. But L is not strongly atomic either. On the other hand, it is possible to construct a strongly dually atomic V_1-lattice the dual of which is not a Q-lattice.

Let L be an algebraic distributive atomistic lattice and let $B=\{p_1, p_2, \dots\}$ be a countable maximal independent set of atoms of L, i. e., 1 is the direct join of B, denoted by $1=\dot{\bigvee} B$. Further we define $a_i:=\dot{\bigvee}(p_j\colon p_j\in B,\ j=1, \dots, i)$. Now we add a new element $0^*\prec 0$ and complete the lattice $a_1/0^*$ by adding a countable chain $a_1\succ b_1\succ b_2\succ \dots > 0^*$. Having completed the lattice $a_i/0^*$, we complete the lattice $a_{i+1}/0^*$ in such a way that we add a countable chain of elements $a_{i+1}\succ c_{b,1}\succ \prec c_{b,2}\prec\dots>b$ to each $b\in a_i/0^*$ with $b\neq a_i$ and $b\notin a_{i-1}/0^*$. In this way we get a new lattice L^* which is a strongly dually atomic V_1-lattice by construction, since every element of $L^* \setminus L$ is completely join-irreducible, the elements $0, p_2, p_3, \dots$ are also completely join-irreducible in L^*, and every element of L^* is a join of such elements. But the dual of L^* is not a Q-lattice, since for each $b \in L^* \setminus L$ we have a chain of elements $c_{b,1} \succ c_{b,2} \succ \dots > b$ such that $b = \bigwedge(c_{b,i}\colon i=1,2,\dots)$ and $b < \bigwedge(c_{b,i}\colon i=1,\dots,n)$ for all n, i.e. each element of $L^* \setminus L$ is not inaccessible in the dual of L^*. Therefore, the set Q of the inaccessible elements of the dual of L^* is contained in L. Thus, if $x = \bigwedge R$ and $R \subseteqq Q$ then $x \in L$, and if $b \in L^* \setminus L$, then there is no subset R of Q with $b = \bigwedge R$. However, L^* is not lower continuous either.

As a corollary to Lemma 9, we get the following

Corollary 12. *Every element of a lower continuous V_1-lattice L has irredundant decompositions iff to each infinite chain C of L with $c \neq \bigvee C$ for all $c \in C$, there exist an element $x < \bigvee C$ and an element $v \in V_1$ with $x \vee v = \bigvee C$.*

Proof. If every element of L has irredundant decompositions, then by Lemma 9 there exist such elements x and v. Suppose now that to each infinite chain C with $c \neq \bigvee C$ for all $c \in C$ there exist $x < \bigvee C$ and $v \in V_1$ with $x \vee v = \bigvee C$. Let a be an arbitrary element of L. If $a/0$ contains a dual atom x, then there exists an element $v \in V_1$ with $x \vee v = a$, since L is a V_1-lattice. If $a/0$ does not contain dual atoms, then there exists a maximal infinite chain C of elements between 0 and a, and $c \in C \setminus \{a\}$ implies $c \neq \bigvee (C \setminus \{a\}) = a$. So our assumption yields the existence of an element $x < a$ and of an element $v \in V_1$ such that $x \vee v = a$. Consequently, by Lemma 9, every element of L has irredundant decompositions.

REFERENCES

[1] Birkhoff, G. and Frink Jr., O., Representations of lattices by sets, *Trans. Amer. Math. Soc.* **64** (1948), 299—316. *MR* **10**—279.

[2] Crawley, P., Decomposition theory of nonsemimodular lattices, *Trans. Amer. Math. Soc.* **99** (1961), 246—254. *MR* **22** # 10930.

[3] Crawley, P. and Dilworth, R. P., *Algebraic theory of lattices*, Prentice-Hall, Englewood Cliffs, N. J., 1973.

[4] Dilworth, R. P., The arithmetical theory of Birkhoff lattices, *Duke Math. J.* **8** (1941), 286—299. *MR* 3—100.

[5] Dilworth, R. P., Note on the Kurosh—Ore theorem, *Bull. Amer. Math. Soc.* **52** (1946), 659—663. *MR* **8**—62.

[6] Dilworth, R. P. and Crawley, P., Decomposition theory for lattices without chain conditions, *Trans. Amer. Math. Soc.* **96** (1960), 1—22. *MR* **22** # 9461.

[7] Kuroš, A., Durchschnittsdarstellungen mit irreduziblen Komponenten in Ringen und sogenannten Dualgruppen. *Rec. Math. Moscou=Mat. Sbornik* **42** (1935), 613—616. *Zbl* **13**, 195.

[8] Ore, O., On the foundation of abstract algebra II, *Ann. of Math.* (2) 37 (1936), 265—292. *Zbl* **14**, 197.

[9] Richter, G., On the structure of lattices in which every element is a join of join-irreducible elements, *Period. Math. Hungar.* **13** (1982), 47—69.

[10] Richter, G. and Stern, M., A unified theory of cyclically generated modular lattices and AC-lattices, *Studia Sci. Math. Hungar.* **11** (1976), 413—429. *MR* **83a**: 06009.

[11] Richter, G. and Stern, M., ZI-lattices, Proc. Colloq. on Universal Algebra, Model Theory and Lattice Theory (held in Esztergom, 1977), North-Holland Publ. Co., Amsterdam, 635—652.

[12] Stern, M., Baer lattices, *Wiss. Z. Martin-Luther-Univ. Halle—Wittenberg Math.-Natur. Reihe* **27** (1978), no. 4, 135—140. *MR* **80k**: 06007. Baer lattices II, *ibid.* **27** (1978),

no. 5, 65—69. *MR* **80e**: 06003a. Baer lattices III, *ibid.* **28** (1979), no. 1, 21—24. *MR* **80e**: 06003b.
[13] STERN, M., Baersche Verbände, Doctoral thesis, Halle, 1978.
[14] STERN, M., Generalized matroid lattices, Proc. Colloq. on Algebraic Methods in Graph Theory, Szeged (Hungary), 1978, North-Holland Publ. Co., Amsterdam, 1981, 727—748.
[15] SZÁSZ, G., *Einführung in die Verbandstheorie,* Akadémiai Kiadó, Budapest; Teubner, Leipzig, 1962. *MR* **25** # 2011.

(Received December 5, 1980)

SEKTION MATHEMATIK
MARTIN-LUTHER-UNIVERSITÄT HALLE—WITTENBERG
UNIVERSITÄTSPLATZ 6
DDR—4010 HALLE
GERMAN DEMOCRATIC REPUBLIC

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 251—282

О МАКСИМАЛЬНОМ ЦИКЛЕ ГРАФА

Ж. Г. НИКОГОСЯН

Summary

Let V denote the order, δ the minimum degree and k the point connectivity of a graph G.

THEOREM 1. *Every 2-connected graph G satisfying $\delta \geqq \frac{v+k}{3}$ is hamiltonian.*

THEOREM 2. *Every 3-connected graph G satisfying $\delta < \frac{v+k}{3}$ has a simple circuit of length $\geqq 3\delta - k$.*

Рассматриваются конечные неориентированные графы без петель и кратных ребер. Все понятия и обозначения, не определяемые здесь, можно найти в книге [1]. Для любого подграфа L графа G через $V(L)$ и $X(L)$ будем обозначать соответственно множество вершин и множество ребер подграфа L. Пусть $\delta(G)$ обозначает минимальную степень, $k(G)$—вершинную связность, $h(G)$—максимальную длину простых циклов и $\alpha(G)$—число вершинной независимости графа G.

В 1952 г. Дирак [2] доказал, что если в графе G имеет место $\delta(G) \geqq v(G)/2$, где $v(G) = |V(G)|$, то G—гамильтонов, т.е. $h(G) = v(G)$. В работе [2] доказано также, что если $\delta(G) < v(G)/2$ и $k(G) \geqq 2$, то $h(G) \geqq 2\delta(G)$.

Нэш-Вильямс [3] в 1971 г. доказал, что если 2-связный граф G удовлетворяет условиям $\delta(G) \geqq (v(G)+2)/3$ и $\delta(G) \geqq \alpha(G)$, то $h(G) = v(G)$.

П. Эрдёш и В. Хватал [4] в 1972 г. доказали гамильтоновость графа G при условии $\alpha(G) \leqq k(G)$.

В настоящей работе доказываются следующие две теоремы:

Теорема 1. *Если 2-связный граф G удовлетворяет условию $\delta(G) \geqq$ $\geqq (v(G)+k(G))/3$, то $h(G) = v(G)$.*

Теорема 2. *Если 3-связный граф G удовлетворяет условию $\delta(G) <$ $< (v(G)+k(G))/3$, то $h(G) \geqq 3\delta(G) - k(G)$.*

Первый и последний элемент любой конечной последовательности l будем обозначать через $F(l)$ и $L(l)$ соответственно. Запись $i = \overline{n_1, n_2}$ будет означать, что индекс i пробегает все значения множества $\overline{n_1, n_2} = \{n_1, n_1+1, \ldots, n_2\}$. Пусть

$$N(v) \equiv \{u \in V(G)/uv \in X(G)\}, \quad N(R) \equiv \bigcup_{v \in R} N(v) \setminus R,$$

где $v \in V(G)$, $R \subseteq V(G)$.

1980 *Mathematics Subject Classification*. Primary 05C38; Secondary 05C45.
Key words and phrases. S-admissible pairs of chains, deviation of the chain from the subgraph.

Рассмотрим произвольный 2-связный граф G, для которого $S=\{v_1, v_2, \ldots, v_k\}$ является некоторым минимальным разделяющим множеством вершин. Пусть $G_1, G_2, \ldots, G_t$—компоненты связности графа $G-S$ и пусть

$$H_1 \equiv \bigcup_{i=1}^{f} G_i, \quad H_2 \equiv \bigcup_{i=f+1}^{t} G_i,$$

где $f \in \overline{1, t-1}$.

Для произвольной простой цепи $P=P_1P_2\ldots P_n$ $(n \geqq 3)$ графа G и для любых i, j таких, что $ij \in \overline{1, n}$, $i<j$, введём обозначения

$$[P_i, P_j]\uparrow P = P_i P_{i+1} \ldots P_j, \quad [P_i, P_j)\uparrow P = [P_i, P_j]\uparrow P - P_j,$$

$$(P_i, P_j]\uparrow P = [P_i, P_j]\uparrow P - P_i, \quad (P_i, P_j)\uparrow P = [P_i, P_j)\uparrow P - P_i.$$

Определение 1. Пару цепей R_1, R_2 графа G назовём S-допустимой, если

$$\{F(R_1),\ L(R_2)\} \subseteqq V(H_r), \quad \{L(R_1),\ F(R_2)\} \subseteqq S, \quad V(R_1)\cap V(R_2) = \emptyset,$$

$$V^* \subseteqq V(H_r)\cup S, \quad N(F(R_1))\cup N(L(R_2)) \subseteqq V^*\cup S,$$

где $V^* = V(R_1)\cup V(R_2)$, $r \in \overline{1, 2}$.

Определение 2. Отключением цепи $P=P_1P_2\ldots P_n$ $(n \geqq 3)$ из подграфа L графа G назовём последовательное удаление из L всех вершин $P_2, P_3, \ldots, P_{n-1}$. Граф, полученный путём отключения цепи P из графа L, будем обозначать через $L-P$.

Пусть $Q \equiv \alpha_1\alpha_2\ldots\alpha_c$—любая простая цепь графа G, для которой $V(Q) \subseteqq V(H_r)\cup S$, $r \in \{1, 2\}$.

Будем предполагать, что для некоторой цепи $Q_0 \equiv \alpha_a\alpha_{a+1}\ldots\alpha_{a+b}$, где $a \geqq 1$, $b \geqq 2$, $a+b \leqq c$, множество $V(Q_0)$ имеет непустое подмножество V_0, для которого $V_0 \setminus \{\alpha_a, \alpha_{a+b}\} \neq \emptyset$.

Определение 3. Простую цепь T графа G назовём эстафетным путём для цепи Q и будем обозначать через (T), если

$$F(T) \in V(Q), \quad L(T) \in V(Q), \quad V(T) \setminus \{F(T), L(T)\} \subseteqq V(H_r) \setminus V(Q).$$

Определение 4. Пусть (T_1) и (T_2)—вершинно непересекающиеся эстафетные пути для цепи Q в графе G. Тогда (T_1) и (T_2) назовём смежными, если либо

$$F(T_2) \in (F(T_1), L(T_1))\uparrow Q, \quad L(T_2) \notin [F(T_1), L(T_1)]\uparrow Q,$$

либо

$$L(T_2) \in (F(T_1), L(T_1))\uparrow Q, \quad F(T_2) \notin [F(T_1), L(T_1)]\uparrow Q.$$

Определение 5. Пусть эстафетные пути (T_1) и (T_2) определены для цепи Q. Скажем, что (T_1) накрывает (T_2), если

$$\{F(T_2), L(T_2)\} \subseteqq (F(T_1), L(T_1))\uparrow Q, \quad V(T_1)\cap V(T_2) = \emptyset.$$

Определение 6. Пусть $(T_1), (T_2), \ldots, (T_d)$—вершинно непересекающиеся эстафетные пути для Q, ни один из которых не накрывает другой. Последовательность $(T_1, T_2, \ldots, T_d)$ назовём эстафетным путём для Q, если эстафетный путь (T_i) при любом $i=\overline{2, d-1}$ смежен только с (T_{i-1}) и (T_{i+1}), при $i=1$ смежен только с (T_2) и при $i=d$ смежен только с (T_{d-1}).

Определение 7. Пусть

$$i_0 = \max\left\{i/i \geqq a+b+1,\ L(T) = \alpha_i,\ F(T) \in \{\alpha_a, \alpha_{a+1}, \ldots, \alpha_{a+b}\} \setminus V_0\right\},$$

$$i_1 = \min\left\{i/i \leqq a-1,\ L(T) = \alpha_i,\ F(T) \in \{\alpha_a, \alpha_{a+1}, \ldots, \alpha_{a+b}\} \setminus V_0\right\},$$

где max и min берётся по всем цепям T, являющимися эстафетными путями для Q. Тогда эстафетный путь (T_0), определённый для Q, назовём эстафетным звеном для тройки (Q_0, V_0, Q), если для некоторого $i \in \{i_0, i_1\}$ имеет место

$$L(T_0) = \alpha_i, \quad F(T_0) \in \{\alpha_a, \alpha_{a+1}, \ldots, \alpha_{a+b}\} \setminus V_0.$$

Определение 8. Эстафетный путь $(T_1, T_2, \ldots, T_d)$ назовём эстафетной цепью для тройки (Q_0, V_0, Q), если либо выполняются следующие три условия

1. $L(T_i) \in \{\alpha_{a+b+1}, \alpha_{a+b+2}, \ldots, \alpha_c\}$, $i=\overline{1, d}$;
2. Цепь T_1 является эстафетным звеном для (Q_0, V_0, Q);
3. При любом $i=\overline{2, d}$ цепь T_i является эстафетным звеном для тройки $([\alpha_a, L(T_{i-1})]\uparrow Q, V_0, Q)$, либо выполняются следующие три условия:
1. $L(T_i) \in \{\alpha_1, \alpha_2, \ldots, \alpha_{a-1}\}$, $i=\overline{1, d}$;
2. Цепь T_1 является эстафетным звеном для (Q_0, V_0, Q);
3. При любом $i=\overline{2, d}$ цепь T_i является эстафетным звеном для тройки $([\alpha_{a+b}, L(T_{i-1})]\uparrow Q, V_0, Q)$.

Число d будем называть длиной эстафетной цепи $(T_1, T_2, \ldots, T_d)$.

Пусть $(T_1, T_2, \ldots, T_d)$—эстафетный путь для Q и пусть

$$F(T_i) \equiv \alpha_{k_i}, \quad L(T_i) \equiv \alpha_{q_i}, \quad \alpha_{k_i} < \alpha_{q_i}, \quad i = \overline{1, d},$$

$$\alpha_{k_1} \leqq \alpha_{k_2} \leqq \ldots \leqq \alpha_{k_d}.$$

Определение 9. Цикл, полученный из $\left(\bigcup\limits_{i=1}^{d} T_i\right) \cup [\alpha_{k_1}, \alpha_{q_d}]\uparrow Q$ путём удаления всех подцепей $[\alpha_{k_i}, \alpha_{q_{i-1}}]\uparrow Q$, $i=\overline{2, d}$, будем называть циклом, порождённым эстафетным путём $(T_1, T_2, \ldots, T_d)$.

Определение 10. Длиной эстафетного пути $(T_1, T_2, \ldots, T_d)$ назовём длину цепи $[\alpha_{k_1}, \alpha_{q_d}]\uparrow Q$.

Определение 11. Эстафетный путь $(T_1, T_2, \ldots, T_d)$, определённый как для Q, так и для $Q' \subseteq Q$, назовём максимальным на Q', если он имеет максимальную длину среди всех путей такого рода.

В дальнейшем мы часто будем пользоваться теоремой Менгера [1], которую мы здесь сформулируем с некоторыми её следствиями:

Теорема Менгера. *Пусть V_1 и V_2—любые непересекающиеся подмножества вершин k-связного графа G. Тогда справедливы следующие утверждения.*

1. *Подграфы $\langle V_1 \rangle$ и $\langle V_2 \rangle$ соединены по крайней мере k простыми цепями, которые попарно могут пересекаться только в концевых вершинах.*

2. *Подграфы* $\langle V_1\rangle$ *и* $\langle V_2\rangle$ *соединены по крайней мере* $\min(|V_1|, |V_2|, k)$ *простыми цепями попарно без общих вершин.*

3. *Если* $|V_1| \geqq k$, *то подграфы* $\langle V_1\rangle$ *и* $\langle V_2\rangle$ *соединены по крайней мере k простыми цепями, которые в подграфе* $\langle V_1\rangle$ *попарно не имеют общих вершин.*

Для S-допустимой пары цепей R_1, R_2 введём обозначения

$$d_1 = |N(F(R_1)) \cap V^*|, \quad d_2 = |N(L(R_2)) \cap V^*|.$$

Лемма 1. *Если для S-допустимой пары цепей* R_1, R_2 *графа G имеет место* $d_1 + d_2 \geqq |V^*|+1$, *то существует цепь* $P = l$, *удовлетворяющая условию*

$$F(P) = L(R_1), \quad L(P) = F(R_2), \quad V(P) = V^*. \tag{1}$$

Доказательство. Введём обозначения

$$R_1 = \xi_1\xi_2 \ldots \xi_m, \quad R_2 = \xi_{m+1}\xi_{m+2} \ldots \xi_n, \quad \xi_1\xi_2 \ldots \xi_n = P_0,$$

$$S_1 = \{i/\xi_1\xi_i \in X(G)\}, \quad S_2 = \{i/\xi_n\xi_{i-1} \in X(G)\},$$

где $\xi_m\xi_{m+1}$—фиктивное ребро. Очевидно, что $S_1 \cup S_2 \subseteqq \{\xi_2, \xi_3, \ldots, \xi_n\}$. Если $|S_1 \cap S_2| \leqq 1$, то

$$|V^*| \geqq |S_1| + |S_2| = d_1 + d_2 \geqq |V^*|+1,$$

что невозможно. Пусть $|S_1 \cap S_2| \geqq 2$, т.е. существуют числа t_1, t_2, удовлетворяющие условию

$$t_1 \neq t_2, \quad \xi_1\xi_{t_i+1} \in X(G), \quad \xi_n\xi_{t_i} \in X(G), \quad i = \overline{1, 2}.$$

Далее, пусть $t_0 \in \{t_1, t_2\}$ и $t_0 \neq m$. Тогда цепь, полученная из подграфа $P_0 \cup \xi_1\xi_{t_0+1} \cup \xi_n\xi_{t_0}$ путём удаления ребер $\xi_{t_0}\xi_{t_0+1}$, $\xi_m\xi_{m+1}$, удовлетворяет условию (1). Лемма доказана.

Лемма 2. *Если для S-допустимой пары цепей* R_1, R_2 *графа G имеет место* $|V^* \cap S| \geqq 3$, *то либо существует цепь* $P = l^1$, *удовлетворяющая условию* (1), *либо существует цепь* $P = l^2$, *удовлетворяющая условию*

$$\begin{gathered} F(P) = L(R_1), \quad L(P) = F(R_2), \quad V(P) \subseteqq V(H_r) \cup S, \\ V(P) \cap S \subseteqq V^* \cap S, \quad |V(P)| \geqq d_1 + d_2 - 1, \end{gathered} \tag{2}$$

либо существует цепь $P = l^3$, *удовлетворяющая условию*

$$\begin{gathered} F(P) \in S, \quad L(P) \in S, \quad V(P) \subseteqq V(H_r) \cup S, \\ V(P) \cap S \subseteqq V^* \cap S, \quad |V(P)| \geqq d_1 + d_2. \end{gathered} \tag{3}$$

Если же для R_1, R_2 *имеет место* $|V^* \cap S| = 2$, *то либо существует цепь* $P = l^4$, *удовлетворяющая хотя бы одному из условий* (1), (2), *либо для любой вершины* $z \in S$, *отличной от* $L(R_1)$ *и* $F(R_2)$, *существует цепь* $P = l^5$, *удовлетворяющая условию*

$$\begin{gathered} F(P) \in S, \quad L(P) \in S, \quad V(P) \subseteqq V(H_r) \cap S, \\ V(P) \cap S \subseteqq \{L(R_1), F(R_2), z\}, \quad |V(P)| \geqq d_1 + d_2 + 1. \end{gathered} \tag{4}$$

Доказательство. Введём обозначения

$$R_1 = \xi_1\xi_2 \dots \xi_m, \quad R_2 = \xi_{m+1}\xi_{m+2} \dots \xi_n, \quad \xi_1\xi_2 \dots \xi_n = P_0,$$

$$\Delta = |V^* \cap S|, \quad a = \max \{i/\xi_1\xi_i \in X(G)\}, \quad b = \min \{i/\xi_n\xi_i \in X(G)\},$$

где $\xi_m\xi_{m+1}$—фиктивное ребро. Рассмотрим следующие случаи:

Случай а1. Для некоторых чисел s, t имеет место

$$\xi_n\xi_s \in X(G), \quad \xi_1\xi_t \in X(G), \quad s < t, \quad m \notin \overline{s, t-1},$$

$$\xi_1\xi_i \notin X(G), \quad \xi_n\xi_i \notin X(G), \quad i = \overline{s+1, t-1}.$$

Цикл, определённый эстафетным путём $(\xi_1\xi_t, \xi_n\xi_s)$ на цепи P_0, обозначим через C_0. Пусть

$$C_0 - \xi_m\xi_{m+1} \equiv P_1, \quad \{i/\xi_1\xi_i \in X(G)\} \equiv S_1, \quad \{i/\xi_n\xi_{i-1} \in X(G)\} \equiv S_2.$$

Очевидно, что

$$S_1 \subseteq \{\xi_2, \xi_3, \dots, \xi_s\} \cup \{\xi_t, \xi_{t+1}, \dots, \xi_n\},$$

$$S_2 \subseteq \{\xi_2, \xi_3, \dots, \xi_{s+1}\} \cup \{\xi_{t+1}, \xi_{t+2}, \dots, \xi_n\}.$$

Если $|S_1 \cap S_2| \leqq 1$, то P_1 удовлетворяет условию (2), так как

$$|V(P_1)| \geqq (|S_1| + |S_2| - 1) + |\{\xi_1\}| + t - s - 2 \geqq d_1 + d_2 - 1.$$

Если же $|S_1 \cap S_2| \geqq 2$, то существуют различные числа t_1, t_2, удовлетворяющие условию

$$\xi_1\xi_{t_i+1} \in X(G), \quad \xi_n\xi_{t_i} \in X(G), \quad i = \overline{1, 2}.$$

Пусть $t_0 \in \{t_1, t_2\}$ и $t_0 \neq m$. Тогда цепь, полученная из подграфа $P_0 \cup \cup \xi_1\xi_{t_0+1} \cup \xi_n\xi_{t_0}$ путём удаления ребер $\xi_{t_0}\xi_{t_0+1}, \xi_m\xi_{m+1}$, удовлетворяет условию (1).

Случай а2. $a \leqq m \leqq b-1$.
Рассмотрим $(\Delta-1)$-связный граф

$$G^{(1)} \equiv G - ((S \setminus V^*) \cup \{\xi_a\}).$$

Так как

$$\min(a, b) \geqq \min(d_1+1, d_2+1) \geqq \delta(G) - (k - \Delta) \geqq \Delta,$$

то по теореме Менгера (пункт 2) в графе $G^{(1)}$ цепи $\xi_1\xi_2 \dots \xi_{a-1}$, $\xi_{a+1}\xi_{a+2} \dots \xi_n$ соединены по крайней мере $\Delta-1$ простыми цепями $l_1, l_2, \dots, l_{\Delta-1}$ попарно без общих вершин. Пусть

$$A \equiv \{i/V(l_i) \cap (V(H_2) \cup (S \setminus V^*)) = \emptyset\}.$$

Поскольку цепи l_i попарно не пересекаются, то

$$|\{i/V(l_i) \cap V(H_2) \neq \emptyset\}| \leqq \left[\frac{\Delta}{2}\right],$$

и следовательно, $|A| \geqq (\Delta-1) - [\Delta/2] \geqq]\Delta/2[$.

Допустим, что $\Delta=2$. Тогда имеет место

$$S\cap\{\xi_1, \xi_2, \ldots, \xi_{a-1}\}=\emptyset \quad \text{или} \quad S\cap\{\xi_{a+1}, \xi_{a+2}, \ldots, \xi_n\}=\emptyset,$$

откуда следует $V(l_1)\cap V(H_2)=\emptyset$, т.е. $|A|\geqq 1$. Если же $\Delta\geqq 3$, то неравенство $|A|\geqq 1$ очевидно следует из $|A|\geqq]\Delta/2[$.

Таким образом $|A|\geqq 1$ при $\Delta\geqq 2$, т.е. тройка

$$A_0\equiv((\xi_1\xi_2\ldots\xi_a),\ \{\xi_a\},\ P_0)$$

обладает некоторым эстафетным звеном Q_1. Пусть $(Q_1, Q_2, \ldots, Q_g)$ является эстафетной цепью максимальной длины для тройки A_0, где $g\geqq 1$. Покажем, что $L(Q_g)=\xi_n$. Допустим обратное, т.е. $L(Q_g)\neq\xi_n$. Если $L(Q_g)\in[\xi_{b+1}, \xi_n]\uparrow P_0$, то $(Q_1, Q_2, \ldots, Q_g, \xi_b\xi_n)$ является эстафетной цепью для A_0 длиной $g+1$, что противоречит максимальности числа g. Пусть $L(Q_g)\in[\xi_a, \xi_b]\uparrow P_0$. Рассмотрим $(\Delta-1)$-связный граф

$$G^{(2)}\equiv G-((S\setminus V^*))\cup\{L(Q_g)\}.$$

По теореме Менгера в графе $G^{(2)}$ цепи $[\xi_1, L(Q_g))\uparrow P_0$ и $(L(Q_g), \xi_n]\uparrow P_0$ соединены по меньшей мере $\Delta-1$ попарно непересекающимися простыми цепями, среди которых существует цепь, не имеющая общих вершин с множеством $V(H_2)$. Тогда для тройки

$$([\xi_1, L(Q_g)]\uparrow P_0,\ \{L(Q_g)\},\ P_0)$$

существует некоторое эстафетное звено Q_{g+1}, что противоречит максимальности числа g. Таким образом $L(Q_g)=\xi_n$.

Введём обозначения

$$F(Q_i)=\xi_{a_i}, \quad L(Q_i)=\xi_{b_i}, \quad i=\overline{0, g},$$

где $Q_0\equiv\xi_{a_0}\xi_{b_0}$, $a_0=1$ и

$$b_0=\min\{i/i>a_1,\ \xi_1\xi_i\in X(G)\}.$$

Без потери общности можем предполагать, что

$$a_g=\max\{i/i<b_{g-1},\ \xi_n\xi_i\in X(G)\}.$$

Цикл, порождённый эстафетным путём $(Q_0, Q_1, \ldots, Q_g)$, обозначим через C_0.

Случай а2.1. $b_i\leqq m\leqq a_{i+2}-1$, $i\in\overline{0, g-2}$.

Цепь $C_0-\xi_m\xi_{m+1}$ удовлетворяет условию (2).

Случай а2.2. Случай а2.1 не имеет место, т.е.

$$a_{h+1}\leqq m\leqq b_h-1, \quad h\in\overline{1, g-2}.$$

Циклы, порождённые эстафетными путями $(Q_0, Q_1, \ldots, Q_h)$, $(Q_{h+1}, Q_{h+2}, \ldots, Q_g)$, обозначим через C_1 и C_2 соответственно. Если $g=2$, то приходим к случаю а2.1. Пусть $g\geqq 3$.

Заменим Q_h через такую цепь Q, для которой последовательность $(Q_1, Q_2, \ldots, Q_{h-1}, Q, Q_{h+1}, \ldots, Q_g)$, как и прежде, определяет эстафетный путь для P_0, а вершина $L(Q)$ на цепи P_0 имеет наименьший возможный номер. Без введения новых обозначений можем предполагать, что $Q=Q_h$. Аналогично, заменим цепь Q_{h+1} через такую цепь (ее также будем обозначать через Q_{h+1}), для которой последовательность $(Q_1, Q_2, \ldots, Q_g)$, как и прежде, определяет эстафетный путь для P_0, а вершина $F(Q_{h+1})$ на цепи P_0 имеет наибольший возможный номер.

Случай а2.2.1. $\Delta \geqq 3$.

Рассмотрим некоторую вершину $\xi_\lambda \in V^* \cap S$, отличную от ξ_m и ξ_{m+1}. В силу симметрии достаточно рассматривать случай $\lambda \geqq a_{h+1}$.

Случай а2.2.1.1. $\lambda > b$.

Цепь

$$(C_1-[\xi_m, \xi_{b_h}]\uparrow P_0)\cup \xi_\lambda \xi_{\lambda+1} \ldots \xi_n \xi_{\lambda_0} \xi_{\lambda_0-1} \ldots \xi_{b_h},$$

где

$$\lambda_0 = \max \{i/i < \lambda,\ \xi_n \xi_i \in X(G)\},$$

удовлетворяет условию (3).

Случай а2.2.1.2. $\lambda \leqq b$, $b_c \leqq \lambda \leqq a_{c+2}$, $c \in \overline{h, g-2}$.

Если $(c-h)$—чётное число, то цепь

$$(C_0-[\xi_\lambda, \xi_{b_h}]\uparrow (C_0-\xi_n))\cup [\xi_{m+1}, \xi_{b_h}]\uparrow P_0$$

удовлетворяет условию (3). Если же $(c-h)$—нечётное число, то условию (3) удовлетворяет цепь

$$(C_0-[\xi_\lambda, \xi_{a_{h+1}}]\uparrow (C_0-\xi_n))\cup [\xi_{a_{h+1}}, \xi_m]\uparrow P_0.$$

Случай а2.2.1.3. $\lambda \leqq b$, $a_{c+1} \leqq \lambda \leqq b_c$, $c \in \overline{h+1, g-2}$.

Если $(c-h+1)$—чётное число, то цепь

$$(C_0-[\xi_{a_{c+1}}, \xi_{b_h}]\uparrow (C_0-\xi_n))\cup ([\xi_{a_{c+1}}, \xi_\lambda]\uparrow P_0)\cup ([\xi_{m+1}, \xi_{b_h}]\uparrow P_0)$$

удовлетворяет условию (3). Если же $(c-h+1)$—нечётное число, то условию (3) удовлетворяет цепь

$$(C_0-[\xi_{a_{c+1}}, \xi_{a_{h+1}}]\uparrow (C_0-\xi_n))\cup ((\xi_{a_{c+1}}, \xi_\lambda)\uparrow P_0)\cup ([\xi_{a_{h+1}}, \xi_m]\uparrow P_0).$$

Случай а2.2.1.4. $a_{h+1} \leqq \lambda \leqq b_h$.

Пусть $V^* \cap S \subseteqq V([\xi_{a_{h+1}}, \xi_{b_h}]\uparrow P_0)$. Допустим, что для цепи P_0 существует эстафетный путь (M), удовлетворяющий условию

$$V^* \cap S \subseteqq V([F(M), L(M)]\uparrow P_0) \subset V([\xi_{a_{h+1}}, \xi_{b_h}]\uparrow P_0). \tag{5}$$

По определению цепей Q_h и Q_{h+1} имеем

$$V(M)\cap (V(Q_i)\setminus \{F(Q_i), L(Q_i)\}) = \emptyset, \quad i = \overline{0, g}.$$

Если существует цепь, удовлетворяющая (5), то очевидно существуют цепи M_1, M_2, удовлетворяющие условию (5), и числа q_1, q_2, для которых имеет место

$$F(M_1) = \xi_{q_1}, \quad q_1 = \min\{i/F(M) = \xi_i\},$$

$$L(M_2) = \xi_{q_2}, \quad q_2 = \max\{i/L(M) = \xi_i\}.$$

Если же не существует цепи, удовлетворяющей условию (5), то будем предполагать, что

$$q_1 = \min\,\{i/i \geqq a_{h+1},\ \xi_i \in V^* \cap S\},$$

$$q_2 = \max\,\{i/i \leqq b_h,\ \xi_i \in V^* \cap S\}.$$

Рассмотрим $(\Delta-2)$-связный граф

$$G^{(3)} \equiv G - ((S \setminus V^*) \cup \{\xi_{q_1}, \xi_{q_2}\}).$$

Так как $q_2 - q_1 - 1 \geqq \Delta - 2 \geqq 1$, то по теореме Менгера в графе $G^{(3)}$ подграфы $([\xi_1, \xi_{q_1}) \uparrow P_0) \cup ((\xi_{q_2}, \xi_n] \uparrow P_0)$ и $(\xi_{q_1}, \xi_{q_2}) \uparrow P_0$ соединены по крайней мере одной цепью, которая ввиду $V^* \cap S \subseteq V((\xi_{q_1}, \xi_{q_2}) \uparrow P_0)$, не имеет общих вершин с множеством $V(H_2)$. Следовательно, тройка

$$A_1 \equiv ([\xi_{q_1}, \xi_{q_2}] \uparrow P_0,\ \{\xi_{q_1}, \xi_{q_2}\},\ P_0)$$

обладает некоторым эстафетным звеном N_1. Пусть $(N_1, N_2, \ldots, N_q)$ является эстафетной цепью минимальной длины среди всех эстафетных цепей $(N_1, N_2, \ldots$ $\ldots, N_i)$, $i \geqq 1$, которые определены для тройки A_1 и удовлетворяют условию $L(N_i) \notin V([\xi_{a_{h+1}}, \xi_{b_h}] \uparrow P_0)$. В силу симметрии можем предполагать, что $L(N_q) \in$ $\in V(\xi_{b_h}, \xi_n] \uparrow P_0$. По определению цепей Q_i имеем

$$V(Q_i) \cap V(N_j) \subseteq V^*, \quad i = \overline{0, g}, \quad j = \overline{0, q}, \quad L(N_q) \in V((\xi_{b_h}, \xi_{a_{h+2}}] \uparrow P_0).$$

Цикл, порождённый эстафетным путём $(N_1, N_2, \ldots, N_q)$, обозначим через C_3. Если $\xi_m \xi_{m+1} \in X(C_3)$, то цепь

$$((C_0 \cup C_3) - [\xi_{b_h}, L(N_q)] \uparrow P_0) - \xi_m \xi_{m+1}$$

удовлетворяет условию (3). Пусть $\xi_m \xi_{m+1} \notin X(C_3)$. Если множество $(V(C_3) \cap S) \setminus$ $\setminus \{F(N_1)\}$ содержит некоторую вершину u, то цепь

$$((C_0 \cup C_3 \cup [\xi_{m+1}, F(N_1)] \uparrow P_0) - ([\xi_{b_h}, L(N_q)] \uparrow P_0 \cup [F(N_1), u] \uparrow P_0)$$

также удовлетворяет условию (3). Если же $(V(C_3) \cap S) \setminus \{F(N_1)\} = \emptyset$, то из определения цепи N_1 следует существование для P_0 эстафетного пути (M), удовлетворяющего условию (5). Тогда условию (3) удовлетворяет цепь

$$((C_0 \cup C_4) - [\xi_{b_h}, L(N_q)] \uparrow P_0) - \xi_m \xi_{m+1},$$

где C_4—есть цикл, порождённый эстафетным путём $(M_2, N_1, N_2, \ldots, N_q)$.

Случай а2.2.2. $\Delta = 2$.

Рассмотрим граф $G^{(4)} \equiv G - (S \setminus \{z\})$, где $z \in S$—произвольная вершина, отличная от $L(R_1)$ и $F(R_2)$. Так как $G^{(4)}$ связен, то существует цепь N, удовлетворяющая условию

(6)
$$F(N) = z, \quad L(N) \in \left(V^* \cup \left(\bigcup_{i=0}^{g} V(Q_i)\right)\right) \setminus \{\xi_m, \xi_{m+1}\},$$
$$V(N) \setminus \{z, L(N)\} \subseteq V(H_1) \setminus \left\{V^* \cup \left(\bigcup_{i=0}^{g} V(Q_i)\right)\right\}.$$

Случай а2.2.2.1. $L(N) \in V(Q_v) \setminus \{F(Q_v), L(Q_v)\}$, $v \in \overline{0, g}$.

В силу симметрии достаточно рассматривать случай $v \geqq h+1$. Если $(v-h)$— нечётное число, то цепь

$$(C_0 \cup N \cup [\xi_{a_{h+1}}, \xi_m] \uparrow P_0) - [L(N), \xi_{a_{h+1}}] \uparrow (C_0 - \xi_1)$$

удовлетворяет условию (4). Если же $(v-h)$—чётное число, то условию (4) удовлетворяет цепь

$$(C_0 \cup N \cup [\xi_{m+1}, \xi_{b_h}] \uparrow P_0) - [L(N), \xi_{b_h}] \uparrow (C_0 - \xi_1).$$

Случай а2.2.2.2. $L(N) \equiv \xi_\tau \in V(P_0)$.

В силу симметрии достаточно рассматривать случай $\tau > m+1$. Тогда для некоторой цепи N^1, удовлетворяющей условию (6), и для некоторого числа τ_1 имеет место

$$L(N^1) = \xi_{\tau_1}, \quad \tau_1 = \max\{i / L(P) = \xi_i\},$$

где max берётся по всем цепьям P, удовлетворяющим условию (6).

Если для некоторой цепи P, удовлетворяющей условию (6), имеет место $L(P) \in [\xi_1, \xi_m) \uparrow P_0$, то существует цепь N^2, удовлетворяющая условию (6), и число τ_2, для которых

$$L(N^2) = \xi_{\tau_2}, \quad \tau_2 = \min\{i / i \leqq m,\ L(P) = \xi_i\},$$

где min берётся по всем цепьям P, удовлетворяющим условию (6).

Если же для любой цепи P, удовлетворяющей условию (6), имеет место $L(P) \notin [\xi_1, \xi_m) \uparrow P_0$, то будем предполагать, что $\tau_2 = m$. Поскольку при необходимости можно к вершине ξ_τ присоединить цепь N_1 с концевой вершиной $z \in S$, то этот случай подобен случаю $\Delta \geqq 3$. Тогда, если $\tau_1 > b_h$ или $\tau_2 > a_{h+1}$, то мы приходим к одному из случаев а2.2.1.1—а2.2.1.3. Пусть $a_{h+1} \leqq \tau_2 < \tau_1 \leqq b_h$. Рассмотрим граф

$$G^{(5)} \equiv G - ((S - \{\xi_m, \xi_{m+1}, z\}) \cup \{\tau_1, \tau_2\}).$$

Так как $k(G^{(5)}) \geqq 1$, то в графе $G^{(5)}$ подграфы

$$([\xi_1, \xi_{\tau_1}) \uparrow P_0) \cup ((\xi_{\tau_2}, \xi_n) \uparrow P_0), \quad (\xi_{\tau_1}, \xi_{\tau_2}) \uparrow P_0$$

соединены некоторой цепью Q. Условие $z \in Q$ противоречит либо максимальности числа τ_1, либо минимальности числа τ_2. Пусть $z \notin Q$. Тогда $V(Q) \cap \cap V(H_2) = \emptyset$ и, следовательно, тройка

$$([\xi_{\tau_1}, \xi_{\tau_2}] \uparrow P_0, \{\tau_1, \tau_2\}, P_0)$$

обладает некоторой эстафетной цепью, что приводит нас к случаю а2.2.1.4.

Случай а3. $b \leqq m \leqq a-1$.

Введём следующие обозначения

$$h_1 = \max \{i/i < a,\ \xi_n \xi_i \in X(G)\}, \quad h_2 = \min \{i/i > b,\ \xi_1 \xi_i \in X(G)\},$$

$$h_3 = \max \{i/i \leqq b,\ \xi_1 \xi_i \in X(G)\}, \quad h_4 = \min \{i/i \geqq a,\ \xi_n \xi_i \in X(G)\}.$$

Если $h_2 \leqq m$ или $h_1 \geqq m+1$, то приходим к случаю а1. Пусть $h_2 \geqq m+1$, $h_1 \leqq m$.

Через Q_1 и Q_2 обозначим цепи, для которых последовательность $(\xi_1 \xi_{h_3},\ Q_1,\ Q_2,\ \xi_{h_4} \xi_n)$ определяет эстафетный путь для P_0 и существуют числа a_1, a_2, b_1, b_2 такие, что

$$F(Q_i) = \xi_{a_i}, \quad L(Q_i) = \xi_{b_i}, \quad i = \overline{1, 2},$$

$$b_1 = \min \{i/L(P) = \xi_i\}, \quad a_2 = \max \{i/F(Q) = \xi_i\},$$

где min или max берётся по всем цепьям P, Q, для которых последовательность $(\xi_1 \xi_{h_3},\ P,\ Q,\ \xi_{h_4} \xi_n)$ определяет эстафетный путь для P_0. Поскольку последовательность $(\xi_1 \xi_{h_3},\ \xi_1 \xi_{h_2},\ \xi_{h_1} \xi_n,\ \xi_{h_4} \xi_n)$ также определяет эстафетный путь для P_0, то из определения чисел b_1, a_2 следует $h_1 \leqq a_2 < b_1 \leqq h_2$.

Введём обозначения

$$a_0 = \min \{i/i > a_1,\ \xi_1 \xi_i \in X(G)\}, \quad b_0 = \max \{i/i < b_2,\ \xi_n \xi_i \in X(G)\}.$$

Цикл, порождённый эстафетным путём $(\xi_1 \xi_{a_0},\ Q_1,\ Q_2,\ \xi_{b_0} \xi_n)$, обозначим через C_0. Если $m \notin \overline{a_2, b_1-1}$, то цепь $C_0 - \xi_m \xi_{m+1}$ очевидно удовлетворяет условию (2). Пусть $a_2 \leqq m \leqq b_1 - 1$.

Случай а3.1. $\Delta \geqq 3$.

Рассмотрим некоторую вершину $\xi_\pi \in V^* \cap S$, отличную от ξ_m и ξ_{m+1}. В силу симметрии можем предполагать, что $\pi \geqq a_2$.

Случай а3.1.1. $\pi > h_4$.

Цепь

$$\xi_\pi \xi_{\pi+1} \cdots \xi_n \xi_{\pi_1} \xi_{\pi_1 - 1} \cdots \xi_{h_2} \xi_1 \xi_2 \cdots \xi_m,$$

где $\pi_1 = \max \{i/i < \pi,\ \xi_n \xi_i \in X(G)\}$, удовлетворяет условию (3).

Случай а3.1.2. $h_2 < \pi \leqq h_4$.

Цепь

$$\xi_\pi \xi_{\pi+1} \cdots \xi_n \xi_{h_1} \xi_{h_1 - 1} \cdots \xi_1 \xi_{\pi_2} \xi_{\pi_2 - 1} \cdots \xi_{m+1},$$

где $\pi_2 = \max \{i/i < \pi,\ \xi_1 \xi_i \in X(G)\}$, удовлетворяет условию (3).

Случай а3.1.3. $b_1 < \pi \leqq h_2$.

Цепь

$$\xi_\pi \xi_{\pi+1} \cdots \xi_n \xi_{h_1} \xi_{h_1 - 1} \cdots \xi_{a_0} \xi_1 \xi_2 \cdots \xi_{a_1} \xi_{b_1} \xi_{b_1 - 1} \cdots \xi_{m+1}$$

удовлетворяет условию (3).

Случай а3.1.4. $a_2 \leqq \pi \leqq b_1$.

Если $(V^* \cap S) \cap V((\xi_{b_1}, \xi_n] \uparrow P_0) \neq \emptyset$, то приходим к одному из случаев a3.1.1—a3.1.3. Пусть $V^* \cap S \subseteq V([\xi_{a_2}, \xi_{b_1}] \uparrow P_0)$.

Допустим, что для цепи P_0 существует эстафетный путь (M), удовлетворяющий условию

$$V^* \cap S \subseteq V([F(M), L(M)] \uparrow P_0) \subset V([\xi_{a_2}, \xi_{b_1}] \uparrow P_0). \tag{7}$$

Если существует цепь, удовлетворяющая условию (7), то очевидно, существуют числа q_1, q_2 и цепи M_1, M_2 (удовлетворяющие условию (7)), для которых

$$F(M_1) = \xi_{q_1}, \quad q_1 = \min \{i / F(M) = \xi_i\},$$

$$L(M_2) = \xi_{q_2}, \quad q_2 = \max \{i / L(M) = \xi_i\},$$

где min и max берутся по всем цепьям M, удовлетворяющим условию (7). Если же не существует цепи, удовлетворяющей условию (7), то будем предполагать, что

$$q_1 = \min \{i / i \geqq a_2, \ \xi_i \in V^* \cap S\},$$

$$q_2 = \max \{i / i \leqq b_1, \ \xi_i \in V^* \cap S\}.$$

Как и раньше, для тройки

$$A_2 = ([\xi_{q_1}, \xi_{q_2}] \uparrow P_0, \{\xi_{q_1}, \xi_{q_2}\}, P_0)$$

существует некоторый эстафетный путь $(N_1, N_2, \ldots, N_q)$, имеющий минимальную длину среди всех $(N_1, N_2, \ldots, N_i)$, которые определены для A_2 и удовлетворяют условию $L(N_i) \notin V([\xi_{a_2}, \xi_{b_1}] \uparrow P_0)$. В силу симметрии достаточно рассматривать случай $L(N_q) \in V([\xi_{b_1+1}, \xi_n] \uparrow P_0)$. Цикл, порождённый эстафетным путём $(N_1, N_2, \ldots, N_q)$, обозначим через C_1. Пусть $L(N_q) \equiv \xi_r$. Неравенство $r > h_4$ противоречит определению цепи Q_2, следовательно, $r \leqq h_4$. При $r > h_2$ через P_1 будем обозначать цепь

$$\xi_r \xi_{r+1} \ldots \xi_n \xi_{h_1} \xi_{h_1-1} \ldots \xi_1 \xi_{r_1} \xi_{r_1-1} \ldots \xi_{b_1},$$

где $r_1 = \max \{i / i < r, \ \xi_1 \xi_i \in X(G)\}$, а при $r \leqq h_2$ через P_1 будем обозначать цепь

$$\xi_r \xi_{r+1} \ldots \xi_n \xi_{h_1} \xi_{h_1-1} \ldots \xi_{a_0} \xi_1 \xi_2 \ldots \xi_{a_1} \xi_{b_1}.$$

Если $\xi_m \xi_{m+1} \in X(C_2)$, где $C_2 = (C_1 - [\xi_{b_1}, \xi_r] \uparrow P_0) \cup P_1$, то цепь $C_2 - \xi_m \xi_{m+1}$ удовлетворяет условию (3). Пусть $\xi_m \xi_{m+1} \notin X(C_2)$. Если множество $(V(C_2) \cap \cap S) \setminus \{F(N_1)\}$ содержит некоторую вершину u_1, то цепь

$$(C_2 \cup [\xi_{m+1}, F(N_1)] \uparrow P_0) - [F(N_1), u_1] \uparrow P_0$$

удовлетворяет условию (3). Если же $(V(C_2) \cap S) \setminus \{F(N_1)\} = \emptyset$, то по определении цепи N_1 существует цепь, удовлетворяющая условию (7) и, следовательно, условию (3) удовлетворяет цепь

$$(C_2 \cup M_2 \cup ([F(M_2), \xi_m] \uparrow P_0) \cup [\xi_{m+1}, F(N_1)] \uparrow P_0) - [F(N_1), L(M_2)] \uparrow P_0.$$

Лемма доказана.

Лемма 3. *Если* (R_1, R_2) *является S-допустимой парой для графа G, то либо существует цепь* $P = l^1$, *удовлетворяющая хотя бы одному из условий* (1), (2), *либо существует цепь* $P = l^2$ *удовлетворяющая условию*

$$F(P) = L(R_1), \quad L(P) = F(R_2), \quad V(P) \subseteq V(H_r) \cup S, \tag{8}$$
$$V(P) \cap S \subseteq V^* \cap S, \quad |V(P)| \geqq \min(d_1, d_2) + \Delta - 1,$$

где $\Delta = |V^* \cap S|$.

Доказательство. Введём следующие обозначения

$$R_1 = \xi_1 \xi_2 \dots \xi_m, \quad R_2 = \xi_{m+1} \xi_{m+2} \dots \xi_n, \quad \xi_1 \xi_2 \dots \xi_n = P_0,$$
$$a = \max \{i / \xi_1 \xi_i \in X(G)\}, \quad b = \min \{i / \xi_n \xi_i \in X(G)\}$$

где $\xi_m \xi_{m+1}$—фиктивное ребро. Рассмотрим следующие случаи:

Случай б1. Для некоторых чисел s, t имеет место

$$\xi_n \xi_s \in X(G), \quad \xi_1 \xi_t \in X(G), \quad s < t, \quad m \notin \overline{s, t-1}.$$

Существует цепь, удовлетворяющая хотя бы одному из условий (1), (2) (см. лемма 2, случай а1).

Случай б2. $a \leqq m \leqq b-1$.

Для тройки

$$A_1 \equiv (\xi_1 \xi_2 \dots \xi_a, \ \{\xi_a\}, \ P_0)$$

существует некоторая эстафетная цепь $(Q_1, Q_2, \dots, Q_g)$, где $L(Q_g) = \xi_n$ (см. лемма 2, случай а2). Введём обозначения

$$F(Q_i) = \xi_{a_i}, \quad L(Q_i) = \xi_{b_i}, \quad i = \overline{0, g},$$

где $Q_0 \equiv \xi_{a_0} \xi_{b_0}$, $a_0 = 1$ и

$$b_0 = \min \{i / i > a_1, \ \xi_1 \xi_i \in X(G)\}.$$

Без потери общности можем предполагать, что

$$a_g = \max \{i / i < b_{g-1}, \ \xi_n \xi_i \in X(G)\}.$$

Если $b_i \leqq m \leqq a_{i+2} - 1$, $i \in \overline{0, g-2}$, то существует цепь, удовлетворяющая условию (2) (см. лемма 2, случай а2.1). Пусть $a_{h+1} \leqq m \leqq b_h - 1$, $h \in \overline{1, g-2}$. Тогда для P_0 существует некоторый эстафетный путь $(Q_1^1, Q_2^1, \dots, Q_{f_1}^1)$ (где $Q_1^1 = Q_{h-1}$ при $h \geqq 2$ и $Q_1^1 = \xi_1 \xi_a$ при $h = 1$), который является максимальным на цепи $[\xi_{a_{h-1}}, \xi_m] \uparrow P_0$. Аналогично, для P_0 существует некоторый эстафетный путь $(Q_1^2, Q_2^2, \dots, Q_{f_2}^2)$ (где $Q_1^2 = Q_{h+2}$ при $g - h \geqq 3$ и $Q_1^2 = \xi_n \xi_b$ при $g - h = 2$), который является максимальным на цепи $[\xi_{b_{h+2}}, \xi_{m+1}] \uparrow P_0$. Циклы, порождённые эстафетными путями

$$(Q_{h+1}, Q_{h+2}, \dots, Q_g), \quad (Q_0, Q_1, \dots, Q_{h-1}, Q_1^1, Q_2^1, \dots, Q_{f_1}^1),$$

обозначим через C_0 и C_1 соответственно. Объединение всех цепей

$$P_0, Q_i (i \neq h, i \neq h+1), \quad Q_j^1, \ j = \overline{1, f_1}, \quad Q_k^2, \ k = \overline{1, f_2},$$

в графе G определяет некоторый подграф W. Через W_1 и W_2 обозначим компоненты связности графа $W - \{L(Q_{f_1}^1)\}$, а через W_3 и W_4 компоненты связности графа $W - \{L(Q_{f_2}^2)\}$. Без потери общности можем предполагать, что $\xi_1 \in V(W_1) \cap V(W_3)$.

Рассмотрим $(\Delta - 1)$-связный граф

$$G^{(1)} \equiv G - (\{L(Q_{f_1}^1)\} \cup (S \setminus V^*)).$$

Так как $a \geqq d_1 + 1 \geqq \delta(G) - (k - \Delta) + 1 \geqq \Delta + 1$, то $|V(W_1)| \geqq \Delta$. Аналогично, $|V(W_i)| \geqq \Delta$, $i = \overline{2, 4}$. Тогда по теореме Менгера (пункт 2) в графе $G^{(1)}$ подграфа W_1 и W_2 соединены по крайней мере $\Delta - 1$ цепями $l_1, l_2, \ldots, l_{\Delta-1}$ попарно без общих вершин. Очевидно, что

$$|\{i / V(l_i) \cap V(H_2) \neq \emptyset\}| \leqq \min(n_1, n_2),$$

где

$$n_1 = |V(W_1) \cap S|, \quad n_2 = |V(W_2) \cap S|.$$

Поэтому в графе $G - L(Q_{f_1}^1)$ подграфы W_1 и W_2 соединены по крайней мере $\Delta - 1 - \min(n_1, n_2)$ вершинно непересекающимися цепями $E_1, E_2, \ldots, E_\omega$, которые не имеют общих вершин с множеством $V(H_2) \cup (S \setminus V^*)$. Аналогично, в графе $G - L(Q_{f_2}^2)$ подграфы W_3 и W_4 соединены по крайней мере $\Delta - 1 - \min(n_3, n_4)$ вершинно непересекающимися цепями $F_1, F_2, \ldots, F_\sigma$, где

$$n_3 = |V(W_4) \cap S|, \quad n_4 = |V(W_3) \cap S|,$$

которые не имеют общих вершин с множеством $V(H_2) \cup (S \setminus V^*)$. Нетрудно проверить, что $\min(n_1, n_2) + \min(n_3, n_4) \leqq \Delta - 2$, откуда

$$\omega + \sigma \geqq 2\Delta - 2 - \min(n_1, n_2) - \min(n_3, n_4) \geqq \Delta.$$

Если $\Delta = 2$, то цепь $C_0 - \xi_m \xi_{m+1}$ удовлетворяет условию (8). Пусть $\Delta \geqq 3$. Тогда без потери общности можем предполагать, что $\omega \geqq 2$.

Случай б2.1. $h \geqq 2$, $g - h \geqq 3$.

Случай б2.1.1. $L(E_i) \in V(Q_j) \setminus \{F(Q_j), L(Q_j)\}$, $i \in \overline{1, \omega}$, $j \in \overline{h+2, g}$.

Этот случай противоречит определению цепи Q_j.

Случай б2.1.2. $L(E_{i_0}) \in V(Q_{j_0}^2) \setminus \{F(Q_{j_0}^2), L(Q_{j_0}^2)\}$, $i_0 \in \overline{1, \omega}$, $j_0 \in \overline{1, f_2}$.

Случай б2.1.2.1. $F(E_{i_0}) \in V(Q_i) \setminus \{F(Q_i), L(Q_i)\}$, $i \in \overline{1, h-1}$.

Этот случай противоречит определению цепи Q_i.

Случай б2.1.2.2. $F(E_{i_0}) \in V(Q_{i_1}^1) \setminus \{F(Q_{i_1}^1), L(Q_{i_1}^1)\}$, $i_1 \in \overline{1, f_1}$.

Объединение цепей

$$[F(Q_{i_1}^1), F(E_{i_0})] \uparrow Q_{i_1}^1, E_{i_0}, [L(E_{i_0}), F(Q_{j_0}^2)] \uparrow Q_{j_0}^2$$

определяет некоторую простую цепь Q^0. Цепь $C_2-\xi_m\xi_{m+1}$, где цикл C_2 порождён эстафетным путём

$$(Q_0, Q_1, \ldots, Q_{h-1}, Q_1^1, Q_2^1, \ldots, Q_{i_1-1}^1, Q^0, Q_{j_0-1}^2, Q_{j_0-2}^2, \ldots, Q_1^2, Q_{h+2}, Q_{h+3}, \ldots, Q_g),$$

удовлетворяет условию (2).

Случай б2.1.2.3. $F(E_{i_0})\in[\xi_1, \xi_{b_{h-2}})\uparrow P_0$ при $h\geqq 3$ и $F(E_{i_0})\in[\xi_1, \xi_a)\uparrow P_0$ при $h=2$.

Этот случай противоречит определению цепи Q_{h-1}.

Случай б2.1.2.4. $F(E_{i_0})\in[\xi_{b_{h-2}}, L(Q_{f_1}^1)]\uparrow P_0$.

Без потери общности можем предполагать, что

$$F(E_{i_0})\in[L(Q_{f_1-1}^1),\ L(Q_{f_1}^1)]\uparrow P_0.$$

Объединение цепей E_{i_0} и $[L(E_{i_0}), F(Q_{j_0}^2)]\uparrow Q_{j_0}^2$ определяет некоторую простую цепь Q^1. Цикл, порождённый эстафетным путём

$$(Q_0, Q_1, \ldots, Q_{h-1}, Q_1^1, Q_2^1, \ldots, Q_{f_1}^1, Q^1, Q_{j_0-1}^2, Q_{j_0-2}^2, \ldots, Q_1^2, Q_{h+2}, Q_{h+3}, \ldots, Q_g),$$

обозначим через C_3. Тогда цепь $C_3-\xi_m\xi_{m+1}$ удовлетворяет условию (2).

Случай б2.1.3. $L(E_{i_0})\in[\xi_n, L(Q_{f_2}^2)]\uparrow P_0$.

Рассуждения в этом случае проводятся подобно случаям б2.1.2.3, б2.1.2.4.

Случай б2.1.4. $L(E_{i_0})\in[\xi_{m+1}, L(Q_{f_2}^2)]\uparrow P_0$.

Без потери общности можем предположить, что

$$L(E_i)\in[\xi_{m+1}, L(Q_{f_2}^2)]\uparrow P_0, \quad i=\overline{1, \omega},$$

$$L(F_j)\in[\xi_m, L(Q_{f_1}^1)]\uparrow P_0, \quad j=\overline{1, \sigma},$$

$$L(Q_h)\in[\xi_{m+1}, L(Q_{f_2}^2)]\uparrow P_0, \quad F(Q_{h+1})\in[\xi_m, L(Q_{f_1}^1)]\uparrow P_0.$$

Случай б2.1.4.1. $F(E_{i_0})\in V(Q_i)\setminus\{F(Q_i), L(Q_i)\}$, $i\in\overline{1, h-1}$.

Этот случай противоречит определению цепи Q_i.

Случай б2.1.4.2. $F(E_{i_0})\in[\xi_1, \xi_{b_{h-2}})\uparrow P_0$ при $h\geqq 3$ и $F(E_{i_0})\in[\xi_1, \xi_a)\uparrow P_0$ при $h=2$.

Этот случай противоречит определению цепи Q_{h-1}.

Случай б2.1.4.3. $F(E_1)\in V(C_1)$, $F(E_2)\in V(C_1)$.

Цепь

$$((C_0\cup C_4)-[L(E_1), L(E_2)]\uparrow P_0)-\xi_m\xi_{m+1},$$

где

$$C_4=(C_1\cup E_1\cup E_2)-(F(E_1), F(E_2))\uparrow(C_6-\xi_1),$$

удовлетворяет условию (2).

Случай б2.1.4.4.

$$F(E_1)\in(F(Q^1_{\mu+1}),\quad L(Q^1_\mu))\uparrow P_0,\quad \mu\in\overline{1, f_1-1},$$

$$F(E_2)\in(F(Q^1_{\nu+1}),\quad L(Q^1_\nu))\uparrow P_0,\quad \nu\in\overline{1, f_1-1}.$$

Цикл, порождённый эстафетным путём

$$(Q_0, Q_1, \ldots, Q_{h-1}, Q^1_1, Q^1_2, \ldots, Q^1_\mu),$$

обозначим через C_5. Если $\mu=\nu$, то цепь

$$(C_5-[F(E_1),\ F(E_2)]\uparrow P_0)\cup E_1\cup E_2\cup(C_0-(\xi_m\xi_{m+1}\cup[L(E_1), L(E_2)]\uparrow P_0))$$

удовлетворяет условию (2). Если же $\mu<\nu$, то условию (2) удовлетворяет объединение следующих двух цепей

$$(C_4-[F(E_1), L(Q^1_\mu)]\uparrow P_0)\cup([L(Q^1_\mu),\ F(E_2)]\uparrow P_0),$$

$$(C_0-(\xi_m\xi_{m+1}\cup[L(E_1), L(E_2)]\uparrow P_0)\cup E_1\cup E_2.$$

Случай б2.1.4.5. $F(E_1)\in(F(Q^1_{i+1}), L(Q^1_i))\uparrow P_0$, $i\in\overline{1, f_1-1}$, $F(E_2)\in V(C_1)$.

Рассуждения в этом случае проводятся подобно случаям б2.1.4.3—б2.1.4.4.

Случай б2.2. $h\geqq 2$, $g-h=2$.

Так как $\omega\geqq 2$, то этот случай сводится к случаю б2.1.

Случай б2.3. $h=1$, $g-h\geqq 3$.

Если $\sigma\geqq 2$, то приходим к случаю б2.2. Пусть $\sigma=1$, т.е. $\omega\geqq\Delta-1$. Тогда цикл C_0 имеет не меньше, чем $d_2+\Delta$ вершин и, следовательно, цепь $C_0-\xi_m\xi_{m+1}$ удовлетворяет условию (8).

Случай б2.4. $h=1$, $g-h=2$.

Без потери общности можем предполагать, что

$$F(E_i)\in\{\xi_1, \xi_2, \ldots, \xi_{a-1}\},\quad L(E_i)\in\{\xi_{m+1}, \xi_{m+2}, \ldots, \xi_b\},\quad i=\overline{1, \omega},$$

$$F(F_j)\in\{\xi_{b+1}, \xi_{b+2}, \ldots, \xi_n\},\ L(F_j)\in\{\xi_a, \xi_{a+1}, \ldots, \xi_m\},\quad j=\overline{1, \sigma}.$$

Пусть

$$L(F_i)\equiv\xi_{\sigma_i},\quad F(F_i)\equiv\xi_{\sigma'_i},\quad i=\overline{1, \sigma},\quad \sigma_1\leqq\sigma_2\leqq\ldots\leqq\sigma_\sigma,$$

$$\sigma_0\equiv\max\{i/i<\sigma'_1,\ \xi_n\xi_i\in X(G)\}.$$

Цепь $C_6-\xi_m\xi_{m+1}$, где C_6 есть цикл, порождённый эстафетным путём $(F_1, \xi_n\xi_{\sigma_0})$, удовлетворяет условию (8), так как

$$|V(C_6)|\geqq d_2+\omega+\sigma=d_2+\Delta.$$

Случай б3. $b\leqq m\leqq a-1$.

Введём обозначения

$$h_1 = \max\{i/i < a,\ \xi_n\xi_i \in X(G)\},\quad h_2 = \min\{i/i > b,\ \xi_1\xi_i \in X(G)\},$$

$$h_3 = \max\{i/i \leqq b,\ \xi_1\xi_i \in X(G)\},\quad h_4 = \min\{i/i \geqq a,\ \xi_n\xi_i \in X(G)\}.$$

Если $h_2 \leqq m$ или $h_1 \geqq m+1$, то приходим к случаю б1. Пусть $h_2 \geqq m+1$, $h_1 \leqq m$. Для P_0 существуют некоторые эстафетные пути $(Q_1^1, Q_2^1, \ldots, Q_{f_1}^1)$, $(Q_1^2, Q_2^2, \ldots, Q_{f_2}^2)$, где $Q_1^1 = \xi_1\xi_{h_3}$, $Q_1^2 = \xi_n\xi_{h_4}$, которые являются максимальными на цепях $[\xi_1, \xi_m]\uparrow P_0$ и $[\xi_n, \xi_{m+1}]\uparrow P_0$ соответственно. В случае

$$L(Q_{f_1}^1) \in (\xi_b, \xi_m]\uparrow P_0 \quad \text{или} \quad L(Q_{f_2}^2) \in (\xi_a, \xi_{m+1}]\uparrow P_0$$

рассуждения проводятся подобно случаю а2.1 (см. лемма 2). Пусть

$$L(Q_{f_1}^1) \in [\xi_1, \xi_b]\uparrow P_0,\quad L(Q_{f_2}^2) \in [\xi_n, \xi_a]\uparrow P_0.$$

Объединение всех цепей

$$P_0,\ Q_i^1,\ Q_j^2,\ i = \overline{1, f_1},\ j = \overline{1, f_2},$$

в графе G определяет некоторый подграф W. Через W_1 и W_2 обозначим компоненты связности графа $W - \{L(Q_{f_1}^1)\}$, а через W_3 и W_4—компоненты связности графа $W - \{L(Q_{f_2}^2)\}$. Без потери общности можем предполагать, что $\xi_1 \in V(W_1) \cap V(W_3)$. Введём обозначения

$$\Delta_1 = |V(W_1) \cap S|,\quad \Delta_2 = |V(W_4) \cap S|,$$

$$G^{(2)} = G - \left(\{L(Q_{f_1}^1)\} \cup (S \setminus V^*) \cup (V(W_1) \cap S)\right).$$

Очевидно, что $\Delta_1 + \Delta_2 \leqq \Delta - 2$. Так как $|V(W_2)| \geqq d_2 \geqq \delta(G) - k + \Delta \geqq \Delta$ и $k(G^{(2)}) \geqq \Delta - \Delta_1 - 1 \geqq \Delta_2 + 1 \geqq 1$, то по теореме Менгера (пункт 3) в графе $G^{(2)}$ подграфы $W_1^* \equiv \langle V(W_1) \setminus S \rangle$ и W_2 соединены по крайней мере $\Delta - \Delta_1 - 1$ цепями l_i^1, $i = \overline{1, \Delta - \Delta_1 - 1}$ (где $L(l_i^1) \in V(W_2)$), которые ввиду $V(W_1^*) \cap S = \emptyset$ не имеют общих вершин с множеством $V(H_2)$ и попарно не пересекаются в подграфе W_2. Из определения эстафетного пути $(Q_1^1, Q_2^1, \ldots, Q_{f_1}^1)$ следует $L(l_i^1) \in [\xi_{m+1}, \xi_n]\uparrow P_0$, $i \in \overline{1, \Delta - \Delta_1 - 1}$. В случае $L(l_i^1) \in \left(L(Q_{f_2}^2), \xi_n\right]\uparrow P_0$, $i \in \overline{1, \Delta - \Delta_1 - 1}$, рассуждения проводятся подобно случаю а2.1 (см. лемма 2). Пусть

$$L(l_i^1) \in [\xi_{m+1}, L(Q_{f_2}^2)]\uparrow P_0,\quad i = \overline{1, \Delta - \Delta_1 - 1}.$$

Тогда

$$\left|V\left([\xi_{m+1}, L(Q_{f_2}^2)]\uparrow P_0\right)\right| \geqq \Delta - \Delta_1 - 1.$$

В силу симметрии имеем

$$\left|V\left([L(Q_{f_1}^1), \xi_m]\uparrow P_0\right)\right| \geqq \Delta - \Delta_2 - 1.$$

Рассмотрим Δ_1-связный граф

$$G^{(3)} = G - \left(\{L(Q_{f_1}^1)\} \cup (S \setminus V^*) \cup (V(W_2) \cap S)\right).$$

Случай б3.1. $|V(W_2) \setminus S| < \Delta_1$ или $|V(W_3) \setminus S| < \Delta_2$.

Без потери общности можем предполагать, что $|V(W_2)\setminus S|<\Delta_1$. Так как $|V(W_1)|\geqq\Delta_1$, то по теореме Менгера (пункт 3) в графе $G^{(3)}$ подграфы W_1 и $W_2^*\equiv\langle V(W_2)\setminus S\rangle$ соединены по крайней мере

$$\min\left(|V(W_1)|, |V(W_2^*)|, \Delta_1\right) = |V(W_2^*)|$$

цепями l_i^2 попарно без общих вершин, которые ввиду $W_2^*\cap S=\emptyset$ не имеют общих вершин с множеством $V(H_2)$. Тогда $\xi_n=L(l_{i_0}^1)$ для некоторого $i_0\in\overline{1, |V(W_2^*)|}$. Из определения цепи P_0 следует, что $l_{i_0}^1\in X(G)$, что противоречит определению числа b.

Случай б3.2. $|V(W_2)\setminus S|\geqq\Delta_1$, $|V(W_3)\setminus S|\geqq\Delta_2$.

По теореме Менгера (пункт 2) в графе $G^{(3)}$ подграфы W_1 и $\langle V(W_2)\setminus S\rangle$ соединены по крайней мере Δ_1 цепями $E_1, E_2, \ldots, E_{\Delta_1}$ попарно без общих вершин. В силу симметрии можем предполагать, что в графе $G^{(3)}$ подграфы W_4 и $\langle V(W_3)\setminus S\rangle$ соединены по крайней мере Δ_2 цепями $F_1, F_2, \ldots, F_{\Delta_2}$ попарно без общих вершин. Введём обозначения

$$F(E_i) = \xi_{\omega_i}, \quad L(E_i) = \xi_{\omega_i'}, \quad i = \overline{1, \Delta_1},$$

$$F(F_j) = \xi_{\sigma_i}, \quad L(F_j) = \xi_{\sigma_j'}, \quad j = \overline{1, \Delta_2}.$$

Без потери общности можем предполагать, что

$$\omega_1 \geqq \omega_2 \geqq \ldots \geqq \omega_{\Delta_1}, \quad \sigma_1 \leqq \sigma_2 \leqq \ldots \leqq \sigma_{\Delta_2},$$

$$\xi_{\omega_i'}\in[\xi_{m+1}, L(Q_{f_2}^2)]\uparrow P_0, \quad i = \overline{1, \Delta_1},$$

$$\xi_{\sigma_i'}\in[\xi_m, L(Q_{f_1}^1)]\uparrow P_0, \quad i = \overline{1, \Delta_2}.$$

Случай б3.2.1. $\omega_1'<a$, $\sigma_1'>b$.

Если $\Delta_1=0$, то

$$\left|V\left([\xi_{m+1}, L(Q_{f_2}^2)]\uparrow P_0\right)\right| \geqq \Delta-1,$$

откуда следует, что цепь

$$\xi_{m+1}\xi_{m+2}\cdots\xi_n\xi_b\xi_{b+1}\cdots\xi_m \tag{9}$$

содержит по меньшей мере $d_2+\Delta-1$ вершин и, следовательно, удовлетворяет условию (8).

Допустим, что $\Delta_1=1$. Отсюда

$$\left|V\left([\xi_{m+1}, L(Q_{f_2}^2)]\uparrow P_0\right)\right| = \Delta-2,$$

так как иначе приходим к случаю $\Delta_1=0$. Тогда для некоторых чисел $i, j\in\overline{1, \Delta-\Delta_1-1}$ имеет место:

$$L(l_i^1) = \xi_{m+1}, \quad L(l_j^1) = \xi_{m+2}.$$

Ребро $\xi_{m+1}\xi_{m+2}$ в цепи (9) заменим цепью

$$l_i^1\cup l_j^1\cup[F(l_i^1), F(l_j^1)]\uparrow P_0.$$

Полученная при этом цепь удовлетворяет условию (8).

Пусть $\Delta_1 \geqq 2$. Без потери общности можем предполагать, что $\Delta_2 \geqq 2$. Введём обозначения

$$\omega_0 = \min \{i/i > \omega_1', \ \xi_1 \xi_i \in X(G)\},$$

$$\sigma_0 = \max \{i/i < \sigma_1', \ \xi_n \xi_i \in X(G)\}.$$

Цепь

$$\xi_m \xi_{m-1} \dots \xi_b \xi_n \xi_{n-1} \dots \xi_{\omega_0} \xi_1 \xi_2 \dots \xi_{\omega_1} \xi_{\omega_1'} \xi_{\omega_1'-1} \dots \xi_{m+1}$$

очевидно содержит по меньшей мере

$$d_2 + \Delta + (\sigma_1' - \sigma_0) - (\omega_0 - \omega_1')$$

вершин, а цепь

$$\xi_{m+1} \xi_{m+2} \dots \xi_a \xi_1 \xi_2 \dots \xi_{\sigma_0} \xi_n \xi_{n+1} \dots \xi_{\sigma_1} \xi_{\sigma_1'} \xi_{\sigma_1'+1} \dots \xi_m$$

содержит по меньшей мере

$$d_1 + \Delta + (\omega_0 - \omega_1') - (\sigma_1' - \sigma_0)$$

вершин. Длиннейшая из этих цепей содержит по меньшей мере $\min(d_1, d_2) + \Delta - 1$ вершин и, следовательно, удовлетворяет условию (8).

Случай б3.2.2. $\omega_1' \geqq a$, $\sigma_1' \leqq b$.

Без потери общности можем предполагать, что

$$\xi\sigma_1 \in (L(Q_{f_2}^2), L(Q_{f_2-1}^2)] \uparrow P_0.$$

Цикл, порождённый эстафетным путём

$$(\xi_1 \xi_{h_0}, Q_1^2, Q_2^2, \dots, Q_{f_2}^2),$$

где

$$h_0 = \max \{i/\xi_n \xi_i \in X(G), \ \xi_i \in (L(Q_1^2), \xi_1) \uparrow P_0\},$$

обозначим через C_7. Цепь $C_9 - \xi_m \xi_{m+1}$, где

$$C_8 = (C_7 - [L(Q_{f_2}^2), \xi_{\sigma_1}] \uparrow P_0) \cup F_1 \cup [\xi_{\sigma_1'}, L(Q_{f_2}^2)] \uparrow P_0,$$

$$C_9 = (C_8 - [\xi_a, \xi_{\omega_1'}] \uparrow P_0) \cup [\xi_1, \xi_{\omega_1}] \uparrow P_0,$$

содержит по меньшей мере $d_2 + \Delta + (b - \sigma_1') - (\omega_1' - a)$ вершин. В силу симметрии существует подобная цепь, имеющая по меньшей мере $d_1 + \Delta + (\omega_1' - a) - (b - \sigma_1')$ вершин. Длиннейшая из этих двух цепей имеет по меньшей мере $\min(d_1, d_2) + \Delta$ вершин и, следовательно, удовлетворяет условию (8).

Случай б3.2.3. $\omega_1' < a$, $\sigma_1' \leqq b$ или $\omega_1' \geqq a$, $\sigma_1' > b$.

Этот случай сводится к случаям б3.2.1 и б3.2.2.

Случай б3.3. $\Delta = 2$.

Этот случай сводится к случаям б2.2.2 и б3.1.

Случай б4. $a < b$, $m \notin \overline{a, b-1}$.

Этот случай сводится и случаям б2 и б3.

Лемма доказана.

Доказательство теоремы 2. Пусть $v(G)=v$, $\delta(G)=\delta$, $k(G)=k$, $h(G)=h$ и пусть $S=\{v_1, v_2, \ldots, v_k\}$ является разделяющим множеством вершин графа G, удаление которого порождает компоненты связности $G_1, G_2, \ldots, G_t$. Введём обозначения

$$H_1 = \bigcup_{i=1}^{t-1} G_i, \quad H_2 = G_t.$$

Без потери общности можем предполагать, что $|V(H_2)| \leqq |V(H_1)|$.

Случай 1. $|V(H_1)| \geqq 2\delta - k - 1$.

Допустим, что некоторая простая цепь $x=P_1$ удовлетворяет следующим условиям:

а1. $F(x) \in S$, $L(x) \in S$, $V(x) \subseteqq V(H_1) \cup S$.

а2. Для любой цепи $x=P$, удовлетворяющей условию а1, имеет место

$$|V(P) \cap S| \leqq |V(P_1) \cap S|.$$

а3. Для любой цепи $x=P$, удовлетворяющей условиям а1, а2, имеет место

$$|\{i / V(P) \cap V(G_i) \neq \emptyset\}| \leqq |\{i / V(P_1) \cap V(G_i) \neq \emptyset\}|.$$

а4. Для любой цепи $x=P$, удовлетворяющей условиям а1, а2, а3, имеет место

$$|V(P)| \leqq |V(P_1)|.$$

Без потери общности можем предполагать, что $V(P_1) \cap S = \{v_1, v_2, \ldots, v_\Delta\}$. Если $N(v_i) \cap V(G_j) = \emptyset$ для некоторых $i \in \overline{1, k}$, $j \in \overline{1, t}$, то множество $S \setminus \{v_i\}$ для графа G может служить $k-1$ вершинным разделяющим множеством, что противоречит k-связности графа G. Пусть

$$N(v_i) \cap V(G_j) \neq \emptyset, \quad i = \overline{1, k}, \quad j = \overline{1, t}.$$

Тогда для любой пары вершин (v_i, v_{i+1}) существует пара (u_i^1, u_i^2), удовлетворяющая условию

$$u_i^1 \in V(G_i), \; u_i^2 \in V(G_i), \; v_i u_i^1 \in X(G), \; v_{i+1} u_i^2 \in X(G),$$

где $i \in \overline{1, \min(t, k) - 1}$. Так как $G_1, G_2, \ldots, G_t$—связные компоненты, то пара вершин (u_i^1, u_i^2) для любого $i \in \overline{1, \min(t, k) - 1}$ соединена цепью N_i, для которой $V(N_i) \subseteqq V(G_i)$. Объединение всех цепей

$$v_i u_i^1, \; v_{i+1} u_i^2, \; N_i, \; i = \overline{1, \min(t, k) - 1},$$

в графе G определяет некоторую простую цепь (которая удовлетворяет условию а1), содержащую $\min(t, k)$ вершин из множества S и проходящую через $\min(t, k) - 1$ компонент связности графа $G - S$. Отсюда следует, что P_1 существует, а из условий а2, а3 следует, что $\Delta \geqq \min(t, k)$ и

$$\gamma_1 \equiv |\{i/V(G_i) \cap V(P_1) \neq \emptyset\}| \geqq \min(t, k) - 1.$$

Допустим, что для некоторого числа $i \in \overline{1, t-1}$ имеет место

$$V(P_1) \cap V(G_i) \neq \emptyset, \quad V(G_i) \setminus V(P_1) \neq \emptyset. \tag{10}$$

Без потери общности можем предполагать, что $i = t-1$. Через T_1 обозначим некоторую компоненту связности графа $\langle V(G_{t-1}) \setminus V(P_1) \rangle$. Введём обозначения

$$M_1 = N(V(T_1)), \quad B = \{v \in M_1 / \exists u \notin S(uv \in X(P_1))\},$$

$$D = M_1 \setminus S, \quad E = M_1 \cap S, \quad F = \{v \in V(P_1) / \exists u \in D(uv \in X(P_1))\}.$$

Из определения множества M_1 и из неравенства $k(G) \geqq k$ следует $M_1 \subseteq \subseteq V(P_1) \cup S$ и $|M_1| \geqq k$. Так как G_{t-1}—связная компонента, то $D \neq \emptyset$. Если $B = \emptyset$, то $F \subseteq S$, $F \cap M_1 = \emptyset$, откуда $|F| \geqq |D| + 1$, $F \cap E = \emptyset$ и, следовательно,

$$|M_1| = |D| + |E| \leqq |F| + |E| - 1 \leqq |S| - 1 = k - 1,$$

что невозможно. Таким образом $B \neq \emptyset$. По определению множества B существуют вершины u_1, u_2, u_3 такие, что

$$u_1 u_2 \in X(P_1), \quad u_2 u_3 \in X(G), \quad u_1 \notin S, \quad u_3 \in T_1.$$

Пример подграфа $(P_1 \cup u_2 u_3) - u_1 u_2$, который состоит из двух непересекающихся цепей, доказывает существование S-допустимой пары цепей $(x, y) = (R_1, R_2)$, удовлетворяющей условиям:

б1. $L(x) = F(P_1)$, $F(y) = L(P_1)$, $F(x) \in V(H_1)$, $L(y) \in V(H_1)$,

$$[F(P_1), u_1] \uparrow P_1 \subseteq x, \quad [L(P_1), u_2] \uparrow P_1 \subseteq y,$$

$$|V_1^*| > |V(P_1)|, \quad V_1^* \subseteq V(H_1) \cup S,$$

где $V_1^* = V(x) \cup V(y)$, $u_1 u_2 \in X(P_1)$.

б2. Для любой S-допустимой пары цепей $(x, y) = (l_1, l_2)$, удовлетворяющей условию б1, имеет место

$$|(V(l_1) \cup V(l_2)) \cap S| \leqq |V_1^* \cap S|.$$

б3. Для любой S-допустимой пары цепей $(x, y) = (l_1, l_2)$, удовлетворяющей условиям б1, б2, имеет место

$$|\{i/V(G_i) \cap (V(l_1) \cup V(l_2))\}| \leqq |\{i/V(G_i) \cap V_1^* \neq \emptyset\}|.$$

б4. Для любой S-допустимой пары цепей $(x, y) = (l_1, l_2)$, удовлетворяющей условиям б1, б2, б3, имеет место

$$|V(l_1) \cup V(l_2)| \leqq |V_1^*|.$$

Без потери общности можем предполагать, что

$$V_1^* \cap S = \{v_1, v_2, \ldots, v_\pi\}, \quad \Delta \leqq \pi \leqq k.$$

Пусть

$$d_1 = |N(F(R_1)\cap V_1^*|, \quad d_2 = |N(L(R_2))\cap V_1^*|.$$

По лемме 2 для S-допустимой пары цепей R_1, R_2 существует цепь $P=P_2$, удовлетворяющая хотя бы одному из условий (1), (2), (3) при $\pi \geqq 3$ и хотя бы одному из условий (1), (2) (4) при $\pi=2$.

Если же условие (10) не имеет место, то будем предполагать, что $P_2=P_1$, $\pi=\Delta$, $V_1^*=V(P_1)$.

Пусть простая цепь $x=P_3$ удовлетворяет условиям:

в1. $$V(x)\subseteqq V(H_2)\cup S,\ F(x)=F(P_2),\ L(x)=L(P_2),$$
$$V(x)\cap V_1^*\subseteqq \{F(P_2),\ L(P_2)\}.$$

в2. Для любой цепи $x=P$, удовлетворяющей условию в1, имеет место

$$|V(P)\cap S| \leqq |V(P_3)\cap S|.$$

в3. Для любой цепи $x=P$, удовлетворяющей условиям в1, в2, имеет место

$$|V(P)| \leqq |V(P_3)|.$$

Без потери общности можем предполагать, что $V(P_3)\cap S=\{F(P_3), L(P_3), v_\pi, v_{\pi+1}, \dots, v_r\}$, если $\pi\geqq 3$ или $\pi=2$, $V(P_2)\cap S=V_1^*\cap S$. Если же $\pi=2$, $V(P_2)\cap S\neq V_1^*\cap S$, то будем предполагать, что $V(P_3)\cap S=\{F(P_3), L(P_3), v_4, v_5, \dots, v_r\}$. При $|V(P_3)\cap S|=2$ будем предполагать, что $r=\pi$.

Если

(11) $$V(G_t)\setminus V(P_3)\neq \emptyset,$$

то как и при условии (10), можем показать, что существует S-допустимая пара цепей $(x, y)=(R_3, R_4)$, удовлетворяющая условиям:

г1. $$L(x)=F(P_3),\ F(y)=L(P_3),\ F(x)\in V(H_2),\ L(y)\in V(H_2),$$
$$[F(P_3), u]\uparrow P_3 \subseteqq x, \quad [L(P_3), v]\uparrow P_3\subseteqq y, \quad V_2^*\subseteqq V(H_2)\cup S,$$
$$V_2^*\cap V_1^*\subseteqq \{F(P_2), L(P_2)\}, \quad |V_2^*| > |V(P_3)|,$$

где $V_2^*=V(x)\cup V(y)$, $uv\in X(P_3)$.

г2. Для любой S-допустимой пары цепей $(x, y)=(l_1, l_2)$, удовлетворяющей условию г1, имеет место

$$|(V(l_1)\cup V(l_2))\cap S| \leqq |V_2^*\cap S|.$$

г3. Для любой S-допустимой пары цепей $(x, y)=(l_1, l_2)$, удовлетворяющей условиям г1, г2, имеет место

$$|V(l_1)\cup V(l_2)| \leqq |V_2^*|.$$

Без потери общности можем предполагать, что

$$V_2^*\cap S=\{F(P_2), L(P_2), v_{\pi+1}, v_{\pi+2}, \dots, v_\omega\},$$

если $\pi\geqq 3$ или $\pi=2$, $P_2\cap S=V_1^*\cap S$. Если же $\pi=2$, $P_2\cap S\neq V_1^*\cap S$, то будем предполагать, что

$$V_2^* \cap S = \{F(P_2), L(P_2), v_4, v_5, \ldots, v_\omega\}.$$

При $|V_2^* \cap S| = 2$ будем предполагать, что $\omega = \pi$.

По лемме 3 для S-допустимой пары цепей R_3, R_4 существует цепь $P = P_4$, удовлетворяющая хотя бы одному из условий (1), (2), (8).

Если условие (11) не имеет место, то будем предполагать, что $P_4 = P_3$, $\omega = r$, $V_2^* = V(P_3)$.

Пусть

$$d_3 = |N(F(R_3)) \cap V_2^*|, \quad d_4 = |N(L(R_4)) \cap V_2^*|.$$

Случай 1.1. $V(H_1) \setminus V(P_1) = \emptyset$, $V(H_2) \setminus V(P_3) = \emptyset$.

По предположению условия (10) и (11) не возможны, т.е. $P_2 = P_1$, $P_4 = P_3$, откуда

$$|V(P_2)| \geqq |V(H_1)| + |V(P_2) \cap S| \geqq 2\delta - k - 1 + \Delta,$$

$$|V(P_4)| \geqq |V(H_2)| + |V(P_4) \cap S| \geqq \delta - k + 1 + r - \Delta + 2.$$

Если $r = k$, то

$$h \geqq |V(P_2 \cup P_4)| \geqq |V(P_2)| + |V(P_4)| - 2 \geqq 3\delta - k.$$

Пусть $r < k$. Через T_2—обозначим компоненту связности графа $\langle v_{r+1}, v_{r+2}, \ldots, v_k \rangle$. Для множества $N(V(T_2)) \equiv M_2$, очевидно, имеет место

$$M_2 = M_2 \cap (V(P_2) \cup V(P_4)), \quad |M_2| \geqq k.$$

Пусть $M_2 \equiv \{\eta_1, \eta_2, \ldots, \eta_\tau\}$, где последовательность определена по некоторым направлением цикла $P_2 \cup P_4$. По определению множества T_2 для любого $i \in \overline{1, \tau}$ существует цепь Q_i, удовлетворяющая условиям

$$F(Q_i) = \eta_i, \; L(Q_i) = \eta_{i+1}, \quad V(Q_i) \setminus \{\eta_i, \eta_{i+1}\} \subseteqq V(T_2),$$

где $\eta_{\tau+1} = \eta_1$. Если для некоторого числа $i_0 \in \overline{1, \tau}$ по некоторым направлением цикла $P_2 \cup P_4$ имеет место

$$S \cap V(\eta_{i_0}, \eta_{i_0+1}) \uparrow (P_2 \cup P_4) = \emptyset,$$

то цикл

$$((P_2 \cup P_4) - [\eta_{i_0}, \eta_{i_0+1}] \uparrow (P_2 \cup P_4)) \cup Q_{i_0}$$

противоречит либо определению цепи P_2 либо цепи P_4. Тогда

$$S \cap V(\eta_i, \eta_{i+1}) \uparrow (P_2 \cup P_4) \neq \emptyset, \quad i = \overline{1, \tau},$$

откуда следует $r \geqq \tau \geqq k$, что противоречит предположению.

Случай 1.2. $V(H_1) \setminus V(P_1) \neq \emptyset$, $V(H_2) \setminus V(P_3) = \emptyset$.

По предположению, условие (11) не имеет место, т.е. $P_4 = P_3$ и

$$|V(P_4)| \geqq (\delta - k + 1) + (\omega - \pi + 2) \geqq \delta - k + 3.$$

Случай 1.2.1. Условие (10) не имеет место, т.е. для любого $i \in \overline{1, t-1}$ имеет место

$$V(P_1) \cap V(G_i) = \emptyset \quad \text{или} \quad V(G_i) \setminus V(P_1) = \emptyset.$$

Поскольку $V(H_1) \setminus V(P_1) \neq \emptyset$, то $\gamma_1 \leqq t-2$. С другой стороны $\gamma_1 \geqq \min(t-1, k-1)$. Тогда $t-1 \geqq k$, откуда

$$\gamma_1 \geqq k-1, \quad \Delta \geqq \min(t, k) \geqq \min(k+1, k) \geqq k.$$

Без потери общности можем предполагать, что

$$\{1, 2, \ldots, k-1\} \subseteqq \{i / V(P_1) \cap V(G_i) \neq \emptyset\}.$$

По предположению $V(G_i) \setminus V(P_1) = \emptyset$, $i = \overline{1, k-1}$. Отсюда ввиду $|V(G_i)| \geqq \delta - k + 1$, $i = \overline{1, t}$, имеем

$$|V(P_1)| \geqq \left| \bigcup_{i=1}^{k-1} V(G_i) \right| + |S| \geqq (k-1)(\delta - k + 1) + k,$$

и, следовательно,

$$h \geqq |V(P_1)| + |V(P_4)| - 2 \geqq (k-1)(\delta - k + 1) + k + (\delta - k + 3) - 2 =$$
$$= (\delta - k)(k-3) + 3\delta - k \geqq 3\delta - k.$$

Случай 1.2.2. Условие (10) имеет место, т.е. $P_2 \neq P_1$.

Цепь P_2 не удовлетворяет условию (1), так как иначе P_2 будет противоречить определению цепи P_1. Следовательно, цепь P_2 по лемме 2 удовлетворяет хотя бы одному из условий (2), (3) при $\pi \geqq 3$ и хотя бы одному из условий (2), (4) при $\pi = 2$.

Для любой вершины $\xi \in \{v_{\pi+1}, v_{\pi+2}, \ldots, v_k\}$ имеет место $F(R_1)\xi \notin X(G)$ или $L(R_2)\xi \notin X(G)$, так как иначе цепь $R_1 \cup R_2 \cup F(R_1)\xi \cup L(R_2)\xi$ будет противоречить определению цепи P_1. Отсюда

$$d_1 + d_2 \geqq 2\delta - (k - \pi).$$

Если $\omega = k$ и $\pi \geqq 3$, то по лемме 2

$$|V(P_2)| \geqq d_1 + d_2 - 1 \geqq 2\delta - k + \pi - 1, \quad |V(P_4)| \geqq \delta - \pi + 3,$$

откуда

$$h \geqq |V(P_2)| + |V(P_4)| - 2 \geqq (2\delta - k + \pi - 1) + (\delta - \pi + 3) - 2 \geqq 3\delta - k.$$

Если $\omega = k$, $\pi = 2$ и $|V(P_4) \cap S| = k$, то

$$|V(P_2)| \geqq 2\delta - k + \pi - 1 = 2\delta - k + 1, \quad |V(P_4)| \geqq (\delta - k + 1) + k$$

и, следовательно, $h \geqq 3\delta - k$.

Если $\omega = k$, $\pi = 2$ и $|V(P_4) \cap S| = k-1$, то по лемме 2 цепь P_2 удовлетворяет условию (4), откуда

$$|V(P_2)| \geqq d_1 + d_2 + 1 \geqq 2\delta - k + 3,$$

$$h \geqq (2\delta - k + 3) + (\delta - k + 1 + k - 1) - 2 \geqq 3\delta - k.$$

Пусть $\omega < k$ и пусть T_3—некоторая компонента связности графа $\langle v_{\omega+1}, v_{\omega+2}, \ldots, v_k\rangle$. Как и раньше,

$$|M_3 \cap (V_1^* \cup V_2^*)| \leqq \omega < k, \quad |M_3| \geqq k,$$

где $M_3 = N(V(T_3))$. Отсюда

$$M_3 \cap (V(H_1) \setminus V_1^*) \neq \emptyset,$$

и, следовательно,

$$(N(F(R_1)) \cap N(L(R_2))) \cap V(T_3) = \emptyset,$$

что приводит нас к случаю $\omega = k$.

Случай 1.3. $V(H_1) \setminus V(P_1) \neq \emptyset$, $V(H_2) \setminus V(P_3) \neq \emptyset$.

Пусть $\pi \geqq 3$. По предположению, условие (11) выполняется, т.е. $P_4 \neq P_3$. Цепь P_4 не удовлетворяет условию (1), так как иначе P_1 будет противоречить определению цепи P_3. Тогда по лемме 3 цепь P_4 удовлетворяет хотя бы одному из условий (2), (8). Если условие (10) не имеет место, то учитывая слабое неравенство $|V(P_4)| \geqq \delta - k + 3$, приходим к случаю 1.2.1. Пусть условие (10) выполняется, т.е. $P_2 \neq P_1$. Так как P_2 не удовлетворяет условию (1), то по лемме 2 удовлетворяет хотя бы одному из условий (2), (3). Для любой вершины $\xi \in \{v_{\pi+1}, v_{\pi+2}, \ldots, v_k\}$ имеет место либо $F(R_1)\xi \notin X(G)$ либо $L(R_2)\xi \notin X(G)$. Поэтому

$$d_1 + d_2 \geqq 2\delta - (k - \Delta_2).$$

Если $\omega = k$, то для P_2 по лемме 2

$$|V(P_2)| \geqq d_1 + d_2 - 1 \geqq 2\delta - k + \pi - 1,$$

а для P_4 по лемме 3 либо

$$|V(P_4)| \geqq d_3 + d_4 - 1 \geqq 2(\delta - \pi + 2) - 1 \geqq \delta - \pi + 3,$$

либо

$$|V(P_4)| \geqq \min(d_3, d_4) + 1 \geqq \delta - \pi + 3,$$

откуда

$$h \geqq |V(P_2)| + |V(P_4)| - 2 \geqq 3\delta - k.$$

Пусть $\omega < k$ и пусть T_4—некоторая компонента связности графа $T = \langle v_{\omega+1}, v_{\omega+2}, \ldots, v_k\rangle$. Как и раньше,

$$|M_4 \cap (V_1^* \cup V_2^*)| \leqq \omega < k, \quad |M_4| \geqq k,$$

где $M_4 = N(V(T_4))$. Отсюда

$$|M_4 \cup (V(H_1) \setminus V_1^*)| \neq \emptyset \quad \text{или} \quad M_4 \cap (V(H_2) \setminus V_2^*) \neq \emptyset.$$

Поскольку T_4—произвольная компонента связности графа T, то существуют множества S_1, S_2, для которых

$$V(T) = S_1 \cup S_2, \quad S_1 \cap S_2 = \emptyset,$$

$$N(S_1) \cap (V(H_1) \setminus V_1^*) \neq \emptyset, \quad N(S_2) \cap (V(H_2) \setminus V_2^*) \neq \emptyset.$$

По определению цепей R_1, R_2, R_3, R_4 имеем

$$(N(F(R_1))\cup N(L(R_2)))\cap S_1 = \emptyset, \quad (N(F(R_3))\cup N(L(R_4)))\cap S_2 = \emptyset.$$

Поэтому

$$d_1+d_2 \geqq 2\delta-|S_2|-(\omega-\pi), \quad \min(d_3, d_4) \geqq \delta-|S_1|-(\pi-2),$$

откуда

$$h \geqq |V(P_2)|+|V(P_4)|-2 \geqq 3\delta-(|S_1|+|S_2|+\omega) \geqq 3\delta-k.$$

При $\pi=2$ рассуждения здесь можно провести аналогично случаю 1.2.2.

Случай 1.4. $V(H_1)\setminus V(P_1)=\emptyset$, $V(H_2)\setminus V(P_3)\neq\emptyset$.

Этот случай сводится к случаям 1.1—1.3.

Случай 2. $|V(H_1)|\leqq 2\delta-2k+1$.

Если $t-1\geqq 2$, то $|V(H_1)|\geqq 2(\delta-k+1)$, что противоречит предположению. Пусть $t=2$ и пусть простая цепь $x=P_1$ удовлетворяет следующим условиям:

д1. $F(x)\in S$, $L(x)\in S$, $V(x)\subseteq V(H_2)\cup S$.

д2. Для любой цепи $x=P$, удовлетворяющей условию д1, имеет место

$$|V(P)\cap S| \leqq |V(P_1)\cap S|.$$

д3. Для любой цепи $x=P$, удовлетворяющей условиям д1, д2, имеет место

$$|V(P)| \leqq |V(P_1)|.$$

Без потери общности можем предполагать, что $V(P_1)\cap S=\{v_1, v_2, \dots, v_\Delta\}$. Если $V(H_2)\setminus V(P_1)\neq\emptyset$, то как и раньше (случай 1) существует S-допустимая пара цепей $(x, y)=(R_1, R_2)$, удовлетворяющая следующим условиям:

е1. $$L(x) = F(P_1), \quad F(y) = L(P_1), \quad F(x)\in V(H_2), \quad L(y)\in V(H_2),$$

$$[F(P_1), u]\uparrow P_1 \subseteq x, \quad [L(P_1), v]\uparrow P_1 \subseteq y,$$

$$|V_1^*| > |V(P_1)|, \quad V_1^* \subseteq V(H_2)\cup S,$$

где $V_1^*=V(x)\cup V(y)$, $uv\in X(P_1)$.

е2. Для любой S-допустимой пары цепей $(x, y)=(l_1, l_2)$, удовлетворяющей условию е1, имеет место

$$|(V(l_1)\cup V(l_2))\cap S| \leqq |V_1^*\cap S|.$$

е3. Для любой S-допустимой пары цепей $(x, y)=(l_1, l_2)$, удовлетворяющей условиям е1, е2, имеет место

$$|V(l_1)\cup V(l_2)| \leqq |V_1^*|.$$

Без потери общности можем предполагать, что $V_1^*\cap S=\{v_1, v_2, \dots, v_\pi\}$, где $\Delta\leqq\pi\leqq k$. Пусть

$$d_1 = |N(F(R_1))\cap V_1^*|, \quad d_2 = |N(L(R_2))\cap V_1^*|.$$

По лемме 2 для S-допустимой пары цепей R_1, R_2 существует цепь P_2, удовлетворяющая хотя бы одному из условий (1), (2), (3) при $\pi \geqq 3$ и хотя бы одному из условий (1), (2), (4) при $\pi = 2$.

Если же $V(H_2) \setminus V(P_1) = \emptyset$, то будем предполагать, что $P_2 = P_1$, $\pi = \Delta$, $V_1^* = V(P_1)$.

Пусть простая цепь $x = P_3$ удовлетворяет следующим условиям:

ж1. $V(x) \subseteqq V(H_1) \cup S$, $F(x) = F(P_2)$, $L(x) = L(P_2)$,
$V(x) \cap V_1^* \subseteqq \{F(P_2),\ L(P_2)\}$

ж2. Для любой цепи $x = P$, удовлетворяющей условию ж1, имеет место

$$|V(P) \cap S| \leqq |V(P_3) \cap S|.$$

ж3. Для любой цепи $x = P$, удовлетворяющей условиям ж1, ж2, имеет место

$$|V(P)| \leqq |V(P_3)|.$$

Без потери общности можем предполагать, что

$$V(P_3) \cap S = \{F(P_3), L(P_3), v_{\pi+1}, v_{\pi+2}, \ldots, v_r\},$$

если либо $\pi \geqq 3$ либо $\pi = 0$, $P_2 \cap S = V_1^* \cap S$. Если же $\pi = 2$, $P_2 \cap S \neq V_1^* \cap S$, то будем предполагать, что

$$V(P_3) \cap S = \{F(P_3), L(P_3), v_4, v_5, \ldots, v_r\}.$$

При $|V(P_3) \cap S| = 2$ будем предполагать, что $r = \pi$.

Если $V(H_1) \setminus V(P_3) \neq \emptyset$, то существует S-допустимая пара цепей $(x, y) = (R_3, R_4)$, удовлетворяющая следующим условиям:

и1. $L(x) = F(P_3)$, $F(y) = L(P_3)$, $F(x) \in V(H_1)$, $L(y) \in V(H_1)$,

$$[F(P_3), u] \uparrow P_3 \subseteqq x, \quad [L(P_3), v] \uparrow P_3 \subseteqq y, \quad V_2^* \subseteqq V(H_1) \cup S,$$

$$V_2^* \cap V_1^* \subseteqq \{F(P_2), L(P_2)\}, \quad |V_2^*| > |V(P_3)|,$$

где $V_2^* = V(x) \cup V(y)$, $uv \in X(P_3)$.

и2. Для любой S-допустимой пары цепей $(x, y) = (l_1, l_2)$, удовлетворяющей условию и1, имеет место

$$\left|(V(l_1) \cup V(l_2)) \cap S\right| \leqq |V_2^* \cap S|.$$

и3. Для любой S-допустимой пары цепей $(x, y) = (l_1, l_2)$, удовлетворяющей условиям и1, и2, имеет место

$$|V(l_1) \cup V(l_2)| \leqq |V_2^*|.$$

Без потери общности можем предполагать, что

$$V_2^* \cap S = \{F(P_2), L(P_2), v_{\pi+1}, v_{\pi+2}, \ldots, v_\omega\},$$

если либо $\pi \geqq 3$, либо $\pi = 2$, $P_2 \cap S = V_1^* \cap S$. Если же $\pi = 2$, $P_2 \cap S \neq V_1^* \cap S$, то будем предполагать, что

$$V_2^* \cap S = \{F(P_2), L(P_2), v_4, v_5, \ldots, v_\omega\}.$$

При $|V_2^* \cap S| = 2$ будем предполагать, что $\omega = \pi$.

По лемме 3 для S-допустимой пары цепей R_3, R_4 существует цепь P_4, удовлетворяющая хотя бы одному из условий (1), (2), (8).

Если $V(H_1) \setminus V(P_3) = \emptyset$, то будем предполагать, что $P_4 = P_3$, $V_2^* = V(P_3)$, $\omega = r$.

Допустим, что $V(H_1) \setminus V(P_3) \neq \emptyset$, т.е. $P_4 \neq P_3$. Если $\pi \geqq 3$, то

$$d_3 + d_4 \geqq 2\delta - 2(\pi - 2) - (k - \omega) \geqq$$

$$\geqq (2\delta - 2k + 1) + (\omega - \pi + 2) + 1 + (k - \pi) \geqq |V_2^*| + 1,$$

где

$$d_3 = |N(F(R_3)) \cap V_2^*|, \quad d_4 = |N(L(R_4)) \cap V_2^*|.$$

Если же $\pi = 2$, то

$$d_3 + d_4 \geqq 2\delta - (k - \omega) - 2 \geqq (2\delta - 2k + 1) + (\omega - 1) + 1 + (k - 3) \geqq |V_2^*| + 1.$$

По лемме 1 для S-допустимой пары цепей R_3, R_4 существует цепь P_4, удовлетворяющая условию (1), что противоречит определению цепи P_3. Таким образом, $V(H_1) \setminus V(P_3) = \emptyset$, $\omega = r$.

Если $V(H_2) \setminus V(P_1) = \emptyset$, то аналогично случаю 1.2.1 можно показать, что $P_2 = P_1$, $r = k$, т.е. $h = v(G)$.

Пусть $V(H_2) \setminus V(P_1) \neq \emptyset$, т.е. $P_2 \neq P_1$. Так как цепь P_2 не удовлетворяет условию (1), то по лемме 2 цепь P_2 удовлетворяет хотя бы одному из условий (2), (3), (4), т.е.

$$|V(P_2)| \geqq d_1 + d_2 - 1 \geqq 2\delta - (k - \pi) - 1,$$

откуда

$$|V(H_1)| \geqq |V(H_2)| \geqq |V(P_2)| - |V(P_2) \cap S| \geqq 2\delta - k - 1,$$

что приводит нас к случаю 1.

Случай 3. $|V(H_1)| = 2\delta - k - 2 - \beta$, где $\beta \in \overline{0, k-4}$.

Пусть простая цепь $x = P_1$ удовлетворяет следующим условиям:

к1. $F(x) \in S$, $L(x) \in S$, $V(x) \subseteqq V(H_2) \cup S$.

к2. $|V(x) \cap S| \leqq \beta + 3$.

к3. Для любой цепи $x = P$, удовлетворяющей условиям к1, к2, имеет место

$$|V(P) \cap S| \leqq |V(P_1) \cap S|.$$

к4. Для любой цепи $x = P$, удовлетворяющей условиям к1, к2, к3, имеет место

$$|V(P)| \leqq |V(P_1)|.$$

Если $V(H_2)\setminus V(P_1)\neq\emptyset$, то существует S-допустимая пара цепей $(x, y)=(R_1, R_2)$, удовлетворяющая следующим условиям:

л1. $$L(x)=F(P_1),\ F(y)=L(P_1),\ F(x)\in V(H_2),\ L(y)\in V(H_2),$$

$$[F(P_1), u]\uparrow P_1 \subseteqq x, \quad [L(P_1), v]\uparrow P_1 \subseteqq y,$$

$$|V_1^*| > |V(P_1)|, \quad V_1^* \subseteqq V(H_2)\cup S,$$

где $V_1^*=V(x)\cup V(y)$, $uv\in X(P_1)$.

л2. $|V_1^*\cap S|\leqq\beta+3$.

л3. Для любой S-допустимой пары цепей $(x, y)=(l_1, l_2)$, удовлетворяющей условиям л1, л2, имеет место

$$\left|(V(l_1)\cup V(l_2))\cap S\right| \leqq |V_1^*\cap S|.$$

л4. Для любой S-допустимой пары цепей $(x, y)=(l_1, l_2)$, удовлетворяющей условиям л1, л2, л3, имеет место

$$|V(l_1)\cup V(l_2)| \leqq |V_1^*|.$$

По лемме 3 для S-допустимой пары цепей $(x, y)=(R_1, R_2)$ существует цепь P_2, удовлетворяющая хотя бы одному из условий (1), (2), (8).

Если же $V(H_2)\setminus V(P_1)=\emptyset$, то будем предполагать, что $P_1=P_2$, $V_1^*=V(P_1)$.

Пусть простая цепь $x=P_3$ удовлетворяет следующим условиям:

м1. $$F(x)=F(P_2),\ L(x)=L(P_2),\ V(x)\subseteqq V(H_1)\cup S,$$

$$V(x)\cap V_1^*=\{F(P_2), L(P_2)\}.$$

м2. Для любой цепи $x=P$, удовлетворяющей условию м1, имеет место

$$|V(P)\cap S| \leqq |V(P_3)\cap S|.$$

м3. Для любой цепи $x=P$, удовлетворяющей условиям м1, м2, имеет место

$$|\{i/V(P)\cap V(G_i)\neq\emptyset\}| \leqq |\{i/V(P_3)\cap V(G_i)\neq\emptyset\}|.$$

м4. Для любой цепи $x=P$, удовлетворяющей условиям м1, м2, м3, имеет место

$$|V(P)| \leqq |V(P_3)|.$$

Допустим, что для некоторого числа $i\in\overline{1, t-1}$ имеет место

(12) $$V(P_3)\cap V(G_i)\neq\emptyset, \quad V(G_i)\setminus V(P_1)\neq\emptyset.$$

Тогда существует S-допустимая пара цепей $(x, y)=(R_3, R_4)$, удовлетворяющая следующим условиям:

н1. $$L(x) = F(P_3), \quad F(y) = L(P_3), \quad F(x)\in V(H_1), \quad L(y)\in V(H_1),$$
$$[F(P_3), u]\uparrow P_3 \subseteqq x, \quad [L(P_3), v]\uparrow P_3 \subseteqq y, \quad V_2^* \subseteqq V(H_1)\cup S,$$
$$V_2^*\cap V_1^* = \{F(P_2), L(P_2)\}, \quad |V_2^*| > |V(P_3)|,$$

где $V_2^*=V(x)\cup V(y)$, $uv\in X(P_3)$.

н2. Для любой S-допустимой пары цепей $(x, y)=(l_1, l_2)$, удовлетворяющей условию н1, имеет место
$$\left|(V(l_1)\cup V(l_2))\cap S\right| \leqq |V_2^*\cap S|.$$

н3. Для любой S-допустимой пары цепей $(x, y)=(l_1, l_2)$, удовлетворяющей условиям н1, н2, имеет место
$$\left|\{i/V(G_i)\cap(V(l_1)\cup V(l_2)) \neq \emptyset\}\right| \leqq |\{i/V(G_i)\cap V_2^* \neq \emptyset\}|.$$

н4. Для любой S-допустимой пары цепей $(x, y)=(l_1, l_2)$, удовлетворяющей условиям н1, н2, н3, имеет место
$$|V(l_1)\cup V(l_2)| \leqq |V_2^*|.$$

Без потери общности можем предполагать, что
$$V_2^*\cap S = \{v_1, v_2, \ldots, v_\Delta\}, \quad V_1^*\cap S = \{v_1, v_2, v_{\Delta+1}, v_{\Delta+2}, \ldots, v_\pi\}.$$
Пусть
$$d_1 = \left|N(F(R_3))\cap V_2^*\right|, \quad d_2 = \left|N(L(R_4))\cap V_2^*\right|.$$

Если $\Delta \geqq k-\beta-1$, то
$$d_1+d_2 \geqq 2\delta-2(k-\Delta) \geqq [(2\delta-k-2-\beta)+\Delta+1]+\Delta-k+\beta+1 \geqq |V_2^*|+1.$$

Если же $\Delta \leqq k-\beta-2$, то для любой вершины $\xi\in\{v_{\pi+1}, v_{\pi+2}, \ldots, v_k\}$ имеет место
$$F(R_3)\xi \notin X(G) \quad \text{или} \quad L(R_4)\xi \notin X(G),$$
откуда, ввиду $|V_1^*\cap S|=\pi-\Delta-2\leqq\beta+3$, имеем
$$d_1+d_2 \geqq 2\delta-2(\pi-\Delta)-(k-\pi) \geqq$$
$$\geqq ((2\delta-k-2-\beta)+\Delta+1)+\Delta-\pi+1+\beta \geqq |V_2^*|+1.$$

Следовательно, по лемме 1 для S-допустимой пары цепей R_3, R_4 существует цепь P_4, удовлетворяющая условию (1), что противоречит определению цепи P_3. Таким образом условие (12) не выполняется.

Если $t-1 \geqq k-\beta-1$, то
$$(\delta-k+1)(k-\beta-1) \leqq |V(H_1)| = 2\delta-k-2-\beta,$$
откуда
$$1 \leqq \delta-k+1 \leqq \frac{k-\beta-4}{k-\beta-3} < 1,$$
что невозможно.

Пусть $t-1 \leqq k-\beta-2$. Тогда из $|V_1^* \cap S| \leqq \beta+3$ следует

$$\Delta \geqq t, \quad |\{i/V(G_i) \cap V_2^* \neq \emptyset\}| = t-1,$$

откуда, ввиду того, что условие (12) не выполняется, имеем $V(H_1) \setminus V(P_3) = \emptyset$, т.е.

$$|V(P_3)| = 2\delta-k-2-\beta+\Delta.$$

Пусть тройка $(x, y, z) = (P_4, R_5, R_6)$ удовлетворяет следующим условиям:

п1. $F(x)=v_1$, $L(x)=v_2$, $V(x) \subseteqq V(H_2) \cup S$.

п2. Пара цепей y, z является S-допустимой и существует только тогда, когда $V(H_2) \setminus V(x) \neq \emptyset$.

п3. $L(y)=v_1$, $F(z)=v_2$, $F(y) \in V(H_2)$, $L(z) \in V(H_2)$,

$$[v_1, u]\uparrow x \subseteqq y, \ [v_2, v]\uparrow x \subseteqq z, \ |V_3^*| > |V(x)|,$$

где $V_3^* = V(y) \cup V(z)$, $uv \in X(x)$.

п4. Если $V(H_2) \setminus V(x) = \emptyset$, то $V_3^* = V(x)$.

п5. $V_3^* \subseteqq V(H_2) \cup S$, $V_3^* \cap V(P_3) = \{v_1, v_2\}$, $V_1^* \cap S \subseteqq V_3^* \cap S$.

п6. Для любой тройки $(x, y, z) = (P, l_1, l_2)$, удовлетворяющей условиям п1—п5, имеет место

$$\left|(V(l_1) \cup V(l_2)) \cap S\right| \leqq |V_3^* \cap S|.$$

п7. Для любой тройки $(x, y, z) = (P, l_1, l_2)$, удовлетворяющей условиям п1—п6, имеет место

$$|V(l_1) \cup V(l_2)| \leqq |V_3^*|.$$

Без потери общности можем предполагать, что

$$V_3^* \cap S = \{v_1, v_2, v_{\Delta+1}, v_{\Delta+2}, \ldots, v_r\}, \quad r \geqq \pi.$$

По лемме 3 для R_5, R_6 существует цепь P_5, удовлетворяющая хотя бы одному из условий (1), (2), (8). Пусть

$$d_5 = \left|N(F(R_5)) \cap V_3^*\right|, \quad d_6 = \left|N(L(R_6)) \cap V_3^*\right|.$$

Если $V(H_2) \setminus V(P_4) = \emptyset$, то будем предполагать, что $P_5 = P_4$.

При $V(H_2) \setminus V(P_4) = \emptyset$ легко убедится, что $r=k$, откуда $h = V(G)$.

Пусть $V(H_2) \setminus V(P_4) \neq \emptyset$, т.е. $V_3^* \neq V(P_4)$. Цепь P_5 не удовлетворяет условию (1), так как иначе P_5 будет противоречить определению цепи P_4. Тогда по лемме 3 цепь P_5 удовлетворяет хотя бы одному из условий (2), (8), т.е.

$$|V(P_5)| \geqq d_5+d_6-1 \quad \text{или} \quad |V(P_5)| \geqq \min(d_5, d_6)+r-\Delta+1.$$

По определению цепи P_5 имеем

$$\min(d_5, d_6) \geqq \delta-(\Delta-2)-(k-r) \geqq \delta-k+r-\Delta+2,$$

откуда

$$d_5+d_6-1 \geqq 2\min(d_5, d_6)-1 \geqq \min(d_5, d_6)+r-\Delta+1.$$

Таким образом

$$|V(P_5)| \geqq \min(d_5, d_6)+r-\Delta+1.$$

Случай 3.2.1. $\Delta \leqq r-\beta-1$.

Если $r=k$, то из $\min(d_5, d_6) \geqq \delta-\Delta+2$ следует

$$h \geqq |V(P_3)|+|V(P_5)|-2 \geqq (2\delta-k-2-\beta+\Delta)+(\delta-\Delta+2+r-\Delta+1)-2 \geqq$$

$$\geqq (3\delta-k)+r-\Delta-\beta-1 \geqq 3\delta-k.$$

Пусть $r<k$ и пусть T_5—некоторая компонента связности графа $\langle v_{r+1}, v_{r+2}, \ldots, v_k\rangle$. По определению цепей P_3, P_4, R_5, R_6 имеем

$$|M_5 \cap (V(P_3) \cup V_3^*)| \leqq r < k,$$

где $M_5 = N(V(T_5))$. Так как $|M_5| \geqq k$ и $|V(H_1) \setminus V(P_3)| = \emptyset$, то

$$M_5 \cap (V(H_2) \setminus V_3^*) \neq \emptyset.$$

Тогда для любой вершины $\xi \in \{v_{r+1}, v_{r+2}, \ldots, v_k\}$ имеет место

$$F(R_5)\xi \notin X(G), \quad L(R_6)\xi \notin X(G),$$

так как иначе цепь $R_5 \cup R_6 \cup F(R_5)\xi \cup L(R_6)\xi$ будет противоречить определению цепи P_4. Отсюда $\min(d_5, d_6) \geqq \delta-\Delta+2$ и, следовательно,

$$h \geqq (2\delta-k-2-\beta+\Delta)+(\delta-\Delta+2+r-\Delta+1)-2 \geqq 3\delta-k.$$

Случай 3.2.2. $\Delta \geqq r-\beta$.

Так как $|V_3^* \cap S| = r-\Delta+2 \leqq \beta+2$, то из условий к2, п1—п7 следует $V_3^* = V_1^*$ и

$$(N(F(R_5)) \cup N(L(R_6))) \cap \{v_{r+1}, v_{r+2}, \ldots, v_k\} = \emptyset,$$

$$(N(F(R_5)) \cap N(L(R_6))) \cap \{v_3, v_4, \ldots, v_\Delta\} = \emptyset.$$

Тогда $d_5+d_6 \geqq 2\delta-\Delta+2$. Учитывая неравенство

$$|V(H_2)| \leqq |V(H_1)| = 2\delta-k-2-\beta,$$

получим

$$|V_3^*| \leqq |V(H_2)|+|V_3^* \cap S| \leqq (2\delta-k-2-\beta)+(r-\Delta+2) \leqq$$

$$\leqq (2\delta-\Delta+2)-(k-r)-\beta-2 \leqq d_5+d_6-2.$$

Следовательно, по лемме 1 для R_5, R_6 существует цепь P, удовлетворяющая условию (1), что противоречит определению цепи P_4. Теорема доказана.

Доказательство теоремы 1 непосредственно следует из доказательства теоремы 2. Заметим однако, что теорема 1 допускает простое доказательство, присланное мне редакцией журнала *J. Combinatorial Theory*[1].

---

[1] Häggkvist, R. and Nicoghossian, G. G., A remark on hamiltonian cycles, *J. Combinatorial Theory* **30** (1981), 118—120.

Следствие 1. *Если 2-связный граф G удовлетворяет условию* $\delta(G) \geqq (v(G)+\alpha(G)-1)/3$, *то* $h(G)=v(G)$.

Доказательство. Если $\alpha(G) \leqq k(G)$, то из теоремы Хватала—Эрдёша [4] следует $h(G)=v(G)$. Пусть $\alpha(G) \geqq k(G)+1$. Тогда

$$\delta(G) \geqq (v(G)+\alpha(G)-1)/3 \geqq (v(G)+k(G))/3$$

и по теореме 1 получим $h(G)=v(G)$.

Следствие 2. *Если 3-связный граф G удовлетворяет условию* $\delta(G) \leqq (v(G)+\alpha(G)-1)/3$, *то* $h(G) \geqq 3\delta(G)-\alpha(G)+1$.

Доказательство. Если $\delta(G) \geqq (v(G)+k(G))/3$ или $\alpha(G) \leqq k(G)$, то по теореме 1 и по теореме Хватала—Эрдёша $h(G)=v(G) \geqq 3\delta(G)-\alpha(G)+1$. Пусть $\delta(G) < (v(G)+k(G))/3$ и $\alpha(G) \geqq k(G)+1$. Тогда по теореме 2 получим

$$h(G) \geqq 3\delta(G)-k(G) \geqq 3\delta(G)-\alpha(G)+1.$$

Рассмотрим графы $G_1=k_{\delta-k}$, $G_2=\bar{k}_{\delta+\varrho}$, $G_3=k_k$, $G_4=k_{\delta-k+1}$, попарно без общих вершин, где $\varrho \geqq 0$. Граф, полученный из $(G_1+G_2) \cup (G_3+G_4)$ посредством добавления всевозможных ребер xy, где $x \in V(G_2)$, $y \in V(G_3)$, обозначим через G_0. Так как удаление δ-вершинного подграфа $G_1 \cup G_3$ из G_0 порождает $\delta+\varrho+1$ компонент связности, то $h(G_0)=3\delta(G_0)-k(G_0)$. Пример графа G_0 показывает, что утверждения теорем 1 и 2 не улучшаемы.

ЛИТЕРАТУРА

[1] Harary, F., *Graph theory*, Addison—Wesley Publishing Co., Reading, Mass.—Menlo Park, Calif. —London, 1969. *MR* **41** # 1566.
[2] Dirac, G. A., Some theorems on abstract graphs, *Proc. London Math. Soc.* (3) **2** (1952), 69—81. *MR* **13**—856.
[3] Nash-Williams, C. St. J. A., Edge-disjoint Hamiltonian circuits in graphs with vertices of large valency, *Studies in Pure Mathematics*, ed. by L. Mirsky, Academic Press, London, 1971. *MR* **44** # 1594.
[4] Chvatal, V. and Erdős, P., A note on Hamiltonian circuits, *Discrete Math.* **2** (1972), 111—113. *MR* **45** # 6654.

(Поступила 22 декабря 1980)

ВЫЧИСЛИТЕЛЬНЫЙ ЦЕНТР
АКАДЕМИИ НАУК АРМЯНСКОЙ ССР И
ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
УЛ. П. СЕВАКА 1
SU—375044 ЕРЕВАН 44
SOVIET UNION

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 283—290

MINIMAL GRAPHS OF DIAMETER TWO AND GIVEN MAXIMAL DEGREE

I. VRŤO and Š. ZNÁM

Denote $F_2(n, k)$ the minimal number of edges of a graph with n vertices, diameter 2 and maximal degree k. This notion was introduced in [1], where the values of $F_2(n, k)$ were determined for $k=n-1, \ldots, n-4$. In [2] is proved that $F_2(n, k)=2n-4$ for $\frac{2n-2}{3} \leqq k \leqq n-5$. In [2] the values for $k<\frac{2n-2}{3}$ are also studied, however, the results given there (without proofs) are not precise. In this paper we prove the following

THEOREM. *If* $n>241$ *then*

$$F_2(n, k)=\begin{cases} 3n-k-9 & \text{if } \frac{2n-6}{3} \leqq k \leqq \frac{2n-3}{3} & (1)\\ 3n-k-8 & \text{if } k=\frac{2n-7}{3} & (2)\\ 3n-k-7 & \text{if } \frac{3n-3}{5} \leqq k \leqq \frac{2n-8}{3}. & (3)\end{cases}$$

PROOF. The extremal graph for the case (1) consists of:

1. vertices a, b, c, d, e, f and edges $ae, ac, bd, ce, df, ef, bf$;
2. a set of $n-k-4$ vertices of degree 3 adjacent to a, b and d;
3. a set of $\left\lfloor\frac{k-2}{2}\right\rfloor$ vertices of degree 2 adjacent to b and c;
4. a set of $\left\lceil\frac{k-2}{2}\right\rceil$ vertices of degree 2 adjacent to c and d. We can easily show that this is a graph with n vertices, diameter 2, maximal degree k, and it has $3n-k-9$ edges.

The extremal graph for the case (2) consists of:

1. vertices a, b, c, d, e, f, g, h and edges $ae, ad, ah, bc, bf, bg, cf, cg, de, dh, ef, fg, gh$;
2. a set of $n-k-5$ vertices of degree 3 adjacent to a, b and d;
3. a set of $\frac{k-3}{2}$ vertices of degree 2 adjacent to a and c;

1980 *Mathematics Subject Classification.* Primary 05C35.
Key words and phrases. Diameter, degree, maximal degree, minimal number of edges.

4. a set of $\frac{k-3}{2}$ vertices of degree 2 adjacent to c and d.

The extremal graph for the case (3) consists of

1. vertices a, b, c, d and edges ab, bc, bd, cd;
2. a set of $2k-n$ vertices of degree 2 adjacent to a and b;
3. a set of $n-k-3$ vertices of degree 3 adjacent to b, c and d;
4. a set of $\left\lceil \frac{n-k-1}{2} \right\rceil$ vertices of degree 2 adjacent to a and c;
5. a set of $\left\lfloor \frac{n-k-1}{2} \right\rfloor$ vertices of degree 2 adjacent to a and d.

In the following let G denote a graph of diameter 2 with $n>241$ vertices and maximal degree k, where

$$\frac{3n-3}{5} \leqq k \leqq \frac{2n-3}{3}. \tag{4}$$

Let V be the set of vertices of G, $e(G)$ the number of edges, and $\delta(G)$ the minimal degree of G. If x is a vertex of G, $O(x)$ denotes the set of all vertices adjacent to x and $d(x)$ the degree of x in G.

To prove the theorem, we shall need eight lemmas.

LEMMA 1. *For every vertex x in G we have*

$$\sum_{v \in O(x)} d(v) \geqq n-1.$$

PROOF. Every vertex can be reached from x by a path of length $\leqq 2$.

COROLLARY 1. *Every vertex of degree 3 in G is adjacent to at least two vertices of degrees at least* $\frac{n-k-1}{2}$.

COROLLARY 2. *Every vertex of degree 4 is adjacent to at least two vertices of degrees at least* $\frac{n-k-1}{3}$.

LEMMA 2. *If* $\delta(G) \geqq 3$ *then* $e(G) \geqq 3n-k-7$.

PROOF. Put $P=\{v \in V, d(v)=3\}$, $Q=\{v \in V, d(v)=4\}$,

$$R=\left\{v \in V,\, 5 \leqq d(v) < \frac{n-k-1}{2}\right\},\quad S=\left\{v \in V,\, d(v) \geqq \frac{n-k-1}{2}\right\}.$$

Let $|P|=p$, $|Q|=q$, $|R|=r$, $|S|=s$.

If $s \geqq 11$, then we have (see (4)): $2e(G) \geqq 3(n-s)+s\,\frac{n-k-1}{2} \geqq 6n-2k-14$ and the assertion follows.

If $s \leqq 10$, then by Lemma 1 every vertex of degree 4 is adjacent to at least one vertex of S. Thus (according to Corollary 1) we get:

$$2e(G) \geqq 3p+4q+5r+2p+q = 5(n-s) \geqq 5n-50 \geqq 6n-2k-14$$

and the proof of Lemma 2 is finished.

Because of (4) no vertex of degree 1 exists in G, hence in the following we shall suppose

(5) $$\delta(G) = 2.$$

Denote $M = \{v \in V,\ d(v) = 2\}$.

LEMMA 3.

(6) $$\bigcap_{v \in M} O(v) \neq \emptyset.$$

PROOF. We shall proceed indirectly, by supposing that $\bigcap_{v \in M} O(v)$ is empty. Then there exist 3 vertices $a, b, c \in V$ so that every couple of them represents the neighbourhood of some vertex of degree 2 in G. Thus every vertex distinct from a, b, c has to be adjacent to at least two of them. Further, there exist at least two edges with both endpoints in $\{a, b, c\}$. Hence the sum of the degrees of a, b, c is at least $2(n-3)+4 = 2n-2$, which contradicts (4).

Denote $O(M) = \bigcup_{v \in M} O(v)$, $|O(M)| = t$.

LEMMA 4. *If* $t \geqq 18$, *then* $e(G) \geqq 3n-k-7$.

PROOF. According to (6) there exists a vertex $a \in O(M)$ adjacent to all vertices of degree 2. A vertex non-adjacent to a has to be adjacent to at least $t-2$ vertices of $O(M)$. Hence the sum of all degrees in G is at least

$$(t-2)(n-k-1) \geqq 16(n-k-1) \geqq 6n-2k-14,$$

and the assertion follows.

LEMMA 5. *If* $5 \leqq t < 18$, *then* $e(G) \geqq 3n-k-7$.

PROOF. Let a be adjacent to all vertices of M. There exist at least $n-d(a)-t$ vertices non-adjacent to a and not belonging to $O(M)$. Each of these vertices has to be adjacent to all remaining vertices of $O(M)$. Thus the sum of all degrees in G is at least $2(t-1)(n-d(a)-t)+4(d(a)-t)$, which is for $t = 5, 6, \ldots, 17$ more than $6n-2k-14$. The proof of the lemma is finished.

LEMMA 6. *If* $t=4$, *then* $e(G) \geqq 3n-k-7$.

PROOF. Suppose a is adjacent to all vertices of M. Let b_1, b_2, b_3 denote the remaining vertices of $O(M)$. Let G_0 be the subgraph of G induced by $O(M)$. Let X be the set of all vertices of $V-(M \cup O(M))$ adjacent to a, and Y be the set of all such vertices non-adjacent to a (note that a vertex of Y is adjacent to all three vertices b_i). Using the notations $|X| = x$, $|Y| = y$, $|M| = m$, we get

(7) $$x+m+y = n-4, \qquad x+m \leqq k-d_0(a),$$

where $d_0(a)$ is the degree of a in G_0.

Now we distinguish two cases:

a) If $d_0(a)=0$, then every vertex $v\in Y$ is of degree at least 4 (exists a path of length $\leqq 2$ from v to a there). Hence the sum of all degrees in G is at least $4m+4x+7y$, which is, by (7), at least $7n-3k-28\geqq 6n-2k-14$.

b) Suppose $d_0(a)\geqq 1$. Let c_i be a vertex of degree 2 adjacent to a and b_i, $i=1, 2, 3$. There exists a path of length $\leqq 2$ from c_i to each of the remaining b_j's thus if b_i is not adjacent to a, it must be adjacent to the remaining b_j's. Hence

$$d_0(a)+e(G_0)\geqq 5. \tag{8}$$

Therefore we have (see (7), (8)): $e(G)\geqq 2m+3y+2x+e(G_0)\geqq 3n-12-k+d_0(a)+e(G_0)\geqq 3n-k-7$. The proof of Lemma 6 is finished.

LEMMA 7. *Let* $|O(M)|=3$, *then*

$$e(G)\geqq\begin{cases}3n-k-9 & \text{if } \dfrac{2n-6}{3}\leqq k\leqq\dfrac{2n-3}{3}\\[2mm] 3n-k-8 & \text{if } k=\dfrac{2n-7}{3}\\[2mm] 3n-k-7 & \text{if } \dfrac{3n-3}{5}\leqq k\leqq\dfrac{2n-8}{3}.\end{cases}$$

PROOF. Suppose b is a vertex adjacent to all vertices of degree 2. Let a, c be the remaining vertices of $O(M)$. Let G_0 be the subgraph of G induced by the vertices a, b, c. Denote by Z the set of vertices in G which are of degree 3 and are not adjacent to b. Obviously, every vertex of Z is adjacent to a and c.

We shall distinguish the cases $Z\neq\emptyset$ and $Z=\emptyset$, respectively.

Case 1. If two vertices v_1, v_2 of Z are adjacent, then deleting the edge v_1v_2 from G, we get a graph G' of diameter 2 and maximal degree k in which (6) does not hold — a contradiction with Lemma 3. Denote by L the set of all vertices from $V-(M\cup\{a, b, c\})$ adjacent to some vertex of Z. Obviously, $L\neq\emptyset$, $L\cap Z=\emptyset$. Further, let P be the set of all vertices of L adjacent to b but non-adjacent to a and c; let Q be the set of vertices of L adjacent to b and to at least one of a and c; let $R=L-(P\cup Q)$. Denote by X_1 the set of all vertices of $W=V-(Z\cup L\cup M\cup\{a, b, c\})$ adjacent to b but not adjacent to a and c, by X_2 the vertices of W adjacent to b and to at least one of a and c, and put $Y=W-(X_1\cup X_2)$. Let $|P|=p$, $|Q|=q$, $|R|=r$, $|Z|=z$, $|X_1|=x_1$, $|X_2|=x_2$, $|Y|=y$. We have

$$m+x_1+x_2+y+z=n-3-p-q-r, \tag{9}$$

$$m+x_1+x_2\leqq k-d_0(b)-p-q, \tag{10}$$

$$m+2y+2z+x_2+2r+q\leqq 2k-d_0(a)-d_0(c). \tag{11}$$

Note that every vertex of Y is of degree at least 4 and has to be adjacent to a and c. Further, every vertex of X_1 is adjacent to all vertices of L. Thus we have

$$e(G)\geqq 2m+(1+p+q+r)x_1+2x_2+3y+3z+e(G_1), \tag{12}$$

where G_1 denotes the subgraph of G induced by the set $L\cup\{a, b, c\}$.

Similarly, if $|L|=1$, then every element of X_1 has to be adjacent to the single element of L and we get

(13) $$e(G) \geqq 2m+\frac{5}{2}x_1+2x_2+3y+3z+e(G_1).$$

Using (9) and (10), from (12) we obtain:

(14) $$e(G) \geqq 3n-k-9-2p-3r-2q+(p+q+r-1)x_1+e(G_1).$$

Every vertex of P is adjacent to all vertices of L, hence we have at least $\binom{p}{2}$ edges with both endpoints in L. Therefore

(15) $$e(G_1) \geqq p+2q+2r+\binom{p}{2}+e(G_0).$$

Now, owing to (14) and (15), we have

(16) $$e(G) \geqq 3n-k-9+(p+q+r-1)x_1+\binom{p}{2}-p-r+e(G_0).$$

Further we shall need the following inequality arising from (9), (10) and (11):

(17) $$2n-6-p-x_1+2e(G_0) \leqq 3k.$$

We shall consider (16) and (17) with respect to various values of parameters. First by, notice that $e(G_0)\geqq 1$ in all cases.

(a) Suppose $x_1=0$;

(i) $p=0$, $r=0$, then from (16) and (17) we get

$$k \geqq \frac{2n-4}{3}, \quad e(G) \geqq 3n-k-8;$$

(ii) $p=0, r\geqq 1$, then $e(G_0)\geqq 2$ (a vertex of Z adjacent to R can be reached from b only if one of the edges ab, bc exists) and hence $k\geqq\frac{2n-2}{3}$;

(iii) $p=1$ or 2, then $k\geqq\frac{2n-6}{3}$, $e(G)\geqq 3n-k-9$;

(iv) $p=3$, then $k\geqq\frac{2n-7}{3}$, $e(G)\geqq 3n-k-8$;

(v) $p\geqq 4$, then $e(G)\geqq 3n-k-6$.

(b) Suppose $x_1\geqq 1$ and $p+q+r\geqq 2$;

(vi) $x_1=1$, $p\leqq 1$, then $k\geqq\frac{2n-6}{3}$, $e(G)\geqq 3n-k-9$;

(vii) $x_1=1$, $p=2$, then $k\geqq\frac{2n-7}{3}$, $e(G)\geqq 3n-k-8$;

(viii) $x_1=1$, $p\geqq 3$, then $e(G)\geqq 3n-k-7$;

(ix) $x_1\geqq 2$, $r\geqq 1$, then $e(G_0)\geqq 2$ (see (ii)), $e(G)\geqq 3n-k-7$;

(x) $x_1\geqq 2$, $q\geqq 1$, then $e(G)\geqq 3n-k-7$;

(xi) $x_1\geqq 2$, $p\geqq 2$, then $e(G)\geqq 3n-k-7$.

(c) Suppose $x_1 \geqq 1$ and $p+q+r=1$. In the same way as above, from (13) we get

$$e(G) \geqq 3n-k-10+q+e(G_0)+\frac{x_1}{2}. \tag{18}$$

We shall consider (17) and (18):

(xii) $q=1$, then $e(G) \geqq 3n-k-7$;

(xiii) $r=1$, then $e(G_0) \geqq 2$ (see (ii)) and $e(G) \geqq 3n-k-7$;

(xiv) $p=1$, $x_1 \leqq 2$, then $k \geqq \frac{2n-7}{3}$, $e(G) \geqq 3n-k-8$;

(xv) $p=1$, $x_1 \geqq 3$, then $e(G) \geqq 3n-k-7$.

Case 2. We have to distinguish two cases again:

Case 2a. Suppose that there exists a vertex e adjacent to b but not to a and c, and $d(e)=3$. Let e be adjacent to the further vertices f, h. Let J (I) be the set of vertices from $V-(M \cup \{a, b, c, e, f, h\})$ adjacent (non-adjacent) to b. Let $|J|=j$, $|I|=i$. Obviously, every vertex of I is of degree at least 4 and is adjacent to a, c and to at least one of vertices f, h. Further, we have $m+j+i=n-6$ and $m+j \leqq k$, thus

$$e(G) \geqq 2m+2j+\frac{7}{2}i \geqq \frac{7}{2}(n-6)-\frac{3}{2}k \geqq 3n-k-7.$$

Case 2b. Suppose that Case 2a does not hold. Denote by X (Y) the set of all vertices of $V-(M \cup \{a, b, c\})$ adjacent (non-adjacent) to b. Every vertex of Y is of degree at least 4 and is adjacent to a and c. Because Case 2a does not hold, every vertex of degree 3 is adjacent to b and to at least one of a, c. Obviously, $m+x+y=n-3$ and $m+x \leqq k$, thus $e(G) \geqq 2m+\frac{5}{2}x+3y+e(G_0) \geqq 3n-k-9+\frac{x}{2}+e(G_0)$.

If $x \geqq 1$, then $e(G) \geqq 3n-k-7$. If $x=0$, then every vertex of $V-\{a, b, c\}$ is adjacent to at least two vertices from $\{a, b, c\}$. Further, in this case we have $e(G_0) \geqq 2$ and we get a contradiction with (4). The proof of Lemma 7 is finished.

LEMMA 8. *Let* $|O(M)|=2$. *Then* $e(G) \geqq 3n-k-7$.

PROOF. Let $O(M)=\{a, b\}$. The sum of the degrees of a and b is at least $n+m-2$, thus

$$m \leqq 2k-n+2. \tag{19}$$

Split the set $U=V-(M \cup \{a, b\})$ into the following subsets:

$W=\left\{v \in U,\ d(v) \geqq \frac{n-k-1}{3}\right\}$;

$P_1=\{v \in U,\ d(v)=3$ and v is adjacent to a vertex of degree $3\}$;

$P_2=\{v \in U, d(v)=3,\ v \notin P_1,\ v$ is adjacent to exactly two vertices of $W \cup \{a, b\}\}$;

$P_3=\{v \in U,\ d(v)=3,\ v$ is adjacent to 3 vertices of $W \cup \{a, b\}\}$;

$Q=\{v \in U,\ d(v)=4\}$;

$R=\left\{v \in U,\ 5 \leqq d(v) < \frac{n-k-1}{3}\right\}$.

Put $|W|=w$, $P=P_1\cup P_2\cup P_3$, $|P_i|=p_i$, $|P|=p$, $|Q|=q$, $|R|=r$. If $w\geqq 7$, then

$$2e(G)\geqq 2m+(n+m-2)+3(n-m-2-w)+w\frac{n-k-1}{3}=$$

$$=6n-2k-10-3w+(w-6)\frac{n-k-1}{3}\geqq 6n-2k-14$$

and the assertion follows.

In the following we shall suppose $w\leqq 6$. Obviously,

$$\text{(20)}\qquad \begin{aligned} p+q+r+w+m &= n-2,\\ p+q+r+w+2m &\leqq 2k. \end{aligned}$$

Denote by $D(R)$ the sum of the degrees of all vertices of R.

By the Corollaries to Lemma 1, every vertex of P and Q has to be adjacent to at least two vertices from the set $W\cup\{a,b\}$. Thus by (20) and (21) we have:

$$\text{(22)}\qquad \begin{aligned} e(G) &\geqq 2m+\frac{5}{2}p+3q+\frac{D(R)-r}{2}+r+\frac{p_3}{2}+w\geqq\\ &\geqq 3n-k-6+\frac{p_3+q+D(R)-4r-3w}{2}. \end{aligned}$$

If $w\leqq 1$, then $e(G)\geqq 3n-k-7$ follows from (22). Hence suppose $w=2,\ldots,6$. If $p_3+q+D(R)-4r\geqq 15$, we get $e(G)\geqq 3n-k-7$ again. Thus consider

$$\text{(23)}\qquad p_3+q+D(R)-4r\leqq 14.$$

Obviously, $D(R)\geqq 5r$, and so from (23) we get

$$\text{(24)}\qquad p_3+q+r\leqq 14.$$

Every vertex of P_2 is adjacent to at least one vertex from R or Q, hence we have

$$\text{(25)}\qquad p_2\leqq 2q+D(R)-r.$$

Finally, from (23), (24) and (25) we get

$$p_2+p_3+q+r\leqq 70,$$

therefore

$$\text{(26)}\qquad p_1+m=n-2-p_2-p_3-q-r-w\geqq n-78>k.$$

Hence in P_1 there exists some vertex non-adjacent to a and also some vertex non-adjacent to b. Further: (19) and (26) imply that $p_1>2$ and so we can choose two non-adjacent vertices $u, v\in P_1$ as follows: u is adjacent to a but not to b, v is adjacent to b but not to a, and both u, v are adjacent to the same vertex c from W. Let u be adjacent to a further vertex f of degree 3, and v to the vertex g, $d(g)=3$.

Every vertex of the set $V-(\{a,b,c,f,g\}\cup(W-\{c\}))$ is adjacent to at least two vertices from $\{a,b,c\}$; the vertex c is adjacent to at least one of a, b; every vertex of the set $(W-\{c\})\cup\{f,g\}$ is adjacent to at least one of a, b. Hence the sum of the degrees of vertices a, b and c is at least

$$2(n-5-w+1)+2+(w-1)+2=2n-5-w.$$

Therefore we have

$$2e(G) \geqq 2m+3(n-m-w-2)+(2n-5-w)+(w-1)\frac{n-k-1}{3} \geqq$$

$$\geqq 6n-2k-37+\frac{n-k-1}{3} \quad \text{(see (19))}$$

and the assertion of Lemma 8 follows.

The assertion of our Theorem follows from Lemmas 2, 4, 5, 6, 7 and 8. The proof of the Theorem is finished.

REMARK. We have found the following further results:

$$F_2(n, k) \leqq 5n-4k-15 \quad \text{if} \quad \frac{3n-8}{5} \leqq k < \frac{3n-3}{5};$$

$$F_2(n, k) \leqq 5n-4k-12 \quad \text{if} \quad \frac{5n-7}{9} \leqq k < \frac{3n-8}{5};$$

$$F_2(n, k) \leqq 4n-2k-14 \quad \text{if} \quad \frac{n+1}{2} \leqq k \leqq \frac{5n-8}{9};$$

$$F_2(n, k) \leqq 3n-15 \quad \text{if} \quad \frac{n-5}{2} \leqq k \leqq \frac{n+1}{2}.$$

Acknowledgement. The authors are grateful to F. Guldan, M. Sudolský and P. Tomasta for their help in constructing the extremal graphs.

REFERENCES

[1] ERDŐS, P. and RÉNYI, A., On a problem in the theory of graphs, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 7/A (1962), 623—641. *MR* 33#1246.
[2] ERDŐS, P., RÉNYI, A. and SÓS, V. T., On a problem of graph theory, *Studia Sci. Math. Hungar.* 1 (1966), 215—235. *MR* 36#6310.

(Received January 6, 1981)

INSTITUTE OF TECHNICAL CYBERNETICS
SLOVAK ACADEMY OF SCIENCES
DÚBRAVSKÁ CESTA 9
ČS—842 37 BRATISLAVA
CZECHOSLOVAKIA

KATEDRA ALGEBRY A TEÓRIE ČÍSEL
MATEMATICKO-FYZIKÁLNEJ FAKULTY
UNIVERSITY KOMENSKÉHO
MLYNSKÁ DOLINA
ČS—842 15 BRATISLAVA

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 291—302

FINDING TWO CONSECUTIVE ELEMENTS

Á. VARECZA

It is a well-known result that $n-2+\lceil\log_2 n\rceil$ pairwise comparisons are needed to determine the first and second elements among n different numbers. In this paper we are considering the problem of finding some ordered consecutive pair. It is proved that the number of necessary pairwise comparisons is the same as in the above weaker case.

It is a frequent occurance that certain element or elements (for example the median or the first two ones) must be sorted out from ordered data stored in the memory unit of a computer.

We can generally apply only one operation: choosing two elements and comparing them. The mathematical problem is obvious: to work out a method, which leads to the target as quickly as possible, in a certain sense. This "sense" for example may be the number of necessary steps. In this matter there only a few exact results have been attained so far. Mostly only lower and upper bounds have been proved. We know for example that if H is a totally ordered finite set with n elements, then $n-1$ pairwise comparisons must be performed for selecting the maximum element (and, of course, for selecting the minimum one).

The problem of selecting the first two elements, was raised by Steinhaus in a mathematical seminar in 1930, non-complete solutions were published by Schreier [7], and Slupecki [6]. The first complete proof is due to Kislicyn [3], who proved that $n-2+\lceil\log_2 n\rceil$ comparisons must be performed for selecting the first two elements in the worst case. Knuth [4] gave a considerably shorter proof for this lower bound. Lately, there has been discovered an accurate, but rather complicated formula for the required number of comparisons for finding the first three elements by Kirkpatrick [2]. If we want to sort out the first and last element of set H simultaneously, then we must make at least $n+\left\lceil\frac{n}{2}\right\rceil-2$ comparisons (J. Pohl [5] and Á. Varecza [8]).

G.O.H. Katona suggested the following sharpening of the Steinhaus problem. How many pairwise comparisons do we need if we want to find some ordered pair of consecutive elements? He conjectured that there is no faster way than to select the first two elements. In this paper we prove this conjecture. We will now introduce the necessary concepts and notations.

First define the concept of strategy suitable for determining some two ordered neighbouring elements of the set H. We start by comparing two elements, for example,

1980 *Mathematics Subject Classification.* Primary 05A05; Secondary 05A15.
Key words and phrases. Finite ordered set, searching, lower bound of the length of strategy.

c and d, where $c, d \in H$ and denote the pair (c, d) by S_0. Let ε_1 (the answer) be 1, if $c > d$, and let ε_1 be 0 if $c < d$. Depending on ε_1 we select a pair $S_1(\varepsilon_1)$, say $e(\varepsilon_1)$ and $f(\varepsilon_1)$. Define ε_2 to be 1 if $e(\varepsilon_1) > f(\varepsilon_1)$ and to be 0 otherwise. Continuing the same way, after a 0,1 sequence $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_i$ of answers, the pair

$$S_i(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_i) \tag{1}$$

is given as the next pair for comparison, with the restriction that if $S_i(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_i)$ is defined, then $S_{i-1}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{i-1})$ is defined, also. ε_{i+1} is 1 or 0 according to whether the first or the second term of the pair (1) is larger. A set of questions given in this way will be called a strategy suitable for determining two neighbouring elements of the set H (simply strategy further on), if for all sequences,

$$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_l$$

when

$$\begin{aligned} &S_{l-1}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{l-1}) && \text{is determined, but} \\ &S_l(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_l) && \text{is not,} \end{aligned} \tag{2}$$

then the answers

(3) $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_l$ (together with the questions $S_0, S_1(\varepsilon_1), \ldots, S_{l-1}(\varepsilon_1, \ldots, \varepsilon_{l-1})$) determine two neighbouring elements of the set H (maybe the first two elements).

If $T_i(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{i+1})$ denotes the inequality made from the pair $S_i(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_i)$ on the basis of the answer ε_{i+1}, then condition (3) can also be formulated in the way that two neighbouring elements of H are derived from the inequalities

$$T_0(\varepsilon_1), T_1(\varepsilon_1, \varepsilon_2), \ldots, T_{l-1}(\varepsilon_{1l}, \varepsilon_2, \ldots, \varepsilon_l). \tag{4}$$

If the sequence $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_l$ satisfies conditions (2), (3) we say: the strategy is finished. Such a sequence is also called a path of $\mathscr{S}$. The maximal length of sequences $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_l$ finishing the strategy is called the length of the strategy.

Let $\mathscr{S}$ denote a strategy and $L(\mathscr{S})$ its length. The situation after answering the question $S_{i-1}(\varepsilon_1, \ldots, \varepsilon_{i-1})$ will be called the state $(\varepsilon_1, \ldots, \varepsilon_i)$ of the strategy $\mathscr{S}$.

We need the following lemma formulated first in Knuth D. E. (1975) (pp. 211—212). We repeat here the details.

LEMMA 1. *Let $\mathscr{S}_1$ be a strategy suitable for selecting the first element of the n element set H. Suppose that the path $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_l$ satisfies the following condition. For any i ($1 \leq i < l$), the answer ε_{i+1} for the comparison $S_i(\varepsilon_1, \ldots, \varepsilon_i) = (c, d)$ is 1 if d appears as smaller but c does not, or neither of them appears as smaller, but c appears more times as larger than d does in the inequalities:*

$$T_0(\varepsilon_1), T_1(\varepsilon_1, \varepsilon_2), \ldots, T_{i-1}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_i) \tag{5}$$

and $\varepsilon_{i+1} = 0$ if the above cases occur when c and d are interchanged.

Then the first element occurs in the inequalities

$$T_0(\varepsilon_1), T_1(\varepsilon_1, \varepsilon_2), \ldots, T_{l-1}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_l) \tag{6}$$

at least $[\log_2 n]$ times.

Proof of Lemma 1. Consider those inequalities from (5) in which some element (in this order) occurs as smaller the first time. Denote by $\mathscr{E}_i$ the set of these inequalities.

We shall prove the following statement about $\mathscr{E}_i$:

If an element a does not act as smaller in (5) *but it does p times as larger, then $b<a$ or $b=a$ can be claimed for elements b, at most 2^p on the basis of $\mathscr{E}_i$.*

Proof. We shall prove the statement by induction on p. For $p=0$ the statement is trivial. Assume that $p>1$ and the statement is true for $p-1$. We shall prove it for p. Let $T_{j-1}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_j)$ be the last inequality in (5) that contains a. Let, for example, $a>c$. We shall see that it is impossible to deduce more $b>a$ or $b=a$ on the basis of $\mathscr{E}_i$, than on the basis of $\mathscr{E}_j$. Namely, let $b>a$ be deducible in $\mathscr{E}_i$ that is, there exist b's such that

$$b = b_0 < b_1 < \ldots < b_k = a$$

and let h be that smallest integer for which the inequalities

$$b_h < b_{h+1} < \ldots < b_k = a$$

still occur in $\mathscr{E}_j$. Problems arise only in case $h>0$. In this case the relation $b_{h-1}<b_h$ must be in $\mathscr{E}_i - \mathscr{E}_j$.

If $h=k$ then this contradicts the assumption that $T_{j-1}(\varepsilon_1, \ldots, \varepsilon_j)$ is the last occurrence of a. Conversely, if $0<h<k$, then b_{h-1} occurs as smaller before $b_{h-1}<b_h$ (which is the first inequality where b_{h-1} occurs as smaller) and this contradicts the conditions on the considered paths. Therefore $h=0$.

If $T_{j-1}(\varepsilon_1, \ldots, \varepsilon_j)$ is not in $\mathscr{E}_j$, that is, it is not the first occurence of c as smaller, then we can immediately take $\mathscr{E}_{j-1}$ instead of $\mathscr{E}_i$ and a appears only $p-1$ times in this $\mathscr{E}_i$. Thus — because of the inductive hypotheses — at most 2^{p-1} pieces of $b<a$ or $b=a$ can be deduced.

Consequently, let us suppose $T_{j-1}(\varepsilon_1, \ldots, \varepsilon_j)$ to be in $\mathscr{E}_i$, that is, c has not been smaller so far. In this case — because of the construction of the ε's — a has been greater at most $p-1$ times before.

How many $b<a$ can be deduced in $\mathscr{E}_j$ by using $c<a$. As many as the number of the inequalities $b<c$ or $b=c$ deducible in $\mathscr{E}_{j-1}$. The number of these relations is at most 2^{p-1} because of the inductive hypotheses.

On the other hand without using $c<a$ there are at most 2^{p-1} deducible $b<a$ or $b=a$ in $\mathscr{E}_{j-1}$. That is, the number of deducible relations $b<a$, or $b=a$ is at most 2^p. Thus the number of relations $b<a$ or $b=a$ deducible in $\mathscr{E}_i$ is at most 2^p. The statement is proved.

Let us return to the proof of the lemma.

To all elements b — except one, say a — a sequence of inequalities from (6) can be found from which we can get $b<a$ by successive application. It is true in $\mathscr{E}_i$, too, because b appears also in $\mathscr{E}_i$ as smaller: $b<b_1$, but b_1 does, too: $b_1<b_2$ and so on.

In this way we reach such an element b_k, which does not act as smaller in $\mathscr{E}_i$. But as a consequence, it does not act in (6) either, consequently $b_k=a$. So if a occurs in (6) p times, then $2^p \geq n$, that is $p \geq \lceil \log_2 n \rceil$.

This completes the proof of the lemma.

Let $\mathcal{S}$ be a strategy suitable for determining some pair of ordered neighbouring elements in H. Let $(\varepsilon_1, \ldots, \varepsilon_i)$ be an arbitrary state of $\mathcal{S}$. We will form a partition of H into the subsets $A_1(\varepsilon_1, \ldots, \varepsilon_i)$, $A_2(\varepsilon_1, \ldots, \varepsilon_i)$, $B_1(\varepsilon_1, \ldots, \varepsilon_i)$, $B_2(\varepsilon_1, \ldots, \varepsilon_i)$, $C(\varepsilon_1, \ldots, \varepsilon_i)$. The last one will be the set of elements not occurring in any comparisons until this state. The definition of the sets A_1, A_2, B_1, B_2 is more complicated and will follow later. If it does not give rise to a misunderstanding we shall omit the ε's, e.g. $A_1(\varepsilon_1, \ldots, \varepsilon_i) = A_1^i$. We say that a comparison (b, c) in the state $(\varepsilon_1, \ldots, \varepsilon_i)$ is of type (e.g.) (A_2, A_1) if

$$S_i(\varepsilon_1, \ldots, \varepsilon_i) = (b, c) \tag{7}$$

and $b \in A_2(\varepsilon_1, \ldots, \varepsilon_i)$, $c \in A_1(\varepsilon_1, \ldots, \varepsilon_i)$. The system of inequalities

$$T_0(\varepsilon_1),\ T_1(\varepsilon_1, \varepsilon_2),\ \ldots,\ T_{i-1}(\varepsilon_1, \ldots, \varepsilon_i) \tag{8}$$

is denoted by $\mathcal{E}_i \cdot a \in K(\varepsilon_1, \ldots, \varepsilon_i) = K^i$ $(N(\varepsilon_1, \ldots, \varepsilon_i) = N^i)$ if $a \in H$ and a occurs in $\mathcal{E}_i$ first as smaller (larger). We also say that a is of type $K(N)$.

We will determine an "adversary" path (more exactly, our definition will be ambiguous, it may lead to different paths). The adversary path and the partition $A_1^i \cup A_2^i \cup B_1^i \cup B_2^i \cup C^i$ will be determined simultaneously and recursively. Suppose that $\varepsilon_1, \ldots, \varepsilon_i$ and the partition $A_1^i \cup A_2^i \cup B_1^i \cup B_2^i \cup C^i$ are determined. Depending on the classes of b and c, resp. (see (7)), we give the result ε_{i+1} and the new classes of b and c. The other elements of A_1^i and A_1^{i+1} (A_2^i and A_2^{i+1}; B_1^i and B_1^{i+1}; B_2^i and B_2^{i+1}; C^i and C^{i+1}) are equal. We do not list the cases following by interchanging the roles of b and c, A_1^i and B_1^i, A_2^i and B_2^i, K^i and N^i, resp.

1. $b, c \in C^i$: $\varepsilon_{i+1} = 1$, $b \in A_1^{i+1}$, $c \in B_1^{i+1}$;

2. $b \in A_1^i(B_1^i)$, $c \in C^i$: $\varepsilon_{i+1} = 1\,(0)$, $c \in B_1^{i+1}(A_1^{i+1})$

3. $b \in A_1^i$, $c \in A_2^i \cup K^i$:

   or

   $b \in A_2^i$, $c \in K^i$:

   or

   $b \in B_2^i$, $c \in B_1^i$: $\varepsilon_{i+1} = 1$.

4. $b, c \in A_1^i(B_1^i)$: $\varepsilon_{i+1} = 1\,(0)$, $h \in A_2^{i+1}(B_2^{i+1})$ if b occurs as larger (smaller) in $\mathcal{E}_i$ more times than c. $\varepsilon_{i+1} = 0\,(1)$, $g \in A_2^{i+1}(B_2^{i+1})$ if b occurs as larger (smaller) as many times as c does and either b occurs in $\mathcal{E}_i$ with at least one element of $B_1^i(A_1^i)$ and c does not or c occurs in $\mathcal{E}_i$ only with elements of $A_2^i(B_2^i)$. In other cases the value of ε_{i+1} is arbitrary; the smaller (larger) element $\in A_2^{i+1}(B_2^{i+1})$.

5. $b, c \in A_2^i(B_2^i)$: ε_{i+1} arbitrary (but not contradicting the previous ε's).

6. $b \in A_2^i(B_2^i), c \in C^i$:

   (i) $\varepsilon_{i+1} = 0\,(1)$, $c \in A_1^{i+1}(B_1^{i+1})$ if no element of $A_2^i(B_2^i)$ occurred with an element of C^i until this state.

   (ii) We distinguish 3 cases if there is exactly one comparison (say $(e > f)$, $e \in A_2^i$ $(e \in B_2^i)$) in $\mathcal{E}_i$ comparing an element of $A_2^i(B_2^i)$ with an element of C^i.

   I. If $b \neq e, f$ then $\varepsilon_{i+1} = 0\,(1)$, $c \in A_1^{i+1}(B_1^{i+1})$.

II. If $b=e$ then $\varepsilon_{i+1}=1(0)$, $c\in B_1^{i+1}(A_1^{i+1})$.

III. In the case $b=f$ we distinguish 2 subcases:

$\varepsilon_{i+1}=1(0)$, $c\in B_1^{i+1}(A_1^{i+1})$ if f does not occur with an element of C^i in $\mathscr{E}_i$. Moreover, in what follows we do not consider this case to be a comparison of an element of $A_2^i(B_2^i)$ with an element of C^i. We proceed in the next steps as if there was no comparison of this type.

$\varepsilon_{i+1}=0(1)$, $c\in A_1^{i+1}(B_1^{i+1})$ if f did occur with an element of C^i. Moreover, we proceed in the next steps as if there was exactly one comparison of an element of $A_2^i(B_2^i)$ with an element of C^i, namely $b<c$.

(iii) If there were exactly two pairs comparing an element of $A_2^i(B_2^i)$ with an element of C^i then

$\varepsilon_{i+1}=1(0)$, $c\in B_1^{i+1}(A_1^{i+1})$ for case I and

$\varepsilon_{i+1}=0(1)$, $c\in A_1^{i+1}(B_1^{i+1})$ for case II.

(iv) If there were at least 3 pairs comparing an element of $A_2^i(B_2^i)$ with an element of C^i (case I or II took place) then we distinguished several cases: $\varepsilon_{i+1}=1(0)$, $b\in B_1^{i+1}(A_1^{i+1})$ if b first occurred with an element of $A_2^{i+1}(B_2^{i+1})$ but did not occur with an element of C^i. In what follows, this case is not considered a "comparison of b with an element of C^i". Consider now the case when b occurred with an element of C^i after getting into $A_2^i(B_2^i)$. Let k and m denote the number of these occurrences where b is smaller (larger) and larger (smaller), resp. If $k<m$ then $\varepsilon_{i+1}=0(1)$, $c\in A_1^{i+1}(B_1^{i+1})$.

If $k\geqq m$ then $\varepsilon_{i+1}=1(0)$, $c\in B_1^{i+1}(A_1^{i+1})$. In the case when b did not occur with an element of C^i after getting into $A_2^i(B_2^i)$ let r and s denote the number of the elements of $A_2^i(B_2^i)$ occurring exactly once with an element of C^i after getting into $A_2^i(B_2^i)$ and being larger (smaller) and smaller (larger) in this comparison, resp. If $s<r$ then $\varepsilon_{i+1}=0(1)$, $c\in A_1^{i+1}(B_1^{i+1})$ and if $s\geqq r$ then $\varepsilon_{i+1}=1(0)$, $c\in B_1^{i+1}(A_1^{i+1})$.

In this way we defined a path (or branch) $\varepsilon_1, \ldots, \varepsilon_l$ of the strategy $\mathscr{S}$. It will be denoted by P. The length $|P|$ of P is l. We shall prove

$$l = |P| \geqq n-2+[\log_2 n].$$

For the sake of easier formulation we introduce the concept of the graph-realization. Correspond the elements of the set H to the vertices of graph G. Let a comparison be an edge of G between the corresponding vertices and let the answer determine the orientation of this edge in the following way: if we compare two elements, say c and d, in some state of G and the result of the comparison is $c>d$ then we direct the edge from c to d, and conversely when $c<d$ we direct the edge from d to c. In the state $(\varepsilon_1, \ldots, \varepsilon_i)$ let $G(\varepsilon_1, \ldots, \varepsilon_i)$ denote the graph derived in this way. H is totally ordered, so oriented circle cannot be produced in $G(\varepsilon_1, \ldots, \varepsilon_i)$. By the above correspondence we uniquely associate an oriented graph to all states of $\mathscr{S}$. It follows from the correspondence that in an arbitrary state $(\varepsilon_1, \ldots, \varepsilon_i)$ of $\mathscr{S}$ the relation $e>f$ is realized if and only if an oriented path leads in $G(\varepsilon_1, \ldots, \varepsilon_i)$ from e to f. If $\mathscr{S}$ is ended and (say) x and y $(x>y)$ are the two consecutive elements, then this means in the corresponding graph that:

(1) there is a directed path from x to y and for all elements $(\neq x, y)$ of H one of the following two conditions is fulfilled:

(2a) an oriented path leads to both x and y.

(2b) an oriented path leads to it from both x and y.

In the following we shall refer to this graph, when necessary.

Now let us suppose that the strategy $\mathcal{S}$ is ended for the sequence

$$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_l \tag{9}$$

along the path determined above. Let

$$T_0(\varepsilon_1), T_1(\varepsilon_1, \varepsilon_2), \ldots, T_{l-1}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_l) \tag{10}$$

be the system of inequalities according to the answer (9).

THEOREM 1. *If $\mathcal{S}$ is a strategy suitable for selecting two arbitrary ordered consecutive elements of the set H, then*

$$L(\mathcal{S}) \geqq n-2+\lceil \log_2 n \rceil$$

and this is the best possible bound.

PROOF. Fix an adversary path (9) defined earlier. Suppose that we get x and y $(x>y)$ as the two consecutive elements along this path. Four cases are possible:
1. x is of type N, y is of type K;
2. x is of type K, y is of type N;
3. both x and y are of type N;
4. both x and y are of type K.

Let us study the cases separately.

Case 1. In the state $(\varepsilon_1, \ldots, \varepsilon_l)$ all elements are either of type K or of type N. Thus using the notation

$$|A_1(\varepsilon_1, \ldots, \varepsilon_l) \cup A_2(\varepsilon_1, \ldots, \varepsilon_l)| = i$$
$$|B_1(\varepsilon_1, \ldots, \varepsilon_l) \cup B_2(\varepsilon_1, \ldots, \varepsilon_l)| = j,$$

$i+j=n$ holds.

We prove that the number of inequalities of type (A, B) in $\mathcal{E}_l$ is at least $\left\lceil \frac{1}{2}\left\lceil \frac{n}{2} \right\rceil \right\rceil$. Let the elements of $N^l = A_1^l \cup A_2^l$ not compared with elements of $K^l = B_1^l \cup B_2^l$ in $\mathcal{E}_l$ be $a_1, \ldots, a_k$. These elements occurred first with an element of A_2^l. Let $b \in A_2^l$ and let $a_{i_1}, \ldots, a_{i_j}$ be the ones from the set $\{a_1, \ldots, a_k\}$ occurring the first time with b. By the definition of P, b occurs in at least j comparisons with the elements of $B_1^l \cup B_2^l$. Hence it follows that the number of inequalities between the elements of $A_1^l \cup A_2^l$ and $B_1^l \cup B_2^l$ is at least $\left\lceil \frac{|A_1^l \cup A_2^l|}{2} \right\rceil$. By symmetry, we may use the lower estimate $\left\lceil \frac{|B_1^l \cup B_2^l|}{2} \right\rceil$, too. That is, the number of inequalities of type (A, B) in $\mathcal{E}_l$ is lower bounded with

$$\max\left(\left\lceil \frac{|A_1^l \cup A_2^l|}{2} \right\rceil, \left\lceil \frac{|B_1^l \cup B_2^l|}{2} \right\rceil\right).$$

As

$$|A_1^l \cup A_2^l| + |B_1^l \cup B_2^l| = n,$$

one of the terms is $\geqq \left\lceil \frac{n}{2} \right\rceil$. We obtained the desired lower bound: $\left\lceil \frac{1}{2}\left\lceil \frac{n}{2} \right\rceil \right\rceil$.

Choose an element $u \neq x$ of type N.

We know that there are directed paths in $G(\varepsilon_1, \ldots, \varepsilon_l)$ either to or from u from both x and y. In the first case there is an edge along the path form y to u which leads from an element of type K to an element of type N. This is contradiction. Consequently, we may suppose that the second case holds: there is a directed path from u to x. Moreover, there is no element of type K along this path, because it would give a pair (an element of type $N <$ an element of type K).

We may conclude that from any element u of type N, there is a directed path to x using elements of type N only. In other words, the graph induced by the elements of type N is connected. That is, it has at least $i-1$ edges.

We can see in the same way, that the graph induced by the elements of type K is also connected. That is, it contains at least $j-1$ edges.

In this way we divide the set of vertices of $G(\varepsilon_1, \ldots, \varepsilon_l)$ into two classes (elements of type N and K, resp.). The number of edges between the two classes is at least $\left[\frac{1}{2}\left[\frac{n}{2}\right]\right]$, and in each of the classes there are at least $i-1$ and $j-1$ edges resp. That is, the graph has at least $\left[\frac{1}{2}\left[\frac{n}{2}\right]\right]+i-1+j-1=n+\left[\frac{1}{2}\left[\frac{n}{2}\right]\right]-2$ edges. In other words, the number of inequalities in (10) is at least

$$n+\left[\frac{1}{2}\left[\frac{n}{2}\right]\right]-2 \geqq n+\lceil\log_2 n\rceil-2 \qquad (n > 12).$$

One can easily check that $l \geqq n+\lceil\log_2 n\rceil-2$ when $n \leqq 12$, too.

Case 2. In $G(\varepsilon_1, \ldots, \varepsilon_l)$ there is a directed path from x to y. If x is of type K and y is of type N, then there is an edge along this path connecting a vertex of type K with a vertex of type N. But this is impossible by the choice of ε's.

Case 3. If there is a directed path between x and y in $G(\varepsilon_1, \ldots, \varepsilon_l)$ then by symmetry we may suppose that it leads from x to y. Let us divide the set $A(\varepsilon_1, \varepsilon_2, \ldots$ $\ldots, \varepsilon_l)-\{x, y\}$ into subclasses:

$$A(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_l)-\{x, y\} = A_1' \cup A_2' \cup A_2''$$

where $A_1'=A_1(\varepsilon_1, \ldots, \varepsilon_l)-\{x, y\}$ and $A_2' \cup A_2''$ is a partition of $A_2(\varepsilon_1, \ldots, \varepsilon_l)$ (A_2' (A_2'') consist of elements having a directed path to (from) both x and y).

Using the notation

$$|A_1'| = i_1, \quad |A_2'| = i_2, \quad |A_2''| = i_3, \quad |B| = j,$$

we have

$$i_1+i_2+i_3+j = n-2. \tag{11}$$

We now prove a lemma about A_1'.

LEMMA 2. *If* $|A_1'|\neq 0$, *then it contains an element* m *occurring in at least*

$$\left\lceil \log_2 \left(n - \left\lfloor \frac{n-|A_1^l|}{2} \right\rfloor \right) \right\rceil - \lceil \log_2 |A_1^l| \rceil$$

inequalities in (10), *and always as greater.*

PROOF. The strategy $\mathcal{S}$ will be completed to be a strategy determining the largest element. For any finishing state of $\mathcal{S}$ let us consider the elements which still can be maximum and organize an "elimination tournament" for them. (That is, the winners are always compared.) If the number of possible maximum elements is k in a certain finishing state of $\mathcal{S}$, then the winner will occur exactly $\lceil \log_2 k \rceil$ times in these comparisons. If we complete all of the finishing states of $\mathcal{S}$, we obtain a strategy $\mathcal{S}_1$ which determines the maximum element. In the state $(\varepsilon_1, \ldots, \varepsilon_l)$ all the elements of $B \cup A_2$ have been occurred as smaller. If $A_1' \neq \emptyset$, x, y occurred, too. Thus the number of possible maximum elements is i_1.

Suppose that $b \in A_2^l \cup B_2^l$ occurs with at least one element of C. Let these inequalities in $\mathcal{E}_l$ be $b < a_1, \ldots, b < a_r, b > c_1, \ldots, b > c_s$, in this order. We omit the inequalities $a_k > b$, $b > c_k$ $(1 \leq j \leq \min(r, s))$ from $\mathcal{E}_l$ and replace the earlier one by $a_k > c_k$. We carry out these changes for all $b \in A_2^l \cup B_2^l$. Denote by $\mathcal{E}_l^*$ the new sequence of inequalities and by S^* $(\subset H)$ the set of elements not occurring in $\mathcal{E}_l^*$.

Let us now apply Lemma 1 for this $\mathcal{S}_1$ and any path going through the earlier specified state $(\varepsilon_1, \ldots, \varepsilon_l)$. It is easy to see that this path satisfies the conditions of Lemma 1 on the set $H - S^*$. It follows that the maximum element occurs in at least $\lceil \log_2 (n - |S^*|) \rceil$ comparisons. On the other hand we know that it occurs exactly $\lceil \log_2 i_1 \rceil$ times after the state $(\varepsilon_1, \ldots, \varepsilon_l)$, that is, it occurs at least $\lceil \log_2 (n - |S^*|) \rceil - \lceil \log_2 |A_1'| \rceil$ times in (10).

We prove that

$$|S^*| \leq \left\lfloor \frac{n-|A_1^l|}{2} \right\rfloor.$$

Let $\varphi\colon S^* \to H - A_1^l$ be a mapping determined in the following way: $\varphi(a) = b$ $(a \in S^*, b \notin S^*, b \notin A_1^l)$ if the element a occurred first time with b. It is easy to see that φ is injective. Hence we obtain

$$2|S^*| \leq |H| - |A_1^l|$$

and the above inequality. The lemma is proved.

Let us return to the proof of Case 3 of the theorem. Suppose first that $A_1' \neq \emptyset$. Choose an element m according to Lemma 2. There is a directed path in $G(\varepsilon_1, \ldots, \varepsilon_l)$ from each element u of $A_1' \cup A_2'$ to both x and y.

We prove that the number of inequalities in $\mathcal{E}_l$ containing an element of $A_1^l \cup A_2' \cup \{x\} - \{m\}$ as larger is at least $2|A_1^l| + |A_2'| - 1$. Let $a \in A_2' \cup \{x\}$ and suppose that a occurs as smaller with k such elements of $A_1^l \cup A_2'$ which occur exactly once as larger in $\mathcal{E}_l$. Then, by Section 6 of the definition of the path, a occurs in at least $k+1$ inequalities as larger (it must occur with an element of $A_2' \cup \{x, y\}$, too). Hence the statement follows. On the other hand, there is a path to each element of $A_2'' \cup B$ from y. y cannot lie along these paths, because it would for a directed cycle with the path from x to y. An element of $A_1' \cup A_2'$ cannot lie along these paths by the same reasoning. That is, the graph induced by $A_2'' \cup B \cup \{y\}$ is connected, consequently it contains at least i_3+j edges.

Summing up the numbers of these 3 classes of edges (inequalities) we obtain

$$l \geqq \left\lceil \log_2 \left(n - \left\lfloor \frac{n-|A_1^l|}{2} \right\rfloor \right) \right\rceil - \lceil \log_2 |A_1^l| \rceil + 2|A_1^l| + |A_2'| - 1 + |B \cup A_2''| =$$

$$= n-3+|A_1^l| + \left\lceil \log_2 \left(n - \left\lfloor \frac{n-|A_1^l|}{2} \right\rfloor \right) \right\rceil - \lceil \log_2 |A_1^l| \rceil \geqq n-2+\lceil \log_2 n \rceil$$

if $|A_1^l|>3$.

Suppose now that $|A_1^l|=3$ and let $A_1^l=\{m, a_1, a_2\}$. Assume first that one of a_1 and a_2, say a_1, occurs only once in $\mathscr{E}_l$: $a_1>b_1$, $b_1 \in A_2' \cup \{x\}$. It is easy to see that b_1 occurs as larger with at least one element of each of B and $A_2' \cup \{x, y\}$. We distinguish several cases.

If b_1 did not occur with an element of C (except a_1) then there is no $c \in S^*$ with $\varphi(c)=b_1$. The element m occurs at least

$$\text{(12)} \qquad \left\lceil \log_2 \left(n-1-\left\lfloor \frac{n-1-3}{2} \right\rfloor \right) \right\rceil - 1 = \left\lceil \log_2 \left(\left\lceil \frac{n}{2} \right\rceil + 1 \right) \right\rceil - 1$$

times in $\mathscr{E}_l^* - \{a_1>b_1\}$. (The set $H-\{a_1\}$ is used.) Consequently, the same is true for $\mathscr{E}_l$ as m, a_2 and b_1 are not in the domain of φ. Now we prove that the graph induced by $H-\{m\}$ contains at least two different non-directed circles and as a consequence at least n edges. b_1 occurs in $\mathscr{E}_l$ with an element c of B. There is a directed path from x to c and from b_1 to x. Denote this circle by K_1. If a_2 also occurs exactly once in $\mathscr{E}_l$(as larger), say $a_2>b_2$ then $b_1 \neq b_2$ by our assumptions. b_2 occurs with an element of B, this ensures another circle $K_2 \neq K_1$. If a_2 occurs twice as larger, say $a_2>c_1$, $a_2>c_2$ then one of c_1 and c_2 is an element of $A_2' \cup \{x\}$. Say e.g. $c_1 \in A_2' \cup \{x\}$. If $c_2 \in A_2' \cup \{x\}$ also holds then there is a path from both c_1, c_2 to x. This gives a new circle. On the other hand, if $c_2 \notin A_2' \cup \{x\}$ we can find another circle similarly. Therefore we have two circles in all cases, so the number of inequalities within $H-\{m\}$ is at least n. The desired inequality

$$\text{(13)} \qquad l \geqq n + \lceil \log_2 n \rceil - 2$$

follows by (12).

If b_1 occurs with an element of C different from a_1 then b_1 occurs (as larger) with at least two different elements of B, by Section 6 of the definition of the path. We distinguish subcases:

Suppose that a_2 occurs either twice or only once (say $a_2>b_2$) but $b_1 \neq b_2$. Then x occurs at least once, b_2 does at least twice as larger. The elements of $A_1^l \cup A_2' \cup \{x\} - \{m\}$ take part in at least $2|A_1^l|+|A_2'|$ inequalities as larger. We can verify (13) in the same way as above.

Suppose now that a_2 occurs exactly once (say $a_2>b_2$) and $b_1=b_2=b$. Then $b<a_1$ and $b<a_2$ are both in $\mathscr{E}_l$, consequently, by Section 6 of the definition of the path, b must occur at least two elements of $B_1^l \cup B_2^l = B^l$.

If b occurs with at least 3 elements of B^l then (13) follows easily.

If b occurs with exactly two elements of B^l and both a_1 and a_2 only once then a_1 and a_2 are not in the domain of φ. It is easy to see that m occurs in at least

$[\log_2 n]-2$ times in the inequalities $\mathscr{E}_l^* - \{a_1 > b\}$. Hence (13) easily follows. So we may suppose that both a_1 and a_2 occur at least twice. If one of them occurs at least 3 times (13) is obvious. We have to consider only the case when both a_1 and a_2 occur exactly twice. Suppose first that $A_2^l \neq \emptyset$. Let $b \in A_2^l$. If b occurs at least twice in $\mathscr{E}_l$ then (13) follows like above. If b occurs only once as larger (say $b > c \in A_2^l \cup \{x\}$) then c occurs at least twice as larger and (13) can be verified again. Consider now the case when $A_2^l = \emptyset$. In this case $x < a_1, x < a_2, a_1 > c_1, a_2 > c_2$ are all in $\mathscr{E}_l$ with some c_1, c_2. If x occurs at least twice as larger then (13) follows like before. Suppose that x occurs as larger exactly once. Then x cannot be compared with an element of S^*. We may suppose that a_1 occurs in $\mathscr{E}_l$ before a_2. The inequality $a_1 > c_1$ comes before $a_1 > x$ and $a_2 > c_2$ precedes $a_2 > x$; otherwise $a_1 < x$ and $a_2 < x$ would follow, by Section 6. Omit a_2 and all inequalities containing it from $\mathscr{E}_l^*$. Replace $a_2 > c_2$ by $a_1 > c_2$ if a_2 is not matched by φ. If, on the contrary, it is matched with some d then $a_2 > c_2$ is replaced by $a_1 > d$. The new system of inequalities given on $H - \{a_2\}$ contains at least

$$\left[\log_2\left(n-1-\left[\frac{n-1-3}{2}\right]\right)\right]-1 = \left[\log_2\left[\frac{n+2}{2}\right]\right]-1 \geqq [\log_2 n]-2$$

inequalities continaing m because there is no $c \in S^*$ satisfying $\varphi(c) = m$, x or a_1. Hence (13) easily follows as the number of inequalities within the set $H - \{m\}$ is at least n.

We can start to examine the case $|A_1^l| = 2$. Here $A_1^l = \{m, a\}$. It is easy to see that if $A_2^l \neq \emptyset$ then the elements of $A_1^l \cup A_2^l \cup \{x\} - \{m\}$ occur in at least $2|A_1^l| + |A_2^l|$ inequalities of $\mathscr{E}_l$. Hence (13) follows. The same can be said if either both a and x occur at least twice as larger in $\mathscr{E}_l$ or one of them occurs once, the other one occurs at least 3 times as larger.

Suppose that a occurs once and x occurs twice as larger in $\mathscr{E}_l$. The element m occurs at least

$$\left[\log_2\left(n-1-\left[\frac{n-1-2}{2}\right]\right)\right] = \left[\log_2\left[\frac{n+1}{2}\right]\right]$$

times in $\mathscr{E}_l^* - \{a > x\}$ ($\varphi(c) \neq m, x$ for $c \in S^*$).

$$l \geqq n-1+\left[\log_2\left[\frac{n+1}{2}\right]\right]$$

implies (13).

If a occurs twice ($a > x$, $a > b$) x only once ($x > y$) as larger in $\mathscr{E}_l$ then let $y < e$ be y's first occurrence as smaller. $y < e$ is before $y < x$ in $\mathscr{E}_l$ because x occurs only once as larger. It is easy to see that $e = m$. Omit the inequalities $a > x$, $a > b$ from $\mathscr{E}_l$. If a is matched with c by φ then replace $a > c$ by $x > c$ in $\mathscr{E}_l^*$. If there is no such c then $a > b$ is replaced by $x > b$. In this system of inequalities m occurs at least

$$\left[\log_2\left(n-1-\left[\frac{n-1-2}{2}\right]\right)\right] = \left[\log_2\left[\frac{n+1}{2}\right]\right]$$

times (as $\varphi(c) \neq x, m$ for $c \in S^*$). The number of inequalities within $H - \{m\}$ is at least $n-1$. (13) follows.

Suppose now that $|A_1^l|=1$, that is $A_1^l=\{x\}$, $A_1'=\emptyset$. (13) obviously follows when $S^*=\emptyset$. We may suppose $S^*\neq\emptyset$. Let $a\in S^*$ be the first such element which go into S^* by a comparison with an element of A_2. Then, by Section 6, there were at least two comparisons of type (A_2, C) where the elements of A_2 were smaller. Let the first two of them be $b_1<a_1$, $b_2<a_2$ $(b_1, b_2\in A_2)$. b_1 must have been smaller before the comparison $b_1<a_1$, say $b_1<e$. The elements a_1, a_2, e are different, at most one of them is equal to x. Therefore, the graph induced by $H-\{x\}$ contains a circle, consequently, the number of its edges is $\geqq n-1$. x occurs in $\lceil\log_2 n\rceil-1$ inequalities. (13) is proved in this case.

We may suppose that if $e\in S^*$ then e occurred first time with an element $\notin A_2$. If B_1^l contains an element occurring 3 times (always as smaller) then the number of inequalities within $H-\{x\}$ is $\geqq n-1$. (13) follows. Suppose that the elements of B_1^l occur at most twice in $\mathscr{E}_l$. Let $a\in S^*$. If $a\in B_2^l$ then a occurs as larger: $a>b$ and b occurs first with an element of B_2^l $(b<c,\ c\in B_2^l)$. Otherwise $a<b$ would follow by Section 4 of the definition of the path. As $c\neq a$, there is a directed path from y to c and to a in the graph induced by $H-\{x\}$. It contains a circle and at least $n-1$ edges. (13) follows in the above way.

We may conclude that $S^*\subseteq B_1^l$. Let $S^*=\{a_1, \ldots, a_k\}$ and suppose that a_i go into S^* by the comparison $a_i>b_i$. We can state $b_i\in B_2^l$, $b_i\neq b_j$ $(i\neq j,\ i, j\in\{1, \ldots, k\})$. b_i must have been larger in $\mathscr{E}_l$ before: $b_i>c_i$ and c_i must have been smaller before: $c_i<e_i$ $(i=1, \ldots, k)$. If $e_i\neq x$ for some i then the graph induced by $H-\{x\}$ contains a circle. The proof of (13) is like before. We may suppose that $e_i=x$ $(i=1, \ldots, k)$. Suppose that the order of these inequalities in $\mathscr{E}_l$ is $x>c_1, \ldots, x>c_k$. If $i>1$ then x occurred before $x>c_i$. Take the inequality $a_i<c_i$ immediately in front of $x>c_i$ in $\mathscr{E}_l^*$. If x occurred before $x>c_1$ then take $a_1<c_1$ before $x>c_1$. In this modified system of inequalities x occurs at least $\lceil\log_2 n\rceil$ times. (13) follows.

We have to consider only the case when x does not occur before $x>c_1$ and $S^*=\{a_1\}$. If $b_1<d_1$ and $b_1>a_1$ are in $\mathscr{E}_l$ in this order (where $d_1\in A_1$) and $S_i(\varepsilon_1, \ldots, \varepsilon_i)=(b_1, c_1)$ then we have to apply Section 4, therefore $b_1, c_1\in B_1^i$. Let $S_j(\varepsilon_1, \ldots, \varepsilon_j)=(b_1, d_1)$. If d_1 occurs as larger in $\mathscr{E}_j$ then put the inequality $a_1<b_1$ immediately in front of $b_1<d_1$ and omit $c_1<b_1$. In this modified system of inequalities x occurs at least $\lceil\log_2 n\rceil$ times. (13) follows. Consequently, we may suppose that d_1 did not occur in $\mathscr{E}_j$. Let $S_k(\varepsilon_1, \ldots, \varepsilon_k)=(c_1, x)$. If $S_r(\varepsilon_1, \ldots, \varepsilon_r)=(x, f)$ where $r>k$ and $f\in C^r$ or $S_p(\varepsilon_1, \ldots, \varepsilon_p)=(d_1, f)$ where $p>j$ and $f\in C^p$ then put this inequality $a<f$ immediately in front of $x>c_1$ or $d_1>f$, respectively. Moreover, if x or d_1 occurs with an element of $A_2\cup B_2\cup B_1$ then replace this inequality with $x>a$ or $d_1>a$, respectively. In this new system of inequalities x occurs at least $\lceil\log_2 n\rceil$ times. (13) follows.

We suppose now that x occurs with elements of A_1 and whenever $d_1\in A_1$ it also occurs with elements of A_1, only. $x>d_1$ can be deduced from $\mathscr{E}_l$, consequently there is a sequence

$$d_1 = g_0 < g_1,\quad g_1 < g_2, \ldots, g_{k-1} < g_k = x$$

in it, where the above inequalities are the first occurrences of g_i as smaller $(0\leqq i\leqq k-1)$. It is easy to see that there is a pair (g_i, g_{i+1}) $(0\leqq i\leqq k-1)$ such that g_{i+1} occurs more times as larger than g_i does. Put $a<g_i$ in front of $g_i<g_{i+1}$. In this modified system of inequalities x occurs at least $\lceil\log_2 n\rceil$ times. (13) follows exactly in the same way as before.

Case 4. The proof of this case is a slight modification of the previous one. In this case we divide $B(\varepsilon_1, \ldots, \varepsilon_l)-\{x, y\}$ into subsets: $B_1(\varepsilon_1, \ldots, \varepsilon_l)-\{x, y\}= =B_1' \cup B_2' \cup B_2''$, where $B_1'=B_1(\varepsilon_1, \ldots, \varepsilon_l)-\{x, y\}$ and $B_2' \cup B_2''$ is a partition of $B_2(\varepsilon_1, \ldots, \varepsilon_l)$ $(B_2'(B_2'')$ consists of elements having a directed path from (to) both x and $y)$.

The proof of Case 3 is repeated with the roles of letters A and B interchanged. Lemma 1 is applied in the opposite way (determining the minimum element), but our "adversary" path $(\varepsilon_1, \ldots, \varepsilon_l)$ satisfies the conditions of this "opposite" lemma also.

This completes the proof of the case and the theorem.

The problem, when two unordered consecutive elements are determined, is slightly different. We obtain the same lower estimations as in the above proof except in Case 3 (and 4) when $A_1'=\emptyset$. In this case we obtain only $[\log_2 (n-1)]$ edges from x, using Lemma 1. That is, the lower estimation is $n-2+[\log_2 (n-1)]$. This coincides with $n-2+[\log_2 n]$ except when $n=2^m+1$. Indeed, there is a strategy with this many comparisons.

Choose any 2^m elements and find the maximum element with 2^m-1 steps with the "elimination tournament". There are exactly m element which proved to be smaller than the maximum. Complete this set with the (2^m+1)-st element and determine the maximum by m steps. The two maximum elements for the pair of the two largest elements. We needed $2^m-1+m=n-2+[\log_2 (n-1)]$ comparisons, indeed. That is, if $\mathscr{S}$ is a strategy determining two unordered consecutive elements in an n element set H, then

$$L(\mathscr{L}) \geqq n-2+[\log_2 (n-1)]$$

and this is the best possible bound.

Finally the author wishes to express his thanks to G.O.H. Katona for his help in writing this paper and to the referees for finding a mistake in the first version.

REFERENCES

[1] KATONA, G. O. H., *Combinatorial search problems*, A survey of combinatorial theory (dedicated to R. C. Bose on the occasion of his seventieth birthday, and containing the Proceedings of the Internat. Sympos. Combinatorial Math. and its Appl., Colorado State Univ., Fort Collins, Colo., 1971), North-Holland, Amsterdam, 1973, 285—308. *MR* **51** #4976.

[2] KIRKPATRICK, D. G., Ph. D. Thesis, University of Toronto, 1979.

[3] KISLICYN, S. S., On the selection of the kth element of an ordered set by pairwise comparisons, *Sibirsk. Math. Ž.* **5** (1964), 557—564 (in Russian). *MR* **29** #2198; errata **29**, p. 1437.

[4] KNUTH, D. E., *The art of computer programming, Vol. 3. Sorting and searching*, Addison—Wesley Series in Computer Science and Information Processing, Addison—Wesley Publishing Co., Reading, Mass—London—Don Mills, Ont., 1973, 211—212. *MR* **56** #4281.

[5] POHL, I., A sorting problem and its complexity, *Comm. ACM* **15** (1972), 462—464.

[6] SŁUPECKI, J., On the systems of tournaments, *Colloquium Math.* **2** (1951), 286—290. *MR* **13**—899.

[7] SCHREIER, I., On tournament elimination system, *Math. Polska* **7** (1932), 154—160.

[8] VARECZA, Á., Módszerek a rendezési algoritmusok elvi korlátainak meghatározására, *Alk. Mat. Lapok* (1980), 191—202.

(Received March 10, 1981; in revised form June 6, 1983)

BESSENYEI GYÖRGY TANÁRKÉPZŐ FŐISKOLA
MATEMATIKA TANSZÉK
SÓSTÓI ÚT 31/B
H—4400 NYÍREGYHÁZA
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 303—318

THE GROMOV—ELIASHBERG PROOF OF HAEFLIGER'S THEOREM

ANDRÁS SZŰCS

1. Introduction

In 1962 A. Haefliger proved a theorem on embeddings of smooth manifolds:

THEOREM (A. Haefliger [1]). *Let M be a closed manifold of dimension m and N an arbitrary manifold of dimension n, where $n \geqq \frac{3}{2}m+2$. Assume that there exists a continuous map F which satisfies the following two conditions:*

1. *the diagram below commutes:*

$$\begin{array}{ccc} M\times M & \xrightarrow{F} & N\times N \\ \downarrow & & \downarrow \\ M\times M & \xrightarrow{F} & N\times N \end{array}$$

where the vertical arrows indicate the mapping $(x, y)\to(y, x)$;

2. $F^{-1}(\Delta(N))=\Delta(M)$

where $\Delta(x)$ stands for the diagonal $\Delta(X)=:\{(x, x)\,|\,x\in X\}$.
Then there exists a differentiable embedding $M \subset\!\!\to N$.

Haefliger's proof is an extension of the method of elimination of double points, due to Whitney [2]. The aim of the present paper is to give a different proof of this theorem for the case $N=R^n$. The basic idea (to be explained in the next section) is due to M. L. Gromov and J. M. Eliashberg (oral communication). This proof has the advantage that it can also be applied to immersions without triple points, when $n<\frac{3}{2}m+2$ (this will be done in a subsequent paper) and to some other questions as well (see [7]).

2. Outline of the proof

2.1. DEFINITIONS. We shall call a mapping F: $X\times X\to Y\times Y$ *Z_2-equivariant* if $F(x_1, x_2)=(y_1, y_2)$ implies $F(x_2, x_1)=(y_2, y_1)$ for any $x_1, x_2\in X$. (This is the same as Condition 1, in the Theorem.) Note that this implies $\Delta(X)\subseteq F^{-1}(\Delta(Y))$.
If F is equivariant and $\Delta(X)=F^{-1}(\Delta(Y))$ then F will be called *isovariant.*

1980 *Mathematics Subject Classification.* Primary 57R40.
Key words and phrases. Imbeddings into Euclidean space, differentiable manifolds.

For $F\colon M\times M\to R^n\times R^n$ let us define the *coordinate functions* $f_i\colon M\times M\to R^1\times R^1$ $(i=1, \dots, n)$ by

$$f_i(u)=(t_i, t_{i+n}),$$

where

$$F(u)=(t_1, \dots, t_{2n}), \quad u\in M\times M.$$

We set $F=(f_1, \dots, f_n)$ in this case.

Observe that *F is equivariant if and only if all coordinate functions of F are equivariant.*

We call a mapping $f\colon M\times M\to R^1\times R^1$ a *holonomic mapping* if $f(x, y)=(h(x), h(y))$ for some $h\colon M\to R^1$. In this case we write $f=h\times h$.

2.2. *Fundamental idea.* Let now $N=R^n$ and M, F as in Theorem 1. Let f_i denote the coordinate functions of F. Our plan is to substitute holonomic mappings $f_i'=h_i\times h_i\colon M\times M\to R^1\times R^1$ for the coordinate functions of F such that $(f_1', \dots, f_i', f_{i+1}, \dots, f_n)\colon M\times M\to R^n\times R^n$ be an isovariant map for every $i=1, \dots, n$. It is then clear that the obtained $h=(h_1, \dots, h_n)\colon M\to R^n$ is an embedding.

2.3. *The induction step for the simplest case.* f_i' will be constructed by induction on i. Assume that the holonomic mappings $f_j'=h_j\times h_j\colon M\times M\to R^1\times R^1$ have already been constructed for $j=1, \dots, i-1$ such that the map $F_{i-1}=(f_1', \dots, f_{i-1}', f_i, \dots, f_n)$ is isovariant. Set $\bar F=(f_1', \dots, f_{i-1}', f_{i+1}, \dots, f_n)\colon M\times M\to R^{n-1}\times R^{n-1}$. For simplicity, we denote f_i by f. By the condition $(\bar F, f)$ is isovariant. We want to replace f by a holonomic map $f'=h\times h$, $h\colon M\to R^1$, such that $(\bar F, f')$ be isovariant. This is particularly simple to do under the following

Assumption (*). The restriction of the projection $\pi_1\colon M\times M\to M$ $(\pi_1(x, y)=x)$ to the set $A_{\bar F}=\overline{(\bar F^{-1}(\Delta(R^{n-1}))\setminus\Delta(M))}$ is injective. (The bar indicates closure.)

Lemma. *Under (*) there exists an $f'=h\times h$ such that $(\bar F, f')\colon M\times M\to R^n\times R^n$ is isovariant.*

Proof. Let $f(x, y)=(\varphi(x, y), \psi(x, y))$ $(x, y\in M)$. (So, $\varphi, \psi\colon M\times M\to R^1$.) For $x\in\pi_1(A_{\bar F})$, define $h(x)$ to be $h(x)=\varphi(x, y)$ where y is the unique point of M such that $(x, y)\in A_{\bar F}$. Extend this to a smooth function $h\colon M\to R^1$ arbitrarily. We assert that $(\bar F, f')$ is isovariant if setting $f'=h\times h$. In fact, the following four statements are equivalent:

(a) $(\bar F, f')$ is isovariant;

(b) $f'(A_{\bar F}\setminus\Delta(M))\cap\Delta(R^1)=\emptyset$;

(c) $f(A_{\bar F}\setminus\Delta(M))\cap\Delta(R^1)=\emptyset$;

(d) $(\bar F, f)$ is isovariant.

The equivalence of (b) and (c) is clear since f and f' agree on $A_{\bar F}$. For (a)⇔(b), observe that,

1° $A_{\bar F}\setminus\Delta(M)=\bar F^{-1}(\Delta(R^{n-1}))\setminus\Delta(M)$

and

2° for any $u \in M \times M$, $(\bar{F}, f')(u) \in \Delta(R^n)$ if and only if $\bar{F}(u) \in \Delta(R^{n-1})$ and $f'(u) \in \Delta(R^1)$.

The proof of equivalence (c)⇔(d) is completely analogous. ∎

Our main task will be to reduce the general case to the one just treated. We shall modify $\bar{F}$ such as to obtain a new $\bar{F}$ satisfying (*). Of course, the first $2(i-1)$ coordinate functions of the new $\bar{F}$ have to remain holonomic ones after this modification.

2.4. *Reduction to* (*). 1. First we attempt to modify $\bar{F}$ such that $A_{\bar{F}}$ becomes a (closed) manifold. A next step of modification yields that $\pi_1|_{A_{\bar{F}}}$: $A_{\bar{F}} \to M$ becomes an immersion and, after a third step, an embedding (Sections 6 and 7). These steps will follow from certain modifications of Thom-Haefliger type transversality theorems (Section 5).

2. However, these transversality theorems work only outside a neighbourhood of the diagonal $\Delta(M)$. Therefore, the mentioned properties of $A_{\bar{F}}$ and $\pi_1|_{A_{\bar{F}}}$ have to be assured by a different method in a neighbourhood N of $\Delta(M)$ (Sections 3 and 4). First of all, we replace F by a map also denoted by F which agrees on N with $g \times g$, for some immersion g: $M \to R^n$. (This will be done by means of a lemma of Haefliger—Hirsch on skew maps [3] and the Hirsch theorem on immersion [4].) The advantage of this new map F is that it is also isovariant and keeps this property after a slight movement, too. After this, it will be easy to assure by slight movements of F that $A_{\bar{F}}$ be a manifold, furthermore $\pi_1|_{A_{\bar{F}}}$ be an immersion, and finally it will be an embedding.

Remark. Denote $\bar{g}$: $M \to R^{n+k-1}$ the map for which $\bar{F}|_N = \bar{g} \times \bar{g}|_N$.

Notice that the derivative of h on $\pi_1(\Sigma^1(\bar{g}))$ in the direction Ker $d\bar{g}$ is not zero. Hence the new map $F = (\bar{F}, h \times h)$ also has the form $g \times g$ in a neighbourhood of $\Delta(M)$ for a (may be new) immersion g of M into R^{n+k}.

3. Preliminaries

3.1. Notation. Let $TM \to M$ denote the tangent bundle of the m-dimensional manifold M. Let $DM \to M$ and $SM \to M$ denote the associated fibre bundle with fibre D^m, the m-dimensional ball and S^{m-1}, the $(m-1)$-sphere, resp.

3.2. We shall often make use of the following well-known fact: The tubular neighbourhood N of the diagonal $\Delta(M)$ of $M \times M$ can be identified with DM. The boundary ∂N of N corresponds to SM.

The space of mappings $X \to Y$ will be denoted by $\{X \to Y\}$. Having identified N with DM, we then identify the spaces of maps

$$\{N \to X\} = \{DM \to X\}, \quad \{\partial N \to X\} = \{SM \to X\}$$

for any space X.

Let ξ: $E(\xi) \to B(\xi)$ be a vector bundle. The above identification allows us to speak about the *linear* mappings $N \to E(\xi)$. The symbol $L(N, E(\xi))$ stands for the space of $N \to E(\xi)$ linear mappings.

3.3. We recall the explicit formula for the identification DM with N. With $v\in T_mM$ $(m\in M)$ let us associate the point $(\exp_m(v), \exp_m(-v))\in M\times M$. (An arbitrary riemannian metric is chosen on M.)

As it is well-known, there exists an $\varepsilon>0$ such that, setting $D_\varepsilon M=\{v\in TM|\,\|v\|\leqq\varepsilon\}$, the restriction to $D_\varepsilon M$ of the above mapping $TM\to M\times M$ is a diffeomorphism onto some tubular neighbourhood N of $\Delta(M)$ (in $M\times M$).

We have recalled this formula in order to emphasize *that the identification map* $v\to(\exp_m v, \exp_m(-v))$ *is equivariant* with respect to the involutions $v\to -v$ $(v\ni TM)$ and $(x, y)\to(y, x)((x, y)\in M\times M)$. (Observe, that $D_\varepsilon M\subset TM$ and $N\subset M\times M$ are invariant under the corresponding involution.) Henceforth we preserve the term tubular neighbourhood for the image of $D_\varepsilon M$ under the above identification. ($\varepsilon>0$ is sufficiently small.)

4. Making $\tilde{F}$ nice within N

The following definition is due to A. Haefliger [5]:

4.1. Definition. Let M, L be manifolds, of respective dimensions m and l. A C^2-map $f\colon M\to L$ is of S-type satisfying

1° rank $df_x\geqq m-1$ $(x\in M)$;

2° there are no triple points;

3° there are no singular double points;

4° the self-intersections are transversal;

5° there exist local coordinates $(x_1, \dots, x_m)$ in the neighbourhood of any singular point such that $\left.\frac{\partial f}{\partial x_1}\right|_0=0$ and the $(2m-2)$ vectors $\left.\frac{\partial f}{\partial x_i}\right|_0$, $\left.\frac{\partial^2 f}{\partial x_1\partial x_i}\right|_0$ $(i=2, \dots, m)$ generate R^l.

Let $pr_i\colon R^n\to R^{n-1}$ denote the projection which deletes the i'th coordinate.

4.2. Definition. An immersion $g\colon M\to R^n$ is *generic,* if

1° its self-intersections are transversal;

2° $pr_i\circ g\colon M\to R^{n-1}$ is an S-type map $(i=1, \dots, n)$.

4.3. Lemma. *Let M be a manifold of dimension m, and $n\geqq\frac{3}{2}m+2$. Given an isovariant map $F\colon M\times M\to R^n\times R^n$ there exist an isovariant map $F'\colon M\times M\to R^n\times R^n$ and a general immersion $g\colon M\to R^n$ such that $F'|_N=g\times g|_N$ for some tubular neighbourhood N of $\Delta(M)$.*

Proof. We break the proof of Lemma in 4.3 into two sublemmas. We note that the tubular neighbourhoods $D_\varepsilon M$ do not depend on ε up to diffeomorphism (for small enough ε). Let DM denote their diffeomorphism class.

4.4. Sublemma. *Let X and Y be subspaces of the space E of isovariant maps $DM\to R^n\times R^n$ such that any $x\in X$ is homotopic in E with some $y\in Y$. Let F be an equivariant map $M\times M\to R^n\times R^n$ such that $F|_{D_\varepsilon M}\in X$. Then there exists an equivariant map $F'\colon M\times M\to R^n\times R^n$ such that $F'|_{D_\delta M}\in Y$ (where $\delta<\varepsilon$ and we identify $D_\delta M$ with DM this time).*

PROOF. $D_\varepsilon M \setminus D_\delta M$ is diffeomorphic to $SM\times[0, 1]$. Let $x=F|_{D_\varepsilon M}$ and let H: $DM\times[0, 1]\to R^n\times R^n$ be a homotopy joining x with an $y\in Y$ $(y=H(.,1))$. We define F' as follows:

1° $F'|_{M\times M\setminus D_\varepsilon M}=F|_{M\times M\setminus D_\varepsilon M}$;
2° We identify $\overline{D_\varepsilon M\setminus D_\delta M}$ with $SM\times[0, 1]$ and set $F|_{D_\varepsilon M\setminus D_\delta M}=H|_{SM\times[0,1]}$;
3° $F'|_{D_\delta M}=y$ (where, again, DM is identified with $D_\delta M$). ▮

We introduce certain subspaces of the space E of isovariant maps $DM\to R^n\times R^n$.

Let Mono = Mono(DM, TR^n) be the space of linear monomorphisms $DM\to TR^n$. The isovariant identification of TR^n with $R^n\times R^n$ $(R^n=T_xR^n\ni v\mapsto(x+v, x-v)\in \in R^n\times R^n$, cf. 3.3) permits us to view Mono as a subspace of E.

Let $D\subset$ Mono consists of those maps G: $DM\to TR^n$ which satisfy

$$d(G|_{\Delta(M)})|_{DM}=G.$$

In other words $G|_{\Delta(M)}$: $\Delta(M)\to\Delta(R^n)$ is an immersion $(M=\Delta(M),\ R^n=\Delta(R^n))$ whose differential agrees on $DM\subset TM$ with G. Let

$Q=\{G\in E|$ G is the restriction of $g\times g$ to the tubular neighbourhood DM of $\Delta(M)$ for some immersion g: $M\to R^n\}$.

Finally, let $Q^0\subset Q$ consist of those maps $g\times g|_{DM}$ where g is a general immersion.

REMARK. For the spaces introduced above — by definition — the following inclusions hold:

$$E\supset \text{Mono}\supset D\supset Q\supset Q^0.$$

4.5. SUBLEMMA. *Let X and Y be any two consecutive members of the sequence of the Remark. Then the conditions of Sublemma in 4.4 are satisfied.*

PROOF. (a) $X=E$, $Y=$Mono. This is a reformulation of Theorem 2.3 of Haefliger—Hirsch [3].

(b) $X=$Mono, $Y=D$. This is a well-known theorem of Hirsch on immersions [4].

(c) $X=D$, $Y=Q$. We observe that for any immersion g: $M\to R^n$, the difference between $dg|_{D_\varepsilon M}$ and $g\times g|_{D_\varepsilon M}$ is $O(\varepsilon^2)$ and therefore, they are homotopic in E if ε is sufficiently small.

(d) $X=Q$, $Y=Q^0$. Any immersion g can be approximated by a generic immersion $\bar g$. If $\bar g$ is sufficiently close to g then $g\times g|_{DM}$ and $\bar g\times\bar g|_{DM}$ are homotopic in E. ▮

Now 4.3 is immediate by 4.4 and 4.5. ▮

4.6. LEMMA. *Let $\bar g$: $M\to R^{n-1}$ be an S-type map. Set*

$$A=\overline{(\bar g\times\bar g)^{-1}(\Delta(R^{n-1}))\setminus\Delta(M)}.$$

Then $\Delta(M)$ has a neighbourhood N in $M\times M$ such that the restriction of the projection π_1: $M\times M\to M$ $(\pi_1(x, y)=x)$ to $A\cap N$ is an embedding.

PROOF. Let $k=n-m-1$ ($m=\dim M$). Introduce the submanifolds $\Sigma^1(\bar g)$ and $\Delta(\bar g)$ of M and an involution $T\colon \Delta(\bar g)\to\Delta(\bar g)$ as follows:

$$\Sigma^1(\bar g) = \{x\in M \mid \operatorname{rank} d\bar g(x) = m-1\};$$

$$\Delta(\bar g) = \Sigma^1(\bar g)\cup\{x\in M \mid \bar g(x) = \bar g(x') \text{ for some } x'\neq x,\ x'\in M\}.$$

$$T(x) = \begin{cases} x & \text{if } x\in\Sigma^1(\bar g) \\ x' & \text{if } x, x'\in\Delta(g)\setminus\Sigma^1(g),\ x\neq x' \text{ and } \bar g(x) = \bar g(x'). \end{cases}$$

We note that

1° $\dim \Sigma^1(\bar g)=m-k-1$,
2° $\dim \Delta(\bar g)=m-k$.

Now we observe that, for any sufficiently small tubular neighbourhood N of $\Delta(M)$ in $M\times M$ there exists a tubular neighbourhood L of $\Sigma^1(\bar g)$ in $\Delta(\bar g)$ such that

$$A\cap N = \{(x, y)\in L\times L \mid x = T(y)\}.$$

Hence $A\cap N$ is the graph of $T|_L$. The projection π_1 restricted to this graph being a diffeomorphism $A\cap N\to L$, the assertion of Lemma in 4.6 follows. ∎

4.7. COROLLARY. *Let $F\in Q^0$. (Q^0 has been defined before 4.5.) Define $\bar F$ and $A_{\bar F}$ as in 2.3. Then $\Delta(M)$ has a neighbourhood N in $M\times M$ such that $\pi_1|_{A_{\bar F}\cap N}$ is an embedding $A_{\bar F}\cap N\to M$.*

PROOF. By the definition of Q^0, $\Delta(M)$ has a neighbourhood in $M\times M$ such that $F|_N=g\times g|_N$ for some general immersion g. Hence $\bar F|_N=\bar g\times\bar g|_N$ where $\bar g=g\circ pr_i\colon M\to R^{n-1}$ is of S-type (cf. 2.4). The conditions of Lemma in 4.6 are fulfilled, whence the corollary follows. ∎

4.8. Given a natural number $1\leq i\leq n-1$ and a neighbourhood N of $\Delta(M)$ in $M\times M$ we define space $C_N(i)$ as the space of pairs (H, h) such that

(1) $H\colon M\times M\to R^{n-i}\times R^{n-i}$ is a Z_2-equivariant map;
(2) $h\colon M\to R^{i-1}$ is an arbitrary smooth map;
(3) the map $\varphi=(H, h\times h)\colon M\times M\to R^{n-1}\times R^{n-1}$ restricted to N coincides with the restriction of a map $\bar g\times\bar g\colon M\times M\to R^{n-1}\times R^{n-1}$, where $\bar g$ is an arbitrary S-type map. (Sometimes we shall identify the pair (H, h) with the map $\varphi=(H, h\times h)$.)

4.9. In 4.3 it was shown that the space $C_N(1)$ is not empty. The argumentation in 2.3 shows that if there exists a map $\varphi\in C_N(i)$ which satisfies assumption (*) then $C_N(i+1)$ is not empty as well $(i+1\leq n)$. In Sections 6, 7, 8 we will show that there exists a neighbourhood N_2 (which may be smaller than N) such that in $C_{N_2}(i)$ there exists a map which satisfies (*). In what follows number i will be fixed and $C_N(i)$ will be denoted by C_N.

5. Transversality theorems

A reader, who is ready to accept all kinds of "general position" argumentations, can omit this section.

In this section first we recall the Thom's and Haefliger's transversality theorems, and then we formulate another one, which is their "mixture". Since the formulation of this mixed transversality theorem is very cumbersome, we start with a special case.

NOTATION. The sign $\pitchfork$ will denote "transversal".

5.1. THEOREM (Thom's transversality theorem). *Let A_i and x_i be smooth manifolds for $i=1, 2, \ldots, k$. Let V be a submanifold of the space of r-jets of maps from the product space $\prod_{i=1}^k A_i$ into the product space $\prod_{i=1}^k X_i$, i.e. $V\subset J^r\left(\prod_{i=1}^k A_i, \prod_{i=1}^k X_i\right)$. Define the subset Ω of the space $C=\prod_{i=1}^k C^\infty(A_i, X_i)$ by*

$$\Omega =: \{(f_1, \ldots, f_k)\in C \mid j^r(f_1\times\ldots\times f_k)\pitchfork V\}.$$

Then Ω is open and dense in C.

PROOF in [6].

5.2. NOTATIONS. For an arbitrary space Y denote by $Y^{(k)}$ the product $Y\times\ldots\times Y$ (k factors) and by $\Delta^{(k)}Y$ the following subset of $Y^{(k)}$.

$$\Delta^{(k)}Y=\{(y_1, \ldots, y_k)\in Y^{(k)} \text{ (there exist indices } i, j \text{ such that } i\neq j \text{ and } y_i=y_j\}$$

($\Delta^{(2)}Y$ will be denoted by $\Delta(Y)$, too, as in 2.3). For an arbitrary map $f\colon X\to Y$ let $f^{(k)}\colon X^{(k)}\to Y^{(k)}$ be defined by

$$f^{(k)}(x_1, \ldots, x_k) = (f(x_1), f(x_2), \ldots, f(x_k)).$$

THEOREM (Haefliger's transversality theorem). *Take $A_i=A$ and $X_i=X$, $i=1, \ldots, k$, in Thom's transversality theorem. Fix an open subset $U\subset A^{(k)}$ disjoint from $\Delta^{(k)}(A)$. Let*

$$\Omega = \{f\in C^\infty(A, X) \mid j^r(f^{(k)}|_U)\pitchfork V\}.$$

Then Ω is open and dense in $C^\infty(A, X)$.

PROOF in [5].

5.3. We shall need a modification of Thom's and Haefliger's transversality theorems in which some of the manifolds A_i (and X_i) will be the same, while others will be different. We shall present two versions of this theorem.

THEOREM (Mixed transversality theorem, special case, SMTT).

SMTT (a). *Let A_1, A_2, X_1, X_2, Y be manifolds and put $B=A_1\times A_2$, $Z=X_1\times X_2$. Let V be a submanifold of $B\times Y\times X$ and*

$$\Omega = \{(f_1, f_2, g)\in C \mid j^0(f_1\times f_2, g)\pitchfork V\}$$

where

$$C = C^\infty(A_1, X_1)\times C^\infty(A_2, X_2)\times C^\infty(B, Y).$$

Then Ω is open and dense in C. (Note that the map $(f_1\times f, g)\colon B\to Z\times Y$ is defined by $(f_1\times f_2, g)(a_1, a_2)=(f_1(a_1), f_2(a_2), g(a_1, a_2))$.)

SMTT (b). *Keeping the conditions of* (a) *with* $A_1=A_2=A$ *and* $X_1=X_2=X$, *fix an open subset U in* $A\times A$ *disjoint from* $\Delta(A)$. *Let* $\Omega_U=\{(f,g)\in C \mid j^0(f\times f, g)|_U \pitchfork V\}$ *where*

$$C = C^\infty(A, X)\times C^\infty(B, Y).$$

Then Ω_U *is open and dense in* C.

PROOF. (a). It is easy to see that

$$j^0(f_1\times f_2, g)\pitchfork V \quad \text{in} \quad B\times Y\times Z \Leftrightarrow j^0(f_1\times f_2\times g)\pitchfork V \quad \text{in} \quad B\times B\times Y\times Z$$

where V is embedded in $B\times B\times Y\times Z$ by the composition

$$V\subset B\times Y\times Z \approx \Delta(B)\times Y\times Z\subset B\times B\times Y\times Z.$$

(Notice the difference between the map $(f_1\times f_2, g)$ and $(f_1\times f_2\times g)$ where f_1: $A_1\to X_1$, f_2: $A_2\to X_2$, g: $B=A_1\times A_2\to Y$. The definition of the map $(f_1\times f_2, g)$ is the following:

$$(f_1\times f_2, g)\colon A_1\times A_2\to X_1\times X_2\times Y$$

$$(a_1, a_2) \rightsquigarrow (f_1(a_1), f_2(a_2), g(a_1, a_2))$$

while the map $f_1\times f_2\times g$ is defined as follows:

$$f_1\times f_2\times g\colon A_1\times A_2\times A_1\times A_2 \to X_1\times X_2\times Y$$

$$(a_1, a_2, a_1', a_2') \rightsquigarrow (f_1(a_1), f_2(a_2), g(a_1', a_2')).$$

From Thom's theorem we obtain that Ω is open and dense in C.

(b) Following Haefliger, from Thom's transversality theorem we reduce SMTT (b) to SMTT (a) who derived this transversality theorem. We give this proof in details.

Consider a finite covering of $U\subset A\times A\setminus\Delta(A)$ with open sets of the form $W_1^\alpha\times W_2^\alpha$ $\alpha=1, \ldots, r$, where W_1^α and W_2^α are disjoint open subsets in A. Let

$$\Omega_\alpha = \{(f, g)\in C \mid j^0(f\times f, g)|_{W_1^\alpha\times W_2^\alpha}\pitchfork V\}.$$

Applying MTT (a) we obtain that Ω_α is open and dense in C. Hence the same holds for $\Omega_U=\bigcap_{\alpha=1}^{r}\Omega_\alpha$.

5.4. THEOREM (Mixed transversality theorem, general case GMTT). (a) *Consider the families of manifolds*

$$A_1^1, \ldots, A_{\alpha_1}^1, A_1^2, \ldots, A_{\alpha_2}^2, \ldots, A_1^s, \ldots, A_{\alpha_s}^s;$$

$$X_1^1, \ldots, X_{\alpha_1}^1, X_1^2, \ldots, X_{\alpha_2}^2, \ldots, X_1^s, \ldots, X_{\alpha_s}^s;\ Y_1, \ldots, Y_s$$

and put

$$B^t = \prod_{j=1}^{\alpha_t} A_j^t, \quad Z^t = \prod_{j=1}^{\alpha_t} X_j^t, \quad t = 1, \ldots, s.$$

Let V be a submanifold in $\prod_{t=1}^{s} B^t \times Y_t \times Z_t$ *and*

$$\Omega = \{(\Phi_1, \dots, \Phi_s, g_1, \dots, g_s) \in C \mid j^0(\Phi_1 \times \dots \times \Phi_s, g_1 \times \dots \times g_s) \pitchfork V\}$$

where

$$C = \prod_{t=1}^{s} \left\{ \prod_{j=1}^{\alpha_t} C^\infty(A_j^t, X_j^t) \times C^\infty(B^t, Y_t) \right\}$$

and Φ_t is a map of the form $f_1^t \times \dots \times f_{\alpha_t}^t \colon A_1^t \times \dots \times A_{\alpha_t}^t \to X_1^t \times \dots \times X_{\alpha_t}^t$ while g_t is a map $B^t \to Y_t$.

Then Ω is open and dense in C.

(b) *Keeping the conditions of* (a) *with $A_1^t = \dots = A_{\alpha_t}^t$, $X_1^t = \dots = X_{\alpha_t}^t$, $t = 1, 2, \dots, s$, and denote these manifolds by A_t and X_t, respectively. Let U be an open subset in* $\prod_{t=1}^{\alpha_t} A_t^{(\alpha_t)}$ *which is disjoint from the set*

$$\bigcup_{t=1}^{s} \Delta^{(\alpha_t)} A_t \times \prod_{t' \neq t} A_{t'}.$$

Define

$$\Omega =: \{(f_1, \dots, f_s, g_1, \dots, g_s) \in C \mid \{(f_1^{(\alpha_1)}, g_1) \times \dots \times (f_s^{(\alpha_s)}, g_s)|_U\} \pitchfork V\}$$

where $C = \prod_{t=1}^{s} C^\infty(A_t, X_t) \times C^\infty(A_t^{(\alpha_t)}, Y_t)$.

Then Ω_U is open and dense in C.

Remark. The corresponding jet-transversality variants of these theorems, which can be formulated in a similar way, will be applied, too.

6. Applications of the transversality theorems

6.1. Here we show that for almost every map $\varphi \in C_N$, A_φ is a manifold.

Proposition. *Denote by Ω^1 the subspace of C_N defined as follows*

$$\Omega^1 = \{\varphi \in C_N \mid \varphi|_{M \times M \setminus \Delta(M)} \pitchfork \Delta(R^{n-1})\}.$$

Then Ω^1 is open and dense in C_N.

Proof. Consider a finite open covering $\{W_1^\alpha \times W_2^\alpha \mid \alpha = 1, \dots, r\}$ of $M \times M \setminus N$, where W_1^α and W_2^α are open disjoint sets in M. Fix an α, $1 \leq \alpha \leq r$, and apply SMTT (a) with

$$A_1 = W_1^\alpha, \quad A_2 = W_2^\alpha, \quad X_1 = X_2 = R^{i-1},$$

$$Y = R^{n-i} \times R^{n-i}, \quad V = W_1^\alpha \times W_2^\alpha \times \Delta(R^{n-i}) \times \Delta(R^{i-1}).$$

We obtain that the corresponding set Ω_α of transversal maps forms an open and dense subset in the space

$$C_\alpha = \{C^\infty(W_1^\alpha \times W_2^\alpha, R^{n-i} \times R^{n-i}) \times C^\infty(W_1^\alpha, R^{n-i}) \times C^\infty(W_2^\alpha, R^{i-1})\}.$$

The restriction induces a map

$$\tau_\alpha\colon C_N \to C_\alpha$$

by the formula

$$\tau_\alpha(H, h) = (H|_{W_1^\alpha \times W_2^\alpha}, h|_{W_1^\alpha}, h|_{W_2^\alpha}).$$

Since this map τ_α is open, the preimage $\tau_\alpha^{-1}(\Omega_\alpha)$ is open and dense in C_N. Hence the set

$$\Omega^1 = \bigcap_{\alpha=1}^{r} \tau_\alpha^{-1}(\Omega_\alpha)$$

is also open and dense.

6.2. Here we show that for almost every $\varphi=(H, h\times h)$ the restriction is one to one on small subsets

PROPOSITION. *Given neighbourhoods N_1, N_2 of $\Delta(M)$ in $M\times M$ such that $N_1 \subset N_2 \subsetneqq N$ and a positive number $\varkappa$ define the space $\Omega^2(N_2, N_1; \varkappa)$ as follows:*

$$\Omega^2(N_2, N_1; \varkappa) = \{\varphi\in C_N \mid \text{if } a_1\in A_\varphi\setminus N_1,\ a_2\in A_\varphi\setminus N_1 \text{ and}$$

$$\varrho(a_1, a_2) \leqq \varkappa \text{ then } \pi_1(a_1) \neq \pi_1(a_2)\}.$$

Then $\Omega^2(N_2, N_1; \varkappa)$ is dense in C_N. (For definition of C_N see 4.8 and 4.9.)

PROOF. Let us consider the subspace $X\subset M^{(4)}$ consisting of the points (m_1, m_2, m_3, m_4) such that:

(1) $m_1=m_3$,
(2) the distances of m_1, m_2, m_4 are greater than $\varkappa$.

This subspace X can be covered by finitely many open sets $W_1^\alpha\times W_2^\alpha\times W_3^\alpha\times W_4^\alpha$, $\alpha=1,\dots, r$, where

(1) W_i^α is open in M^n.
(2) $W_i^\alpha\cap W_j^\alpha=\emptyset$ unless $i=1$ and $j=3$ (or $i=3$ and $j=1$).
(3) $W_1^\alpha=W_3^\alpha$.
(4) $W_i^\alpha\times W_{i+1}^\alpha\subset N$ or $(W_i^\alpha\times W_{i+1}^\alpha)\cap N_2=\emptyset$ for $i=1$ and 3.

(a) First suppose that α is such that $(W_i^\alpha\times W_{i+1}^\alpha)\cap N_2=\emptyset$ for $i=1$ and 3. Then we apply GMTT (a) with $s=2$, $\alpha_1=2$, $\alpha_2=2$.

$$A_1^1 = W_1^\alpha,\quad A_2^1 = W_2^\alpha,\quad A_1^2 = W_3^\alpha,\quad A_2^2 = W_4^\alpha,$$

$$X_1^1 = X_2^1 = X_1^2 = X_2^2 = R^{i-1},$$

$Y_1=Y_2=R^{n-1}$ and $V=V_1\times V_2\times V_3$ where

$$V_1 = \{(m_1, m_2, m_3, m_4) \mid m_i\in W_i^\alpha \text{ and } m_1 = m_3\}$$

$$V_2 = \{(x_1, x_2, x_3, x_4) \mid x_i\in R^{i-1},\ x_1 = x_2 \text{ and } x_3 = x_4\},$$

$$V_3 = \{(y_1, y_2, y_3, y_4) \mid y_i\in R^{n-i},\ y_1 = y_2 = y_4\}. \quad (\text{No } y_3!)$$

We obtain, that the corresponding set Ω_α of the maps with transversal to $V=V_1\times V_2\times V_3$ graphs is an open and dense subset of the space

$$C_\alpha = C^\infty(W_1^\alpha\times W_2^\alpha, R^{n-i}\times R^{n-i})\times C^\infty(W_3^\alpha\times W_4^\alpha, R^{n-i}\times R^{n-i})\times$$

$$\times C^\infty(W_1^\alpha, R^{i-1})\times C^\infty(W_2^\alpha, R^{i-1})\times C^\infty(W_3^\alpha, R^{i-1})\times C^\infty(W_4^\alpha, R^{i-1}).$$

The transversality to V is nothing but disjointness from V, because codim $V=m+2(i-1)+2(n-i)>4m=\dim W_1^\alpha\times W_2^\alpha\times W_3^\alpha\times W_4^\alpha$. Let $H\colon M\times M\to R^{n-i}\times\times R^{n-i}$ be as before a Z_2-equivariant map, and $h\colon M\to R^{i-1}$, $\hat{h}\colon M\to R^{i-1}$ arbitrary smooth maps such that $(H,h)\in C_N$. Then for any α, $1\leq\alpha\leq r$ the restriction

$$(H,h,\hat{h})\xrightarrow{\tau_\alpha}(H|_{W_1^\alpha\times W_2^\alpha}, H|_{W_3^\alpha\times W_4^\alpha}, h|_{W_1^\alpha}, h|_{W_2^\alpha}, h|_{W_3^\alpha}, h|_{W_4^\alpha})$$

induces a map

$$\tau_\alpha\colon C_N\times C^\infty(M;\, R^{i-1}).$$

Since τ_α is an open surjective map $\tau_\alpha^{-1}(\Omega_\alpha)$ is open and dense in $C_N\times C^\infty(M;\, R^{i-1})$. Hence the intersection $Q_a=\bigcap\tau_\alpha^{-1}(\Omega_\alpha)$ is open and dense in $C_N\times C^\infty(M;\, R^{i-1})$.

(b) If $W_1^\alpha\times W_2^\alpha\subset N$ and $(W_3^\alpha\times W_4^\alpha)\cap N_2=\emptyset$ then we change the definition of the space C_α: replace its first factor $C^\infty(W_1^\alpha\times W_2^\alpha, R^{n-i}\times R^{n-i})$ by the product $C^\infty(W_1^\alpha, R^{n-i})\times C^\infty(W_2^\alpha, R^{n-i})$ and in this new C_α consider the subspace Ω_α of those maps whose graphs are transversal to V.

(c) If $(W_1^\alpha\times W_2^\alpha)\cap N_2=\emptyset$ and $W_3^\alpha\times W_4^\alpha\subset N$ then we change in the (original) definition of C_α the second factor, replace it by $C^\infty(W_3^\alpha, R^{n-1})\times C^\infty(W_4^\alpha, R^{n-i})$.

(d) If $W_1^\alpha\times W_2^\alpha\subset N$ and $W_3^\alpha\times W_4^\alpha\subset N$ then we change both the first and the second factors of C_α in the previous way.

In all the three cases (b), (c), (d) — like in case (a) — we consider the spaces Ω_α and the maps τ_α and obtain that the intersection $Q_\varepsilon=\bigcap\tau_\alpha^{-1}(\Omega_\alpha)$ is open and dense in $C_N\times C^\infty(M,R)$ for $\varepsilon=a,b,c,d$.

Then $Q=Q_a\cap Q_b\cap Q_c\cap Q_d$ is open and dense in $C_N\times C^\infty(M,R^{i-1})$.

Lemma. $(H,h,\hat{h})\in Q\Rightarrow(H,h)\in\Omega^2(N_2,N_1,\varkappa)$.

Proof. Suppose indirectly that $(H,h)\notin\Omega^2(N_2,N_1,\varkappa)$. This means that there exist points $a_1=(m_1,m_2)$ and $a_2=(m_3,m_4)$ in $M\times M$ such that

(i) $\pi_1(a_1)=\pi_1(a_2)$, i.e. $m_1=m_3$;

(ii) $\varrho(m_1,m_2)$, $\varrho(m_3,m_4)$, $\varrho(m_2,m_4)$ are greater than XY;

(iii) $a_1\in A_\varphi$ (where $\varphi=(H,h\times h)$), i.e. $H(m_1,m_2)\in\Delta(R^{n-i})$ and $h(m_1)==h(m_2)$;

(iv) $a_2\in A_\varphi$, i.e. $H(m_3,m_4)\in\Delta(R^{n-i})$ and $h(m_3)=h(m_4)$. For some α $(m_1,m_2,m_3,m_4)\in W_1^\alpha\times W_2^\alpha\times W_3^\alpha\times W_4^\alpha$. The map

$$(H|_{W_1^\alpha\times W_2^\alpha}, H|_{W_3^\alpha\times W_4^\alpha}, h|_{W_1^\alpha}, h|_{W_2^\alpha}, \hat{h}|_{W_3^\alpha}, h|_{W_4^\alpha})$$

maps the point (m_1,m_2,m_3,m_4) into V and hence $(H,h,\hat{h})$ does not belong to Q. Hence $\Omega^2(N_2,N_1;\varkappa)$ is dense in C_N. ∎

6.3. Here we show that the restriction $\pi_1|_{A_\varphi}$ of the projection $\pi_1\colon M\times M\to M$ $(\pi_1(x, y)=x)$ is an immersion for almost every map $\varphi\in C_N$.

PROPOSITION. *Define the subspace Ω^3 of C_N as follows:*

$$\Omega^3 = \{\varphi\in C_N \mid \pi_1|_{A_\varphi} \text{ is an immersion}\}.$$

Then Ω^3 is open and dense in C_N.

PROOF. First of all recall that inside N (i.e. on $A_\varphi\cap N$) the projection π_1 is an immersion for any element of C_N. To prove that $\pi_1|_{A_\varphi\setminus N_2}$ is an immersion we make use again the transversality theorems. Define the submanifold $V\subset j^1(M\times M, R^{n-1}\times\times R^{n-1})$ as follows:

$$V = \{(m_1, m_2, y_1, y_2, r) \mid \text{where } m_1, m_2\in M,\quad y_1, y_2\in R^{n-1},$$

$$y_1=y_2,\ r \text{ is the matrix } \frac{\partial(y_1, y_2)}{\partial(m_1, m_2)},\ \text{and rank}\left(\frac{\partial y_1}{\partial m_1}-\frac{\partial y_2}{\partial m_1}\right)<\dim M=m\}.$$

(Actually V is not a submanifold, but it is a stratified subset, with finitely many strata and so the transversality theorems can be applied to V. Notice that codim $V=n-1+$ $+n-m=2n-m>\dim M\times M=2m$.) Consider a finite open covering $\{W_1^\alpha\times W_2^\alpha|\alpha$ $=1, \ldots, r\}$ of $U=M\times M\setminus N$, where W_1^α and W_2^α are disjoint open sets in M. Let

$$\Omega_\alpha = \{(H_{12}, h_1, h_2)\in C_\alpha | j^1(H_{12}, h_1\times h_2)\pitchfork V = \emptyset\}$$

where

$$C_\alpha = C^\infty(W_1\times W_2, R^{n-i}\times R^{n-i})\times C^\infty(W_1, R^{i-1})\times C^\infty(W_2, R^{i-1}).$$

Apply the jet variant of SMTT(a) with the substitutions:

$$A_1 = W_1^\alpha,\quad X_1 = R^{i-1},\quad A_2 = W_2^\alpha,\quad X_2 = R^{i-1},$$

$$B = W_1^\alpha\times W_2^\alpha,\quad Y = R^{n-i}\times R^{n-i}.$$

We obtain that Ω_α is open and dense in C_α.

The restriction

$$(H, h)\overset{\tau_\alpha}{\longmapsto}(H|_{W_1^\alpha\times W_2^\alpha}, h|_{W_1^\alpha}, h|_{W_2^\alpha})$$

defines a map

$$\tau_\alpha\colon C_N\times C^\infty \to C_\alpha.$$

Since the map τ_α is open, the preimage $\tau_\alpha^{-1}(\Omega_\alpha)$ is open and dense in $C_N\times C^\infty(M, R^{i-1})$ and the same holds for the set $Q=\bigcap_{\alpha=1}^{3}\tau_\alpha^{-1}(\Omega_\alpha)$.

It has remained to prove the following

LEMMA. $\Omega^3=Q$.

PROOF. By definition $(H, h)\in Q$ means that $\{j^1(\tilde F)|_U\}^{-1}(V)=\emptyset$ $(\tilde F=(H, h\times h))$ and $(H, h)\in\Omega^3$ means that $\pi_1|_{A_F\cap U}$ is immersion $(U=M\times M\setminus N)$. So we have to show that

$$\{j^1(\tilde F)|_U\}^{-1}(V)\neq\emptyset \Leftrightarrow \pi_1|_{A_F\cap U}$$

is not an immersion.

($\Leftarrow$) The right side means that there exist

(α) a point $a \in A_F$ and
(β) a non-zero vector v of the tangent space $T_a(A_F)$ such that $d\pi_1(v)=0$.

For the coordinates $(\bar m_1, \bar m_2, \bar y_1, \bar y_2, \bar r)$ of the point $j^1(\tilde F)(a)$ hold:

(α') $\bar y_1 = \bar y_2$ by (α) and
(β') $d\tilde F(v) \in T_{\tilde F(a)}(\Delta(R^{n-1}))$ by (β).

(β') means that the image of the vector v under the differential: $d\tilde F\colon T(M\times M)\to T(R^{n-1}\times R^{n-1})$ is actually contained in $T(\Delta(R^{n-1}))\subset T(R^{n-1}\times R^{n-1})$. Using the canonical decomposition of the tangent space $T_a(M\times M) = T_{\bar m_1}M \oplus T_{\bar m_2}M$ $(a = (\bar m_1, \bar m_2)\in M\times M)$ the vector v can be represented by $v=(v_1, v_2)$, where $v_i \in T_{\bar m_i}M$. By (β) we have $v_1=0$, and so

$$d\tilde F(v) = d\tilde F\begin{pmatrix}0\\ v_2\end{pmatrix} = \begin{pmatrix} \dfrac{\partial y_1}{\partial m_1}(a) & \dfrac{\partial y_1}{\partial m_2}(a) \\ \dfrac{\partial y_2}{\partial m_1}(a) & \dfrac{\partial y_2}{\partial m_2}(a)\end{pmatrix}\begin{pmatrix}0\\ v_2\end{pmatrix} = \begin{pmatrix}\left(\dfrac{\partial y_1}{\partial m_2}(a)\right)\cdot v_2 \\ \left(\dfrac{\partial y_2}{\partial m_2}(a)\right)\cdot v_2\end{pmatrix}.$$

This vector by (β') belongs to $T(\Delta(R^{n-1}))$, i.e.

$$\left(\frac{\partial y_1}{\partial m_2}(a)\right)\cdot v_2 = \left(\frac{\partial y_2}{\partial m_2}(a)\right)\cdot v_2.$$

Hence

$$\operatorname{rank}\left(\frac{\partial y_1}{\partial m_2} - \frac{\partial y_2}{\partial m_2}\right) < m.$$

We obtained that $a \in \{j^1(\tilde F)\}^{-1}(V)$.

($\Rightarrow$) is obtained by reversing the previous argument. ∎

7. A global estimation for small sets of immersion

It is well-known that any immersion is one to one on the small sets. In fact, this statement is true for small families of immersions, too. This section is devoted to a modified variant of this statement.

7.1. Proposition. *Given a map $\varphi \in \Omega^3$ (Ω^3 has been defined in 6.3) there exist a number $\varkappa > 0$ and a neighbourhood G of φ in C_N such that*

(* *) *if $a_1 \in A_\varphi$, $a_2 \in A_\varphi$ and $\varrho(a_1, a_2) < \varkappa$ then $\pi_1(a_1) \neq \pi_1(a_2)$.*

Proof. Consider a finite covering $\{W_1^\alpha \times W_2^\alpha \mid \alpha = 1, \ldots, r\}$ of the set $M\times M \setminus N_1$ where W_1^α and W_2^α are disjoint open sets in M, diffeomorphic to the ball D^m. It will be sufficient to show that for any α $1 \leq \alpha \leq r$ the restriction $\tilde F|_{W_1^\alpha\times W_2^\alpha}$ has a neighbourhood U_α (in the space C_α (see 6.1)) consisting of maps, which satisfy (* *). This will be shown in the following lemma.

7.2. LEMMA. *Let $f\colon D^s\times D^m\to R^q$ be smooth map. Denote $R^p=:\{(x_1,\dots,x_q)\in R^q|x_1=\dots x_{q-p}=0\}$. Let $m+p<q$. Suppose that for any $d\in D^s$ the following assumptions hold:*

(a) *The restriction of f to $d\times D^m$ is transversal to R^p in R^q:*

(b) card $|(d\times D^m)\cap f^{-1}(R^p)|\le 1$.

Then there exists a neighbourhood U of the map f in $C^\infty(D^s\times D^m\to R^q)$ such that if $q\in U$ then g satisfies (b) *(with g instead of f).*

(To prove the lemma in 7.1 it will be sufficient to apply this Lemma in 7.2 with $D^s=W_1^\alpha$, $D^m=W_2^\alpha$, $R^q=R^{n-1}\times R^{n-1}$, $R^p=\Delta(R^{n-1})$, $f=\tilde F$, $g=\varphi$.)

PROOF. Denote the coordinates in D^s, D^m, R^{q-p}, R^p by $d=(d^1,\dots,d^s)$, $x=(x^1,\dots,x^m)$, $y=(y^1,\dots,y^{q-p})$, $z=(z^1,\dots,z^p)$, respectively. The assumption (a) means that for the map

$$f(d,x)=\big(y_f(d,x),\,z_f(d,x)\big)$$

(where for $y_f(d,x)\in R^p$ and $z_f(d,x)\in R^{q-p}$) the following holds: If $y_f(d,x)=0$ then rank $\left(\frac{\partial y_f}{\partial x}(d,x)\right)=m$.

It follows that there exists $\varepsilon>0$ such that if $\|y_f(d,x)\|<\varepsilon$ then rank $\left(\frac{\partial y_f}{\partial x}(d,x)\right)=m$.

The compactness of $D^s\times D^m$ implies that there exists $\delta>0$ such that if $\|y_f(d,x)\|<\varepsilon$ then

$$\left\|\left(\frac{\partial y_f}{\partial x}(d,x)\right)\cdot\bar h\right\|>2\delta\cdot\|\bar h\|\quad\text{for any } \bar h\in R^m.$$

Let $U'(f)$ be a neighbourhood of f such that if $g\in U'(f)$ and $y_g(d,x)=0$ then $y_f(d,x)<\varepsilon$.

Let $U''(f)$ be a neighbourhood of f such that if $g\in U''(f)$ and

$$\left\|\left(\frac{\partial y_f}{\partial x}(d,x)\right)\cdot\bar h\right\|>2\delta\cdot\|\bar h\|$$

then

$$\left\|\left(\frac{\partial y_g}{\partial x}(d,x)\right)\cdot\bar h\right\|>\delta\cdot\|\bar h\|.$$

Put $U_1(f)=U'(f)\cap U''(f)$.

REMARK. If $g\in U_1(f)$ and $y_g(d,x)=0$ then

$$\left\|\left(\frac{\partial y_g}{\partial x}(d,x)\right)\cdot\bar h\right\|>\delta\cdot\|\bar h\|.$$

Indeed,

$$y_g(d,x)=0\Rightarrow y_f(d,x)<\varepsilon\Rightarrow\left\|\left(\frac{\partial y_f}{\partial x}(d,x)\right)\cdot\bar h\right\|>2\delta\cdot\|\bar h\|\Rightarrow$$

$$\Rightarrow\left\|\left(\frac{\partial y_g}{\partial x}(d,x)\right)\cdot\bar h\right\|>\delta\cdot\|\bar h\|.$$

There exists a constant C such that

$$(***) \qquad \left\| \left(\frac{\partial^2 f}{\partial x^i \partial x^j} (d, x) \right) \right\| < C$$

for any $d \in D^s$ and $x \in D^m$. Let $U_2(f)$ be a neighbourhood of the map f such that for any $g \in U_2(f)$ $(***)$ holds (with g instead of f). Put $\varepsilon = \frac{\delta}{C}$ and let W_ε be the $\frac{\varepsilon}{2}$-neighbourhood of the set $f^{-1}(R^p)$ in $D^s \times D^m$. There exists a neighbourhood $U_3(f)$ of f such that for any $g \in U_2(f)$ we have $g^{-1}(R^p) \subset W_\varepsilon$. We show that the neighbourhood $U(f) = U_1(f) \cap U_2(f) \cap U_3(f)$ satisfies the statement of Lemma in 7.2.

Indeed, let $g \in U(f)$ and denote $(y_g(d, x), z_g(d, x))$ its coordinate functions. We shall make use of the Taylor formula

$$y(x+\bar{h}) = y(x) + \left(\frac{\partial y}{\partial x}(x) \right) \bar{h} + \int_0^1 \sum_{i,j} (1-t) \frac{\partial^2 y(x+t\bar{h})}{\partial x^i \partial x^j} h^i h^j \, dt$$

where y stands for y_g and we dropped d in (d, x). Assume in the contrary, that (d, x), $(d, x+\bar{h}) \in g^{-1}(R^p)$. Then $y_g(d, x) = y_g(d, x+\bar{h})$ and we have

$$\delta \|\bar{h}\| < \left\| \left(\frac{\partial y}{\partial x}(x) \right) \bar{h} \right\| = \left\| \int_0^1 (1-t) \frac{\partial^2 y}{\partial x^i \partial x^j} (x+t\bar{h}) h^i h^j \, dt \right\| < C \|\bar{h}\|^2,$$

i.e. $\frac{\delta}{C} = \varepsilon < \|\bar{h}\|$, which contradicts to the fact that $g \in U_3(f)$.

8. Proof of Haefliger's theorem

As it was explained in Section 2, it is sufficient to prove that there exists an isovariant map $F\colon M \times M \to R^n \times R^n$ such that $F = (\tilde{F}, f)$, $f\colon M \times M \to R^1 \times R^1$, $\tilde{F} = (H, h \times h)\colon M \times M \to R^{n-1} \times R^{n-1}$, $H\colon M \times M \to R^{n-i} \times R^{n-i}$ is equivariant, $h\colon M \to R^{i-1}$ and $\pi_1|_{A_F}$ is an embedding of A_F into M (see 2.3). Now we show the existence of such a map F.

(1) Denote D_N the subset of C_N consisting of maps $\varphi = (H, h \times h)\colon M \times M \to R^{n-1} \times R^{n-1}$ such that

$$\pi_1(\varphi^{-1}(\Delta(R^{n-1}) \setminus N)) \cap \Sigma^1(\tilde{g}_\varphi) = \emptyset.$$

(Recall that $\varphi|_N$ coincides with $\tilde{g}_\varphi \times \tilde{g}_\varphi|_N$ where $\tilde{g}_\varphi\colon M \to R^{n-1}$ is an S-map. $\Sigma^1(\tilde{g}_\varphi)$ denotes the (Σ^1-type) singular point set of the map $\tilde{g}_\varphi$.)

Then D_N is open and dense in C_N.

(2) There exist
(i) neighbourhoods $N_1 \subset N_2 \subsetneqq N$ of $\Delta(M)$ in $M \times M$ and
(ii) an open subset $\mathcal{U}$ of D_N such that

$$\varphi \in \mathcal{U} \Rightarrow \pi_1(A_\varphi \setminus N_2) \cap \pi_1(A_\varphi \cap N_1) = \emptyset.$$

(3) In C_N there exists a dense subset $\mathscr{V}$ consisting of maps φ such that $\pi_1|_{A_\varphi \setminus N_1}$ is one-to-one.

(4) The intersection $\mathscr{V} \cap \mathscr{U}$ is non-empty and any element of this intersection satisfies assumption (*) in 2.3.

PROOF.

(1) To prove (1) we have to recall only that $\Sigma^1(\tilde{g}_\varphi)$ is a submanifold (in M) of codimension $n-m$, $\Delta(R^{n-1})$ is a submanifold of $R^{n-1} \times R^{n-1}$ of codimension $n-1$ and $(n-1)+(n-m) > 2m = \dim M \times M$.

(2) and (4) are obvious.

(3) By 6.1 and 7.1 for any $\psi \in \Omega^1$ there exists a neighbourhood G of ψ in C_N and a number $\varkappa$ such that if $\varphi \in G$, $a_1 \in A_\varphi$, $a_2 \in A_\varphi$ and $\varrho(a_1, a_2) < \varkappa$ then $\pi_1(a_1) \neq \pi_1(a_2)$. By 6.2 the set

$$\Omega^2(N_2, N_1; \varkappa) = \{\varphi \in C_N \text{ if } a_1 \in A_\varphi \setminus N_1,\ a_2 \in A_\varphi \setminus N_2 \text{ and } \varrho(a_1, a_2) > \varkappa \text{ then } \pi_1(a_1) \neq \pi_1(a_2)\}$$

is dense in C_N.

Hence $G \cap \Omega^2(N_2, N_1;\ \varkappa)$ is not empty and any element of this intersection satisfies 3 (i.e. belongs to $\mathscr{V}$) if $\varkappa$ is sufficiently small. Therefore in any neighbourhood of any element of the dense subset Ω^1 of C_N there exists an element of $\mathscr{V}$. Hence $\mathscr{V}$ is dense in C_N. ∎

REFERENCES

[1] HAEFLIGER, A., Plongements différentiables dans le domaine stable, *Comment. Math. Helv.* **37** (1962—63), 155—176. *MR* **28** # 625.

[2] WHITNEY, H., The self-intersections of a smooth n-manifold in $2n$-space, *Ann. of Math.* (2) **45** (1944), 220—246. *MR* **5**—273.

[3] HAEFLIGER, A. and HIRSCH, M. W., Immersions in the stable range, *Ann. of Math.* (2) **75** (1962), 231—241. *MR* **26** # 784.

[4] HIRSCH, M. W., Immersions of manifolds, *Trans. Amer. Math. Soc.* **93** (1959), 242—276. *MR* **22** # 9980.

[5] HAEFLIGER, A., Plongements différentiables de variétés dans variétés, *Comment. Math. Helv.* **36** (1961), 47—82. *MR* **26** # 3069.

[6] THOM, R., Un lemme sur les applications différentiables, *Bol. Soc. Mat. Mexicana* (2) **1** (1956), 59—71. *MR* **21** # 910.

[7] SZŰCS, A., On detached immersions, *Studia Sci. Math. Hungar.* **18** (1983) (to appear).

(Received April 16, 1981)

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

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 319—323

AN OPTIMAL GROUP-TESTING PROCEDURE

RATKO TOŠIĆ

1. The problem and some preliminaries

Consider the following problem:

There are exactly two odd elements, one defective and one mediocre, in the set $X=\{e_1, e_2, ..., e_n\}$ $(n\geqq 2)$, all possibilities occurring with equal probabilities. We want to identify both odd elements (defective and mediocre) by testing some subsets A of X, and for each such set A determining whether A contains some of them. A subset A of X is said to be defective if it contains the defective element, mediocre if it contains the mediocre element but not the defective one, otherwise we say that A is good. The test on an individual subset A informs us whether A is defective, mediocre or a good set. Note that without additional information we do not know whether a defective subset of X contains the mediocre element or not. Our aim is to find an optimal procedure to identify both odd elements, i.e. to minimize the maximum number of tests.

This question was posed by Katona ([2] p. 306, Problem 13). In this paper we investigate a somewhat modified problem. Namely, we suppose that after using a subset A the subsequent tests may check only the subsets of A and $\bar{A}$.

We denote by $\mu(n)$ the maximal test length of an optimal procedure for the modified problem. In this paper we obtain the value of this function for an infinite sequence of values of n. For all other integers greater than 2, we determine a lower bound and an upper bound for $\mu(n)$ which differ by just one unit. The corresponding procedures are constructed inductively. Of course, the constructive part of our proof gives an upper estimate on the original problem of Katona, but it is very likely too rough.

The problem considered in this paper fall under the general heading of dynamic programming. For some discussion of these matters in greater detail, see Bellman [1], Katona [2] and Sobel [3, 4].

Let $P_n(l)$ denote any procedure which enables us to identify both odd elements in the set X $(|X|=n)$ using maximally l tests. $P_n(l)$ is optimal (i.e. $\mu(n)=l$) if there does not exist a procedure $P_n(l_1)$ with $l_1<l$.

Similarly, $P_n^1(r)$ denotes any procedure which enables us to identify the odd element in the set of n elements, given information that there is exactly one of them present, the maximum test length being r. It is well-known that $P_n^1(r)$ is optimal if and only if $2^{r-1}<n\leqq 2^r$.

1980 *Mathematics Subject Classification.* Primary 90B40, 62C20; Secondary 68H05.

Key words and phrases. Optimal strategy, defective element, mediocre defective element, test length.

2. The results

LEMMA. *Let* $n_1 < n$ $(n \geqq 3)$ *and suppose that there is a procedure* $P_n(l)$. *Then there exists a procedure* $P_{n_1}(l_1)$ *such that* $l_1 \leqq l$.

PROOF. Let $X_1 = \{e_1, e_2, \ldots, e_{n_1}\}$ contain exactly one defective and exactly one mediocre element. Consider $X = \{e_1, e_2, \ldots, e_{n_1}, \ldots, e_n\}$ for which a procedure $P_{|X|}(l) = P_n(l)$ exists. Now, a procedure $P_{|X_1|}(l_1) = P_{n_1}(l_1)$ can be constructed by imitating $P_{|X|}(l)$, testing $A \cap X_1$ instead of A at the corresponding place. Some tests may occur to be fictitious. Note that the empty set is a good set. ▮

We denote by $P_n(\leqq l)$ any procedure which enables us to identify both odd elements in the set of n elements, the maximal test length being at most l.

THEOREM. *Let*

$$t_k = F_{\left\lfloor \frac{k}{2} \right\rfloor} + 2^{\left\lfloor \frac{k}{2} \right\rfloor} + (1 + (-1)^{k+1}) 2^{\frac{k-5}{2}} \tag{1}$$

for $k = 2, 3, \ldots$, *where* F_j *is the jth member of the Fibonacci sequence*

$$F_1 = 1, \quad F_2 = 1; \quad F_j = F_{j-1} + F_{j-2}, \qquad j = 3, 4, \ldots, \tag{2}$$

then $\mu(t_k) = k$.

PROOF. The sequence (1) can be written in the form

$$t_{2m} = F_m + 2^m, \quad t_{2m+1} = F_m + 2^m + 2^{m-1}, \quad m = 1, 2, \ldots \tag{3}$$

and the following relations can be easily verified:

$$F_m \leqq 2^{m-2}, \qquad m = 2, 3, \ldots \tag{4}$$

$$2^m < t_{2m} \leqq 2^{m+1}, \qquad m = 1, 2, \ldots \tag{5}$$

$$2^m < t_{2m+1} \leqq 2^{m+1}, \qquad m = 1, 2, \ldots \tag{6}$$

$$t_{2m} = t_{2m-1} + t_{2m-4}, \qquad m = 3, 4, \ldots. \tag{7}$$

In order to prove the theorem, we shall prove that for $k = 2, 3, \ldots$

(i) a procedure $P_{t_k}(k)$ can be constructed, and

(ii) $P_{t_k}(k)$ is optimal, i.e. $P_{t_k}(l)$ does not exist if $l < k$. The proof uses mathematical induction.

(i) *Existence and construction*

For the first three members of the sequence (1) $t_2 = 3$, $t_3 = 4$, $t_4 = 5$, the statement is true. In these cases $P_{t_k}(k)$ is the "element by element" procedure.

Suppose that $P_{t_k}(k)$ can be constructed for each $k \leqq 2m$ $(k \geqq 2, m \geqq 2)$. Then $P_{t_{2m+1}}(2m+1)$ and $P_{t_{2m+2}}(2m+2)$ can be constructed according to the following schemes:

(A) Construction of $P_{t_{2m+1}}(2m+1)$.

Step A1. Test $A = \{e_1, e_2, \ldots, e_{2^{m-1}}\}$ and go to Step A2.

Step A2. If A is good (it means that both odd elements are in $X\setminus A$, where $|X\setminus A|=t_{2m+1}-2^{m-1}=t_{2m}$), apply $P_{t_{2m}}(2m)$ to the set $X\setminus A$. This procedure can be constructed by the induction hypothesis.

If A is mediocre (it means that the mediocre element is in A while the defective one is in $X\setminus A$, where $|A|=2^{m-1}$ and $|X\setminus A|=t_{2m}$), continue by applying two independent procedures $P^1_{2^{m-1}}(m-1)$ and $P^1_{t_{2m}}(m+1)$ to the sets A and $X\setminus A$, respectively. The estimate of the maximal test length of the latter procedure follows from (5).

If A is defective, test $A_2=\{e_{2^{m-1}+1}, \ldots, e_{2^{m-1}+2^m}\}$ ($|A_2|=2^m$) and go to Step A3. (Note that A_2 must be either good or mediocre because the defective element is in A.)

Step A3. If A_2 is mediocre (it means that the mediocre element is in A_2 while the defective is in A), continue by applying two independent procedures $P^1_{2^{m-1}}(m-1)$ and $P^1_{2^m}(m)$.

If A_2 is good, test $A_{20}=\{e_{2^{m-1}+2^m+1}, \ldots, e_{t_{2m+1}}\}$ ($|A_{20}|=F_m$) and go to Step A4. (A_{20} must be either good or mediocre.)

Step A4. If A_{20} is mediocre (we conclude that the mediocre element is in A_{20} while the defective is in A), continue by applying two independent procedures $P^1_{2^{m-1}}(m-1)$ and $P^1_{F_m}(\leq m-2)$ to the sets A and A_{20}, respectively. The estimate of the maximal test length of the latter procedure follows from (4).

If A_{20} is good (it means that both odd elements are in A), apply $P_{2^{m-1}}(2m-2)$ to A. This procedure can be constructed by the induction hypothesis (see (5)).

So, the procedure $P_{t_{2m+1}}(2m+1)$ is constructed.

(B) Construction of $P_{t_{2m+2}}(2m+2)$.

Step B1. Test $B=\{e_1, e_2, \ldots, e_{t_{2m-2}}\}$ and go to Step B2.

Step B2. If B is good (we conclude that both odd elements are in $X\setminus B$, and it follows from (7) that $|X\setminus B|=t_{2m+1}$), apply to $X\setminus B$ the procedure $P_{t_{2m+1}}(2m+1)$ just constructed in Section (A).

If B is mediocre (it means that the mediocre element is in B and the defective is in $X\setminus B$), continue by applying the independent procedures $P^1_{t_{2m-2}}(m)$ to B and $P^1_{t_{2m+1}}(m+1)$ to $X\setminus B$. The estimates of the maximal test lengths follow from (5) and (6).

If B is defective, test $B_2=\{e_{t_{2m-2}+1}, \ldots, e_{t_{2m-2}+2^m}\}$ ($|B_2|=2^m$) and go to Step B3. (B_2 must be either good or mediocre.)

Step B3. If B_2 is mediocre, continue by applying the independent procedures $P^1_{t_{2m-2}}(m)$ to B and $P^1_{2^m}(m)$ to B_2.

If B_2 is good, test

$$B_{20} = \{e_{t_{2m-2}+2^m+1}, \ldots, e_{t_{2m-2}+2^m+2^{m-1}}\} \qquad (|B_{20}| = 2^{m-1})$$

and go to Step B4. (B_{20} must be either good or mediocre.)

Step B4. If B_{20} is mediocre, apply $P^1_{t_{2m-2}}(m)$ to B and $P^1_{2^{m-1}}(m-1)$ to B_{20}.

If B_{20} is good, test

$$B_{200} = \{e_{t_{2m-2}+2^m+2^{m-1}+1}, \ldots, e_{t_{2m+2}}\} \qquad (|B_{200}| = F_m)$$

and go to Step B5.

Step B5. If B_{200} is mediocre, apply $P^1_{t_{2m-2}}(m)$ to B and $P^1_{F_m}(\leqq m-2)$ to B_{200}. If B_{200} is good (it means that both odd elements are in B), continue by applying $P_{t_{2m-2}}(2m-2)$ which can be constructed by the induction hypothesis.

So, the procedure $P_{t_{2m+2}}(2m+2)$ is constructed.

(ii) *Optimality*

To prove the optimality of $P_{t_k}(k)$, it suffices to prove that there does not exist a procedure $P_{t_k}(l)$, where $l<k$ $(k=2, 3, \ldots)$.

It follows by information-theoretical reasonings that the existence of $P_n(l)$ implies its optimality provided $n(n-1)>3^{l-1}$. So, $P_{t_2}(2)=P_3(2)$ and $P_{t_3}(3)= =P_4(3)$ are optimal because $3\cdot 2=6>3^1=3$ and $4\cdot 3=12>3^2=9$. However, $t_k(t_k-1)<3^{k-1}$ for all $k>3$. So, in order to prove the optimality of $P_{t_4}(4)=P_5(4)$, we must analyse it in detail.

Let $C\subset X=\{e_1, e_2, e_3, e_4, e_5\}$ be the set tested in the first step. Consider the following cases:

(1) $|C|=1$. If C is identified as a good set, it means that both odd elements are in $X\setminus C$, where $|X\setminus C|=4$, and now we cannot find the odd elements in $X\setminus C$ by less than three tests.

(2) $|C|=2$. If C is identified as a mediocre set, we need at least three additional tests, because we must apply $P^1_2(1)$ to C and $P^1_3(2)$ to $X\setminus C$.

(3) $|C|=3$. If C is identified as a mediocre set, we must apply $P^1_3(2)$ to C and $P^1_2(1)$ to $X\setminus C$.

(4) $|C|=4$. If C is identified as a defective set, it is possible that the mediocre element is in C, too, and we must accomplish at least three more tests to identify both odd elements.

So, in any case; we need to accomplish at least four tests. It means that $P_5(4)$ is optimal. The basis of mathematical induction is established.

Suppose that the optimality of $P_{t_k}(k)$ is proved for all $k\leqq 2m$ $(k\geqq 2, m\geqq 2)$. We are going to prove that, under these conditions, $P_{t_{2m+1}}(2m+1)$ and $P_{t_{2m+2}}(2m+2)$ are optimal, too.

(A′) Optimality of $P_{t_{2m+1}}(2m+1)$.

Let $D\subset X$ $(|X|=t_{2m+1})$ be tested in the first step of a procedure $P_{|X|}(l)$. Consider the following cases:

(1) $|D|\leqq 2^{m-1}$. Then $|X\setminus D|\geqq t_{2m}$, and if D is identified as a good set, then by the induction hypothesis, we must accomplish at least $2m$ additional tests to identify both odd elements in $X\setminus D$.

(2) $2^{m-1}<|D|<t_{2m}$. Then $2^{m-1}<|X\setminus D|<t_{2m}$, and if D is identified as a mediocre set, we must apply $P^1_{|D|}$ $(\geqq m)$ to identify the mediocre element in D and $P^1_{|X\setminus D|}(\geqq m)$ to identify the defective element in $X\setminus D$. Thus, we need at least $2m$ additional tests, i.e. at least $2m+1$ tests altogether.

(3) $|D|\geqq t_{2m}$. If D is identified as a defective set then it is possible that the mediocre element is in D, too. By the induction hypothesis, we must accomplish at least $2m$ additional tests to identify both odd elements in D.

Thus, we have proved that $P_{t_{2m+1}}(l)$ does not exist if $l<2m+1$, i.e. $P_{t_{2m+1}}(2m+1)$ is optimal.

(B′) Optimality of $P_{t_{2m+2}}(2m+2)$.

Let $E\subset X$ ($|X|=t_{2m+2}$) be tested in the first step of a procedure $P_{|X|}(l)$. Consider the following possibilities:

(1) $|E|\leqq t_{2m-2}$. Then $|X\setminus E|\geqq t_{2m+1}$, and if E is identified as a good set, then according to (A′), we must accomplish at least $2m+1$ additional tests to identify both odd elements in $X\setminus E$.

(2) $t_{2m-2}<|E|\leqq 2^m$. Then $F_{m+1}+2^m\leqq|X\setminus E|<t_{2m+1}$, and if E is identified as a mediocre set, we must apply $P^1_{|E|}(m)$ to identify the mediocre element (see (5)) and $P^1_{|X\setminus E|}(m+1)$ to identify the defective element in $X\setminus E$.

(3) $2^m<|E|<t_{2m+1}$. Then $t_{2m-2}<|X\setminus E|<F_{m+1}+2^m$, and if E is identified as a mediocre set, we must apply $P^1_{|E|}(m+1)$ and $P^1_{|X\setminus E|}(m)$ (see (5)).

(4) $|E|\geqq t_{2m+1}$. If E is identified as a defective set, then it is possible that the mediocre element is in E, too. In that case, according to (A′), we must accomplish at least $2m+1$ additional tests to identify both odd elements in E.

Thus, we have proved that $P_{t_{2m+2}}(l)$ does not exist if $l<2m+2$. It means that $P_{t_{2m+2}}(2m+2)$ is optimal. ∎

COROLLARY. *If* $t_k<n\leqq t_{k+1}$, *then* $k\leqq\mu(n)\leqq k+1$.

PROOF. The proof follows immediately from the Lemma and the Theorem. ∎

It means that, if n is not a member of (1), then the procedure $P_n(l)$ constructed according to the Theorem and the Lemma is either an optimal or an almost optimal procedure, i.e. such that $(l-\mu(n))\leqq 1$.

REFERENCES

[1] BELLMAN, R., *Dynamic programming*, Princeton University Press, Princeton, N. J., 1957. *MR* **19**—820.

[2] KATONA, G. O. H., Combinatorial search problems, *Survey of Combinatorial Theory* (Proc. Internat. Sympos., Colorado State Univ., Fort Collins, Colo., 1970), edited by J. N. Srivastava, North-Holland, Amsterdam, 1973, 285—308. *MR* **51** # 4976.

[3] SOBEL, M., Group testing to classify efficiently all units in a binomial sample, *Information and Decision Processes*, ed. by R. E. Machil, McGraw-Hill, New York, 1960, 127—161. *MR* **24** # A2506.

[4] SOBEL, M., Binomial and hypergeometric group-testing, *Studia Sci. Math. Hungar.* **3** (1968), 19—42. *MR* **38** # 5325.

[5] TOŠIĆ, R., An optimal search procedure, *J. Statist. Plann. Inference.* **4** (1980), 169—171. *MR* **82b** : 62131.

(Received April 21, 1981)

INSTITUTE OF MATHEMATICS
UNIVERSITY OF NOVI SAD
YU—21 001 NOVI SAD
YUGOSLAVIA

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 325—330

UN THÉORÈME SUR LES SYSTÈMES LINÉAIRES DE QUADRIQUES À JACOBIENNE INDÉTERMINÉE

LANDO DEGOLI

On démontre une condition nécessaire et suffisante pour qu'un système linéaire de quadriques de S_r soit à matrice Jacobienne identiquement nulle en supposant d'abord que la caractéristique de la Jacobienne soit r et après qu'elle soit $r-k$.

Dans l'espace complexe linéaire S_r, rapporté aux coordonnées projectives homogènes x_i $(i=0, 1, 2, \ldots, r)$ un système linéaire L_d $(d\geqq r)$ de quadriques linéairement indépendantes est exprimé par l'équation:

$$\lambda^0 f_0+\lambda^1 f_1+\lambda^2 f_2+\ldots+\lambda^d f_d = 0$$

avec

$$f_q = a_q^{ik} x_i x_k,$$

où le symbole de sommation est sous-entendu pour l'écriture abrégée désormais usuelle.

Prenons en considération la matrice Jacobienne à $d+1$ lignes et $r+1$ colonnes:

$$J = \left\| \frac{\partial f_i}{\partial x_s} \right\| \qquad \begin{pmatrix} i = 0, 1, 2, \ldots, d \\ s = 0, 1, 2, \ldots, r \end{pmatrix}. \tag{1}$$

En général la matrice Jacobienne égalisée à zéro est le lieu géométrique des points de S_r conjugués entre eux-mêmes par rapport à toutes les quadriques du système. Si la matrice Jacobienne est identiquement nulle, cela signifie que tout l'espace est le lieu de points conjugués.

Si la caractéristique de la Jacobienne est r, un point générique de S_r est conjugué avec un seul point. Si au contraire la caractéristique est $r-h$ avec $h>0$ un point quelconque de S_r est conjugué avec un espace S_h.

THÉORÈME. *Considérons un système linéaire de quadriques L_d $(d\geqq r)$ qui ne possède aucun système subordonné L_g $(2\leqq g\leqq d-1)$ possédant une Jacobienne identiquement nulle de caractéristique c: $c\leqq g$; $2\leqq c\leqq r-k-1$ $(k>0)$. La condition nécessaire et suffisante pour que le système linéaire de quadriques L_d $(d\geqq r)$, satisfaisant aux prémisses précédentes, ait la Jacobienne identiquement nulle de caractéristique $r-k$ $(k\geqq 0)$, c'est que les quadriques du système qui passent par un point quelconque de S_r possèdent en commun un S_{k+1}.*

1980 *Mathematics Subject Classification.* Primary 51N15.
Key words and phrases. Quadric, linear system of quadrics.

Avant tout, nous démontrons le cas particulier:

Si le système L_d est à Jacobienne identiquement nulle de caractéristique r, les quadriques du système qui passent par un point ont en commun une droite.

Si la Jacobienne est identiquement nulle de caractéristique r cela signifie que tous les déterminants de la matrice (1) d'ordre $m+1$ sont identiquement nuls.

Considérons le déterminant donné par $r+1$ quadriques quelconques. En effet, on peut choisir les premières r quadriques du système:

$$f_0, f_1, f_2, \ldots, f_{r-1}, f_r.$$

On aura:

$$D = \begin{vmatrix} \frac{\partial f_0}{\partial x_0} & \frac{\partial f_1}{\partial x_0} & \cdots & \frac{\partial f_r}{\partial x_0} \\ \frac{\partial f_0}{\partial x_1} & \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_r}{\partial x_1} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_0}{\partial x_r} & \frac{\partial f_1}{\partial x_r} & \cdots & \frac{\partial f_r}{\partial x_r} \end{vmatrix} = 0. \tag{2}$$

Les mineurs d'ordre r extraits du déterminant D ne peuvent pas être tous nuls, autrement le système linéaire L_r déterminé par les $r+1$ quadriques précédentes aurait la Jacobienne identiquement nulle de caractéristique $<r$, et pour cela il existerait dans L_d un système subordonné avec la Jacobienne identiquement nulle de caractéristique inférieure à r contre l'hypothèse.

Il est donc nécessaire qu'au moins l'un des déterminants d'ordre r soit différent de zéro. Nous pouvons supposer qu'il soit le mineur obtenu en éliminant la dernière ligne et la dernière colonne. Nous l'indiquerons par A:

$$A = \begin{vmatrix} \frac{\partial f_0}{\partial x_0} & \frac{\partial f_1}{\partial x_0} & \cdots & \frac{\partial f_{r-1}}{\partial x_0} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_0}{\partial x_{r-1}} & \frac{\partial f_1}{\partial x_{r-1}} & \cdots & \frac{\partial f_{r-1}}{\partial x_{r-1}} \end{vmatrix}.$$

Prenons en considération la matrice extraite du déterminant D formée avec les premières r ligne et indiquons avec:

$$A_0, A_1, A_2, \ldots, A_{r-1}$$

les mineurs d'ordre r que l'on obtient en substituant à la première, deuxième, etc. colonne de A la dernière colonne de la matrice. Un seul de ces déterminants peut être identiquement nul, parce que, en supposant que deux déterminants, par exemple A_0 et A_1, soient identiquement nuls un point de S_r, qui n'annulle pas tous les mineurs d'ordre $r-1$ communs à A_0 et A_1 et pour le reste générique, rendrait nul A. (Voire: L. Kronecker, Werke I^0, Leipzig, 1885, p. 238.) Mais A est une forme d'ordre r dans les x_i et pour cela serait identiquement nul, ce qui est impossible.

Puisque le déterminant D est identiquement nul, les quadriques, tandis qu'elles sont linéairement indépendantes, elles deviennent aussi fonctionnellement dépendantes et, une quelconque, par exemple f_r, sera fonction des autres. Il sera:

$$f_r = F(f_0, f_1, f_2, \ldots, f_{r-1}). \tag{3}$$

Cette relation est exacte pour tous les groupes de $r+1$ quadriques choisis entre L_d, mais il n'est pas possible que des quadriques en nombre $< r+1$ soient fonctionnellement dépendantes entre elles, autrement il existerait dans L_d des systèmes subordonnés contre les prémisses au théorème.

Mais parce que les f_i sont toutes des formes de deuxième ordre en remplaçant au lieu de $x_0, x_1, \ldots, x_r$: $tx_0, tx_1, \ldots, tx_r$, cela donne:

$$t^2 f_r = F(t^2 f_0, t^2 f_1, \ldots, t^2 f_{r-1})$$

qui montre que F est une fonction homogène de premier degré.

En dérivant l'expression (3) on obtient:

$$\begin{gathered} \frac{\partial F}{\partial f_0}\frac{\partial f_0}{\partial x_0}+\frac{\partial F}{\partial f_1}\frac{\partial f_1}{\partial x_0}+\ldots+\frac{\partial F}{\partial f_{r-1}}\frac{\partial f_{r-1}}{\partial x_0}=\frac{\partial f_r}{\partial x_0} \\ \cdot \qquad \cdot \qquad \cdots \qquad \cdot \qquad \cdot \\ \frac{\partial F}{\partial f_0}\frac{\partial f_0}{\partial x_{r-1}}+\frac{\partial F}{\partial f_1}\frac{\partial f_1}{\partial x_{r-1}}+\ldots+\frac{\partial F}{\partial f_{r-1}}\frac{\partial f_{r-1}}{\partial x_{r-1}}=\frac{\partial f_r}{\partial x_{r-1}}, \end{gathered} \tag{4}$$

système de premier degré, qui donne aisément les derivées partielles de la fonction F; c'est-à-dire:

$$\frac{\partial F}{\partial f_0} = -\frac{A_0}{A}, \quad \frac{\partial F}{\partial f_1} = -\frac{A_1}{A}, \ldots, \frac{\partial F}{\partial f_{r-1}} = -\frac{A_{r-1}}{A}. \tag{5}$$

Considérons un point x de S_r de coordonnées: $x_0, x_1, \ldots, x_r$, et soit x' de coordonnées: $x'_0, x'_1, \ldots, x'_r$ son conjugué par rapport à toutes les quadriques du système.

La droite qui joint les deux points sera donnée par:

$$y_i = t_1 x_i + t_2 x'_i \qquad (i = 0, 1, 2, \ldots, r). \tag{6}$$

En remplaçant (6) dans toutes les quadriques, on obtient pour la quadrique générique f_m:

$$f_m(y) = f_m(x)t_1^2 + f_m(x')t_2^2 \qquad (m = 0, 1, 2, \ldots, r) \tag{7}$$

parce que les termes $2a_m^{ik}x_i x'_k$ sont nuls les points x et x' étant conjugués.

En remplaçant (7) dans (3) et aussitôt en dérivant par rapport à t_1 et t_2 on obtient:

$$\begin{gathered} \frac{\partial f_r}{\partial t_1} = \frac{\partial F}{\partial f_s}\frac{\partial f_s}{\partial t_1} \\ \frac{\partial f_r}{\partial t_2} = \frac{\partial F}{\partial f_s}\frac{\partial f_s}{\partial t_2} \end{gathered} \qquad (s = 0, 1, 2, \ldots, r-1). \tag{8}$$

En dérivant (7) on a :

$$\frac{\partial f_m}{\partial t_1} = 2t_1 f_m(x) \qquad \frac{\partial f_m}{\partial t_2} = 2t_2 f_m(x') \qquad (m = 0, 1, 2, \ldots, r).$$

Remplaçons ces dernières dans (8). On obtient :

$$f_r(x) = \frac{\partial F}{\partial f_s} f_s(x) \qquad f_r(x') = \frac{\partial F}{\partial f_s} f_s(x') \qquad (s = 0, 1, 2, \ldots, r-1).$$

Et enfin pour (5) :

(9)
$$A_0 f_0(x) + A_1 f_1(x) + \ldots + A_{r-1} f_{r-1}(x) + A f_r(x) = 0$$
$$A_0 f_0(x') + A_1 f_1(x') + \ldots + A_{r-1} f_{r-1}(x') + A f_r(x') = 0.$$

Les déterminants $A, A_0, A_1, \ldots, A_{r-1}$ sont calculés dans un point générique de S_r par exemple en x et un seul d'entre eux est au maximum nul.

Les expressions précédentes sont identiquement nulles par rapport à t_1 et t_2. Ces variables se trouvent seulement dans les déterminants $A, A_0, A_1, \ldots, A_{r-1}$, qui résultent des fonctions homogènes de t_1 et t_2 et un seul d'entre eux est au maximum nul.

Parce que les deux identités coexistent il faut qu'il soit :

$$f_m(x) = c f_m(x') \qquad (m = 0, 1, 2, \ldots, r)$$

avec c constante pas nulle.

En effet nous pouvons donner au rapport t_1/t_2 infinis valeurs quelconques et, en particulier, r valeurs diverses et par conséquent obtenir deux systèmes algébriques de premier degré aux r équations et r inconnues. Ces dernières sont respectivement les quotients des $f_k(x)$, $f_k(x')$ par rapport à une quelconque d'entre elles, par exemple $f_r(x)$ et $f_r(x')$ c'est-à-dire :

(10) $$f_k(x)/f_r(x), \quad f_k(x')/f_r(x') \qquad (k = 0, 1, 2, \ldots, r-1).$$

Puisque les deux systèmes algébriques ont les mêmes coefficients constants $A, A_0, A_1, \ldots, A_{r-1}$ la solution des deux systèmes est la même.

Il en résulte :

(11) $$\frac{f_k(x)}{f_r(x)} = \frac{f_k(x')}{f_r(x')} \qquad (k = 0, 1, 2, \ldots, r-1).$$

Mais cette égalité est vérifiée seulement si :

(12) $$f_m(x) = c f_m(x'), \qquad (m = 0, 1, 2, \ldots, r)$$

comme nous l'avons signalé.

On en déduit que toutes les quadriques du système L_d qui passent par un point x passent aussi par son conjugué x' et réciproquement. Si x est situé sur la quadrique

f_m, on aura:

$$f_m(x) = 0$$

et pour (12):

$$f_m(x') = 0,$$

en remarquant que x' est le conjugué de x par rapport à toutes les quadriques de L_d. Il en résulte:

$$t_1^2 f_m(x) + t_2^2 f_m(x') = 0$$

et pour (7):

$$f_m(y) = 0$$

où y est le point générique de la droite xx'.

La droite en question appartient donc tout entière à la quadrique f_m. On en déduit que toutes les quadriques qui passent par x contiennent la droite xx'.

Supposons maintenant que la Jacobienne du système L_d soit identiquement nulle de caractéristique $r-k$.

Cela signifie qu'un point x de S_r a pour conjugué un S_k.

Considérons un générique S_{r-k} qui passe par x. Le système L_d sera entrecoupé par S_{r-k} suivant un système linéaire L'_d de quadriques de S_{r-k}, qui à son tour coupera S_k dans un point x', qui résulte le conjugué de x par rapport à toutes les quadriques du système L'_d.

Nous pourrons choisir pour coordonnées de S_{r-k} les $x_0, x_1, \ldots, x_{r-k}$, en annulant toutes les autres coordonnées, c'est-à-dire en écrivant:

$$x_{r-k+1} = x_{r-k+2} = \ldots = x_r = 0.$$

Les équations des quadriques $f_0, f_1, \ldots, f_d$ seront du type:

$$f_i(x_0, x_1, \ldots, x_{r-k}, 0, 0, 0, \ldots, 0) = 0.$$

Les dérivées partielles:

$$\frac{\partial f_i}{\partial x_s} \qquad (i = 0, 1, 2, \ldots, d)$$

pour $x_{r-k+1} = x_{r-k+2} = \ldots = x_r = 0$, seront toutes nulles. La matrice Jacobienne du système L'_d:

$$\left\| \frac{\partial f_i}{\partial x_s} \right\| \qquad \begin{pmatrix} i = 0, 1, 2, \ldots, d \\ s = 0, 1, 2, \ldots, r-k \end{pmatrix}$$

sera identiquement nulle.

Elle ne pourra pas avoir de caractéristique supérieure à $r-k$, parce que ses lignes ne sont qu'en nombre $r-k+1$; elle ne pourra pas avoir de caractéristique inférieure à $r-k$, sinon le point x aurait pour conjugué un S_g avec $g > 0$ et non le seul point x'.

Il en résulte que le système L'_d de S_{r-k} a la caractéristique $r-k$. Cela porte à conclure pour la première partie du théorème, que les quadriques de L'_d qui passent par x auront en commun la droite xx'.

Puisque nous pouvons dire la même chose pour tous les S_{r-k} qui passent par x, on en déduit que les quadriques de L_d qui passent par x auront en commun S_{k+1} joignant le point x avec S_k.

La condition du théorème est donc nécessaire.

Elle est suffisante aussi. En effet, si toutes les quadriques de L_d, qui ont en commun un point x, ont en commun un S_{k+1} il est évident que le point x a pour conjugué le même S_{k+1} par rapport au système L_{d-1} de quadriques qui passent par x. Une autre quadrique du système L_d qui ne passe pas par x a pour conjugué de x un hyperplan qui coupera le S_{k+1} dans un S_k et il en résulte que x aura pour conjugué par rapport au système L_d un S_k.

Cela signifie que la matrice Jacobienne est identiquement nulle de caractéristique $r-k$, comme on voulait démontrer.

Remarque. Les systèmes linéaires de quadriques à Jacobienne identiquement nulle de caractéristique $r-k$, qui possèdent des systèmes subordonnés à Jacobienne identiquement nulle de caractéristique inférieure, ne satisfont pas au théorème.

Par exemple, considérons un système L'_d avec $d \geqq r+1$ à Jacobienne identiquement nulle de caractéristique r, qui possède un système subordonné L_{d-1} à Jacobienne identiquement nulle de caractéristique $r-1$. Les quadriques de L_{d-1} qui passent par un point x ont pour le théorème démontré un plan en commun.

Alors une quadrique ultérieure qui n'appartienne pas à L_{d-1} et qui passe par le point x sera coupée par le plan dans une conique et, par conséquent, les quadriques de L_d qui passent par x ont en commun une conique.

RÉFÉRENCES

[1] Terracini, A., Alcune questioni sugli spazi tangenti e osculatori ad una varietà, *Atti R. Accad. Sci. Torino,* Nota II, **51** (1916); Nota III, **55** (1919—20).

[2] Bonferroni, C., Sui sistemi lineari di quadriche la cui Jacobiana ha dimensione irregolare, *R. Accad. Sa. Torino* **50** (1914—15).

[3] Muracchini, L., Le varietà V_5 i cui spazi tangenti ricoprono una varietà W di dimensione inferiore alla ordinaria, II, *Rivista Mat. Univ. Parma* **3** (1952), 75—89. *MR* **14**—316.

[4] Degoli, L., Sui sistemi lineari di quadriche di S_r a Jacobiana identicamente nulla di caratteristica $\leqq r$, *Atti Accad. Sci. Ist. Bologna Cl. Sci. Fis. Rend.* (11) **10** (1962/63), no. 2, 224—234. *MR* **31** #1592.

(Reçu le 25 avril 1981)

VIA BERENGARIO 82/C
I—41012 CARPI
ITALY

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 331—333

SOME RAMSEY NUMBERS FOR FAMILIES OF CYCLES

BRIAN ALSPACH* and KATHERINE HEINRICH

In this note we show that the edges of K_{p-1}, $p=2\left\lceil\frac{3k}{2}\right\rceil$, can be coloured from a set of k colours so that no monochromatic subgraph contains an even cycle, whereas such a colouring is not possible for K_p.

Throughout this note we use the graph-theoretic terminology of Bondy and Murty [1] and the Ramsey theory notation of Parsons [4]. We denote by $\mathscr{C}_0$ and $\mathscr{C}_e$ the families of graphs consisting of all odd and all even cycles, respectively. Then the family $\mathscr{C}=\mathscr{C}_0\cup\mathscr{C}_e$ consists of all cycles. The Ramsey numbers $r_k(\mathscr{C})$, $r_k(\mathscr{C}_0)$ and $r_k(\mathscr{C}_e)$ denote the least integers p such that if the edges of K_p are coloured from a set of k colours at least one of the monochromatic subgraphs must contain a cycle, an odd cycle or an even cycle, respectively.

One sees immediately that $r_k(\mathscr{C})=2k+1$, and it is not difficult to show that $r_k(\mathscr{C}_0)=2^k+1$ as has been remarked both in [2] and [3]. It is the purpose of this note to show that $r_k(\mathscr{C}_e)=2\left\lceil\frac{3k}{2}\right\rceil$; where $\lceil\cdot\rceil$ denotes the least integer function.

THEOREM. *For* $k\geqq 1$, $r_k(\mathscr{C}_e)=2\left\lceil\frac{3k}{2}\right\rceil$.

PROOF. We show first that $r_k(\mathscr{C}_e)\leqq 2\left\lceil\frac{3k}{2}\right\rceil$. Suppose the edges of K_p, $p=2\left\lceil\frac{3k}{2}\right\rceil$, are coloured with k colours so that the resulting monochromatic subgraphs have no even cycles. Let G be such a monochromatic subgraph. Since G has no even cycles, each of its blocks must be either a single edge or an odd cycle. If the blocks of G contain exactly h odd cycles, then G has at most $p-1+h$ edges. When $k=2m-1$, $m\geqq 1$, then $p=6m-2$ and G has at most $3m-2$ odd cycles (the maximum being attained when all are triangles) and hence at most $p-1+(3m-2)=9m-5$ edges. Thus we require $k(9m-5)\geqq\frac{p(p-1)}{2}$ which is impossible. Similarly, when $k=2m$, $m\geqq 1$, then $p=6m$, G has at most $3m-1$ odd cycles and hence at most $9m-2$ edges. We then require $k(9m-2)\geqq\frac{p(p-1)}{2}$ which is also impossible.

* Partial support was provided by the Natural Sciences and Engineering Research Council of Canada under Grant A—4792.

1980 *Mathematics Subject Classification.* Primary 05C55; Secondary 05C38.
Key words and phrases. Graphs, cycle, Ramsey number.

To complete the proof we need only to construct appropriate k-colourings of K_{p-1}, $p=2\left[\frac{3k}{2}\right]$.

When $k=2m-1$, $p-1=6m-3$ and we partition the vertices of K_{6m-3} into three sets $A=\{a_1, a_2, \ldots, a_{2m-1}\}$, $B=\{b_1, b_2, \ldots, b_{2m-1}\}$ and $C=\{c_1, c_2, \ldots, c_{2m-1}\}$. Then $\mathscr{E}_i$, the edges of colour i, $1\leqq i\leqq 2m-1$, are given by the triangles $\{a_i, b_i, c_i\}$, $\{a_i, b_{i+1}, b_{i-1}\}$, $\{a_i, b_{i+2}, b_{i-2}\}, \ldots, \{a_i, b_{i+m-1}, b_{i+m}\}$, $\{b_i, c_{i+1}, c_{i-1}\}$, $\{b_i, c_{i+2}, c_{i-2}\}, \ldots, \{b_i, c_{i+m-1}, c_{i+m}\}$, $\{c_i, a_{i+1}, a_{i-1}\}$, $\{c_i, a_{i+2}, a_{i-2}\}, \ldots, \{c_i, a_{i+m-1}, a_{i+m}\}$, where subscript addition is performed modulo $2m-1$ on the residues $1, 2, \ldots, 2m-1$. To clarify the construction the edges of $\mathscr{E}_i$ are exhibited in Figure 1.

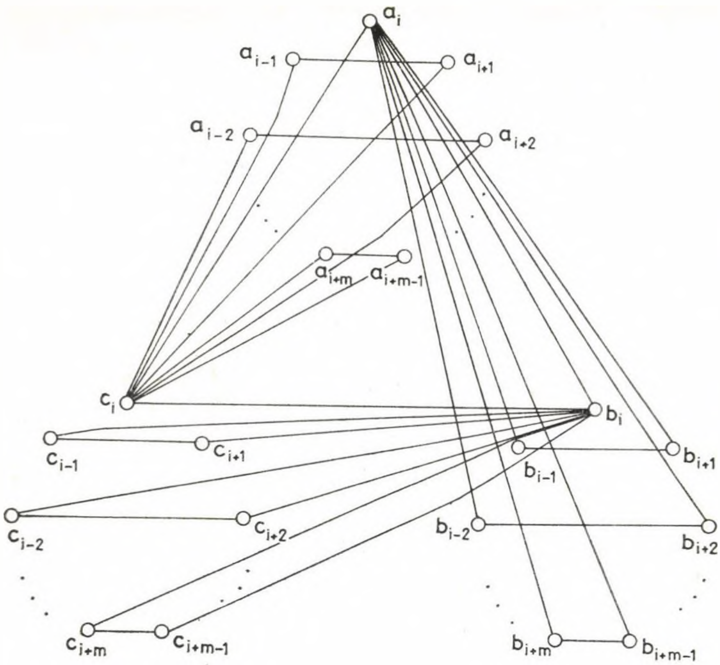


Fig. 1

When $k=2m$, $p-1=6m-1$ and we partition the vertices of K_{6m-1} into three sets $A=\{a_1, a_2, \ldots, a_{2m-1}\}$, $B=\{b_1, b_2, \ldots, b_{2m-1}, b_\infty\}$ and $C=\{c_1, c_2, \ldots, c_{2m-1}, c_\infty\}$. Then $\mathscr{E}_i$, the edges of colour i, $1\leqq i\leqq 2m-1$, are given by the edges $\{b_i, b_\infty\}$, $\{c_i, c_\infty\}$ and the triangles $\{a_i, b_\infty, c_i\}$, $\{a_i, b_{i+1}, b_{i-1}\}$, $\{a_i, b_{i+2}, b_{i-2}\}, \ldots, \{a_i, b_{i+m-1}, b_{i+m}\}$, $\{b_i, c_{i+1}, c_{i-1}\}$, $\{b_i, c_{i+2}, c_{i-2}\}, \ldots, \{b_i, c_{i+m-1}, c_{i+m}\}$, $\{c_i, a_{i+1}, a_{i-1}\}$, $\{c_i, a_{i+2}, a_{i-2}\}, \ldots, \{c_i, a_{i+m-1}, a_{i+m}\}$. The edges $\mathscr{E}_i$ are shown in Figure 2. Finally, $\mathscr{E}_{2m}$, the edges of colour $2m$, consists of the edges $\{b_1, c_1\}$, $\{b_2, c_2\}, \ldots, \{b_{2m-1}, c_{2m-1}\}$, $\{b_\infty, c_\infty\}$ and the triangles $\{a_1, b_1, c_\infty\}$, $\{a_2, b_2, c_\infty\}, \ldots, \{a_{2m-1}, b_{2m-1}, c_\infty\}$. ∎

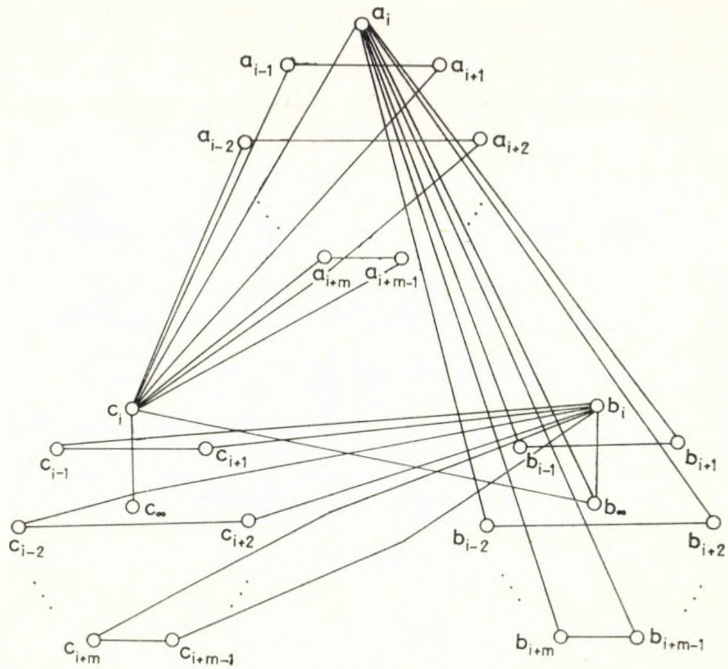


Fig. 2

REFERENCES

[1] BONDY, J. A. and MURTY, U. S. R., *Graph theory with applications,* American Elsevier Publishing Co., Inc., New York, 1976. *MR* **54** # 117.

[2] ERDŐS, P. and GRAHAM, R. L., On partition theorems for finite graphs, *Infinite and finite sets* (Colloq., Keszthely, 1973), edited by A. Hajnal, R. Rado and V. T. Sós, Colloq. Math. Soc. J. Bolyai, **10**, North-Holland Publishing Co., Amsterdam, 1975, 515—527. *MR* **51** # 10159.

[3] HARARY, F., HSU, D. and MILLER, Z., The biparticity of a graph, *J. Graph Theory* **1** (1977), 131—133. *MR* **56** # 2874.

[4] PARSONS, T. D., Ramsey graph theory, *Selected topics in graph theory,* edited by L. W. Beineke and R. J. Wilson, Academic Press, London, 1978, 361—384. *MR* **81e**: 05059.

(Received July 8, 1981)

DEPARTMENT OF MATHEMATICS
SIMON FRASER UNIVERSITY
BURNABY, B.C.
V5A 1S6
CANADA

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 335—340

MATCHINGS AND k-FACTORS IN A RANDOM GRAPH

ANDRZEJ RUCIŃSKI

We consider a random graph obtained from a complete labelled graph by an independent deletion of each edge with the same probability. Conditions for almost sure occurence of a number of disjoint, large matchings are given. We also discuss the problem of existence of a k-factor in a random graph.

1. Introduction

One of the most important and interesting topics in the random graph theory is the problem of the almost sure existence of subgraphs of given type in a random graph $K_{n,p}$. The results in this field give usually a threshold function $q(n)$ such that $K_{n,p(n)}$ for $p(n)$ greater than $q(n)$ contains considered subgraphs with probability tending to 1 as $n\to\infty$. Sometimes, additionally, this limit probability for $K_{n,q(n)}$ is calculated. This problem becomes more difficult if we consider "large" subgraphs, i.e. subgraphs which grow as $n\to\infty$. There are only few such results in the random graph literature (paths [1], trees [5], [9], connected components [5], cliques [4], [12], 1-factors [7], Hamilton cycles [10], [11], [13]).

In this paper we show results about maximal matchings and k-factors, two kinds of the "large" subgraphs of a random graph. Theorem 1 deals with the almost sure existence of a number of disjoint matchings. Main result (Theorem 2) is an attempt to confirm a conjecture of Erdős and Rényi [8]. It is also an approximation of the solution of k-factor problem for a random graph $K_{n,p}$. The analogous problem for a bipartite random graph was solved in [8].

2. Notations and definitions

For elementary definitions from graph theory see [2]. A set of disjoint edges we call *matching* and a k-regular spanning subgraph we call *k-factor*. We use the typical graph theory notation. In particular, we denote by $e(G)$, $\omega(G)$, $\delta(FG)$ and $\Delta(G)$ the number of edges, number of connected components, minimal and maximal degree in a graph G, respectively. We will write $e(S, T)$ for the number of edges between disjoint vertex-subsets S and T and $d_G(v)$ for the degree of vertex v in a graph G. By $G[S]$ we mean a subgraph of G with vertex-set (ev. edge-set) S and all edges of G with both end-points in S (resp. all vertices of G which are endpoints of edges from S).

1980 *Mathematics Subject Classification.* Primary 05C70; Secondary 60C05.
Key words and phrases. Matching, k-factor, random graph, almost sure property.

Now we establish the probabilistic notation. A binomial random variable with parameters n and p we denote by $X(n, p)$. By a *random graph* $K_{n,p}$ we mean the probabilistic space $(\mathcal{G}, \mathsf{P})$, where $\mathcal{G}$ is the family of all graphs with vertex-set V, $|V|=n$ and for every $G \in \mathcal{G}$:

$$\mathsf{P}(G) = p^{e(G)}(1-p)^{\binom{n}{2}-e(G)}.$$

We may interpret $K_{n,p}$ as a result of a random, independent deletion of each edge of a complete graph K_n with the same probability $1-p$. A property $\mathcal{A}$ of a random graph is the set of all graphs among $\mathcal{G}$ which possess this property. We say $K_{n,p}$ has a property $\mathcal{A}$ almost surely (in short a.s.) iff $\mathsf{P}(\mathcal{G}-\mathcal{A})=o(1)$ as $n\to\infty$.

Finally, let $p_k(n)=\frac{1}{n}(\log n+k\log\log n+w(n))$, where $w(n)\to\infty$ as $n\to\infty$, $k=0, 1, 2, \ldots$.

3. Useful results

The fundamental tools needed in our proofs are the following formulas and results from probability, graph and random graph theories.

(1) $$1-x \leq e^{-x}, \quad \binom{n}{m} \leq \left(\frac{en}{m}\right)^m, \quad \binom{n}{m} \leq \binom{n}{m+1} \quad \text{for} \quad m < \frac{n}{2},$$

(2) $$\mathsf{P}(X(n, p) < j-1) \leq \mathsf{P}(X(n, p) < j) \leq \mathsf{P}(X(n-1, p) < j),$$

(3) Bernstein's improvement of Chebyshev inequality (see [14]):

$$\mathsf{P}(|X(n, p)-np| \geq \varepsilon n) \leq 2\exp\{-2pq\varepsilon^2 n(2pq+\varepsilon)^{-2}\} \quad \text{for} \quad 0<\varepsilon<pq, \quad q=1-p,$$

(4) Tutte's k-factor theorem (see [2]):

A graph G contains a k-factor iff for every disjoint vertex-subsets S and D the following inequality holds:

$$q_k(D, S)+k|S|-\sum_{v\in S} d_{G-D}(v) \leq k|D|,$$

where $q_k(D, S)$ is the number of these connected components C of the graph $G-D-S$ for which the number $k|S|+e(C, S)$ is odd;

(5) if a graph G has a Hamilton cycle, then for every non-empty vertex-subset S we have $\omega(G-S) \leq |S|$;

(6) if $p=p_0(n)$, then $K_{n,p}$ contains a 1-factor a.s. (see [7]);

(7) if $p=p_1(n)$, then $K_{n,p}$ contains a Hamilton cycle a.s. (see [10]);

(8) if $p=p_{k-1}(n)$, then $\delta(K_{n,p}) \geq k$ a.s. (a simple corollary of results from [6]).

4. Matchings

Now we formulate our "matching theorem" in its general form:

THEOREM 1. *For every* $0<a<b<1$, *if* $m=m(n), l=l(n), p=p(n)$ *fulfil the following conditions:*

1° $\lim_{n\to\infty} m=\infty$,

2° $l\leqq(1-b)mp$,

3° $lm<n$,

4° $\lim_{n\to\infty} Bn\exp\{-\log m-Apm\}<1$,

where $B=e2^a$, $A=2(a-b)^2(1-a)(2-a-b)^{-2}$, *then a random graph* $K_{n,p}$ *almost surely contains* l *disjoint matchings* $M_1, \ldots, M_l$ *such that:*

(i) $V_1\supset\ldots\supset V_l$, *where* $V_i=V(G[M_i])$, $i=1, \ldots, l$,

(ii) $|V_i|>n-im$, $i=1, \ldots, l$.

We can interpret the above result in the following way: there is a.s. in $K_{n,p}$ under conditions 1°—4° a subgraph F which is almost an l-factor, i.e. F possesses only few points $(o(n))$ with degree in F not equal to l (in fact, they have degrees less than l — see proof). Theorem 1 can be simplified if we assume that $np\to\infty$ and $l<\log(np)$. Then there are a.s. in $K_{n,p}$ l disjoint matchings $M_1, \ldots, M_l$ such that $V_1\supset\ldots\supset V_l$ and $|V_1|=n-o(n)$. From Theorem 1 it easily follows that if m, l and p fulfil the conditions 1°–4° then $K_{n,p}$ a.s. contains (a) $n-lm$ vertices of degree at least l and (b) $\left[\frac{n}{m}\right]$ vertex-disjoint stars on $l+1$ vertices. For $p=$constant a stronger result than (a) can be found in [3].

Let us illustrate the above theorem for typical values of the probability $p=p(n)$.

EXAMPLE 1. If $p=\frac{c}{n}$, then from conditions 1°–4° follows that

$$m=\frac{n}{d},\quad l\leqq(1-b)\frac{c}{d},\quad d(a\log 2+1+\log d)<Ac,$$

where c and d are constants.

EXAMPLE 2. If $p=\frac{w}{n}$, $w=w(n)\to\infty$ as $n\to\infty$, then

$$m=\frac{n}{f},\quad f=f(n)\to\infty\quad\text{as}\quad n\to\infty,\quad l\leqq(1-b)\frac{w}{f},\quad f\log f=o(w).$$

In particular, if $w=\log n$, then

$$m=\frac{n\log\log n}{\log n},\quad l=(1-b)\log\log n.$$

EXAMPLE 3. If $p=$const, then

$$m=\frac{n}{f},\quad l\leqq(1-b)\frac{np}{f},\quad f\log f<Apn,\quad f=f(n)\to\infty\quad\text{as}\quad n\to D.$$

In particular, if $m=\log n$, then $l=(1-b)p\log n$, whereas if $m=\sqrt{n}$, then $l=(1-b)p\sqrt{n}$.

PROOF of Theorem 1.

LEMMA 1. *If a graph G has l disjoint matchings $M_1, \dots, M_l$ and M is a maximal matching in the graph* $G-\bigcup_{i=1}^{l} M_i$, *then* $\Delta(G[V-V(G[M])])\leq l$.

PROOF. By contradiction. ∎

We shall use also an algorithm which, step by step, constructs disjoint matchings and at the same time deletes from a graph all its unmatched vertices.

The algorithm.

1. Set $i=1$, $V_0=V$, $M_0=\emptyset$;
2. Find a maximal matching M_i in the graph $G[V_{i-1}]-M_{i-1}$ and set $V_i=V(G[M_i])$;
3. If $V_i\neq\emptyset$ then go to 2, else stop.

By Lemma 1 for every i we have $\Delta(G[V_{i-1}-V_i]\leq i-1)$. Notice that the set $V_{i-1}-V_i$ is the same as the set of vertices deleted from the graph at the "ith" step of our algorithm.

Let $\mathsf{P}(n, p, m, l)$ denote the probability that $K_{n,p}$ contains a subgraph H with m vertices and $\Delta(H)<l$. We search for such values of m, l and p that $lm<n$ and $\mathsf{P}(n, m, l, p)=o(1)$. Then our algorithm can be repeated l times and everytime it a.s. deletes less than m vertices. But

$$\mathsf{P}(n, p, m, l)\leq\binom{n}{m}\mathsf{P}(\Delta(K_{m,p})<l)$$

and therefore the following lemma completes the proof of Theorem 1.

LEMMA 2. *For every* $0<a<b<1$, *if* $l\leq(1-b)mp$, *then*

$$\mathsf{P}(\Delta(K_{m,p})<l)\leq 2^{am}\exp\{-Apm^2\}.$$

PROOF. Let $V(K_{m,p})=\{1, \dots, m\}$ and A_i denote the events "$d(i)<l$", $i=1, \dots$ $\dots, m$. We have

$$\mathsf{P}(\Delta(K_{m,p})<l)=\mathsf{P}(A_1\cap\dots\cap A_m)=\mathsf{P}(A_1)\mathsf{P}(A_2/A_1)\dots\mathsf{P}(A_m/A_1\cap\dots\cap A_{m-1})$$

and

$$\mathsf{P}(A_{i+1}/A_1\cap\dots\cap A_i)=\sum_{j=1}^{l} q_j\mathsf{P}(d_{K_{m-1,p}}(i+1)<j)\overset{(2)}{\leq}$$

$$\leq\mathsf{P}(d_{K_{m-1,p}}(1)<l)=\mathsf{P}(i),$$

$i=1, \dots, m$ and q_j is the probability that the $(i+1)$-th vertex is joined with exactly $l-j$ vertices among the set $\{1, \dots, i\}$. Thus

$$\mathsf{P}(\Delta(K_{m,p})<l)\leq\mathsf{P}(0)\dots\mathsf{P}(am)\overset{(2)}{\leq}\{\mathsf{P}(am)\}^{am}.$$

Moreover,

$$\mathsf{P}(am) = \mathsf{P}\big(X((l-a)\cdot m-l,\, p) < l\big) \overset{(3)}{\leqq} 2\exp\{-Apm\}.$$

So Lemma 2 is proved. ▮

5. k-factors

Erdős and Rényi in [7] solved the problem of a.s. existence of a 1-factor in a random graph $K_{n,p}$ and formulated in [8] the

CONJECTURE. If $p=p_{k-1}(n)$, then $K_{n,p}$ contains k disjoint 1-factors a.s., $k=1, 2, \ldots$.

From the result of Komlós and Szemerédi [10] see $\big(3$—$(7)\big)$ it follows that the above conjecture is also true for $k=2$. We give a partial solution of this problem

THEOREM 2. *If* $p=p_{k+2}(n)$, *then* $K_{n,p}$ *contains a k-factor a.s.* $k=1, 2, \ldots$.

PROOF. Basing on results (4), (5), (6), (7) and the fact that $q_r(D, S)=q_s(D, S)$ iff $r=s(\mathrm{mod}\ 2)$ one can deduce the following a.s. implication: if $K_{n,p}$ has no k-factor, then there exist disjoint vertex-subsets D and S such that $|D|<|S|$ and

$$(k-1)|D|+\sum_{v\in S} d_{K_{n,p}-D}(v) < (k+1)|S|. \tag{9}$$

Furthermore, if there is a vertex v in the set D for which $e(v, S)<k$, then we can transfer v from D to $V-D-S$ and note that the inequality (9) for the set $D-v$ also holds. Moreover, it follows from (8) that

$$\sum_{v\in S} d_{K_{n,p}}(v) \geqq (k+3)|S|$$

and thus

$$e(S, D) \geqq 2|S|+1.$$

So it must be: $|D|\geqq 3$, $e(v, S)\geqq k$ for every $v\in D$ and $|S|\geqq k$. Summarizing, if $K_{n,p}$ has no k-factor, then the event A below occurs a.s.

Event A="there exist disjoint vertex-subsets D and S such that
1° $3\leqq|D|<|S|$, $|S|\geqq k$,
2° $e(S, D)\geqq 2|S|+1$,
3° inequality (9) holds".

We have to show that $\mathsf{P}(A)=o(1)$. Let us write $A=\bigcup_{s=k}^{n} A_s$, where A_s denote the event A with additional condition that $|S|=s$. Since the inequality (9) implies two others:

$$e(S, V-D-S) < (k+1)|S|-(k-1)|\bar{S}| \tag{9'}$$

and

$$2e(K_{n,p}[S]) < (k+1)|S|, \tag{9''}$$

we can define two other events B_s and C_s which are similar to A_s but instead of (9) they are described by (9′) and (9″), respectively.

Let $s_0=\left\lfloor\frac{n}{k}\right\rfloor$. Using (1) it is possible to obtain that

$$\mathsf{P}\left(\bigcup_{s=k}^{s_0} A_s\right) \leqq \sum_{s=k}^{s_0} \mathsf{P}(B_s) \leqq \sum_{s=k}^{s_0}\sum_{d=3}^{s-1}\binom{n}{s}\binom{n-s}{d}\binom{sd}{2s+1}\times\ldots$$

$$\ldots\times\sum_{j=0}^{s_1}\binom{s(n-s-d)}{j}p^{j+2s+1}q^{s(n-s-d)-j} = o(1),$$

where $s_1=(k+1)s-3(k-1)-1$, $q=1-p$.

From the other side we can calculate with the help of (1) and (3) that

$$\mathsf{P}\left(\bigcup_{s=s_0}^{n} A_s\right) \leqq \sum_{s=s_0}^{n} \mathsf{P}(C_s) \leqq \sum_{s=s_0}^{n}\binom{n}{s}\mathsf{P}\left\{2X\left(\binom{s}{2},p\right)<(k+1)s\right\} = o(1).$$

Thus with probability tending to 1 as $n\to\infty$ the random graph $K_{n,p}$ contains a k-factor and Theorem 2 is proved. ∎

Added in proof. The conjecture from [8] has been completely confirmed by E. Shamir and E. Upfal in „On factors in random graphs", *Israel J. Math.* **39** (1981), 296—302 (*MR* **83i**: 05061). Their proof is of algorithmic nature.

REFERENCES

[1] Ajtai, M., Komlós, J. and Szemerédi, E., The longest path in a random graph, *Combinatorica* **1** (1981), 1—12. *MR* **82d**: 05074.

[2] Bollobás, B., *Extremal graph theory*, London Mathematical Society Monographs, 19, Academic Press, Inc., London—New York, 1978. *MR* **80a**: 05120.

[3] Bollobás, B., Degree sequences of random graphs, *Discrete Math.* (to appear).

[4] Bollobás, B. and Erdős, P., Cliques in random graphs, *Math. Proc. Cambridge Philos. Soc.* **80** (1976), 419—427. *MR* **58**#16408.

[5] Erdős, P. and Rényi, A., On the evolution of random graphs, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5** (1960) 17—61. *MR* **23**#A2338.

[6] Erdős, P. and Rényi, A., On the strength of connectedness of a random graph, *Acta Math. Acad. Sci. Hungar.* **12** (1961), 261—267. *MR* **24**#A54.

[7] 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.

[8] Erdős, P. and Rényi, A., On random matrices II, *Studia Sci. Math. Hungar.* **3** (1968), 459—464. *MR* **39**#5389.

[9] Karoński, M. and Palka, Z., On the size of a maximal induced tree in a random graph, *Math. Slovaca* **30** (1980), 151—155. *MR* **81k**: 05093.

[10] Komlós, J. and Szemerédi, E., Limit distribution for the existence of Hamilton cycles in a random graph, *Discrete Math.*

[11] Koršunov, A. D., Solution of a problem of P. Erdős and A. Rényi on Hamiltonian cycles in undirected graphs, *Dokl. Akad. Nauk SSSR* **228** (1976), 529—532 (in Russian). *MR* **55**#7842. *Soviet. Math. Dokl.* **17**, (1976), 760—764.

[12] Matula, D. W., The largest clique size in a random graph, Technical Report, Department of Computer Science Southern Methodist Univ., Dallas, 1976.

[13] Pósa, L., Hamiltonian circuits in random graphs, *Discrete Math.* **14** (1976), 359—364. *MR* **52**#10497.

[14] Rényi, A., *Foundations of probability*, Holden-Day, Inc., San Francisco, Calif.—London—Amsterdam, 1970. *MR* **41**#9314.

(Received July 16, 1981)

INSTYTUT MATEMATYKI
UNIWERSYTET IM. ADAMA MICKIEWICZA
UL. MATEJKI 48/49
PL—60-769 POZNAŃ
POLAND

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 341—351

THE CONSTRUCTION OF LEAST FAVOURABLE DISTRIBUTIONS IS TRACEABLE TO A MINIMAL PERIMETER PROBLEM

F. ÖSTERREICHER

Huber and Strassen have shown in [2] that a composite testing problem can be replaced by an equivalent single one (in terms of least favourable pairs of distributions), when both hypotheses are given by 2-alternating capacities. In this paper a technique is presented in order to construct the least favourable distribution, if one of the hypotheses is a simple one. The mentioned technique is based on the perimeter of the risk sets in question and applies not only to the total-variation and the ε-contamination model on a general measurable space (cf. [1], [8] and [5]) but also to local-variation and Prochorov-neighbourhoods of distributions on the real line. It is shown that in these cases one can by-pass the application of deeper existence theorems (as Theorem 4.1 in [2]) by elementary geometry. As indicated by an example, the technique can be adapted for future applications.

1. Preliminaries

Let $\mathcal{M}_1$ be the set of all probability measures (pm's) on a measurable space $(\Omega, \mathfrak{A})$, let $\mathcal{Q}$ be a subset of $\mathcal{M}_1$ with elements Q, Q' and let $P \in \mathcal{Q}^c$ be a further pm. Then let us consider the testing problem (tp) $(P, \mathcal{Q})$ of the simple hypothesis P against the composite $\mathcal{Q}$.

The main ideas of the present paper can be developed most instructively and powerfully in terms of the risk sets and risk functions of the testing problems in question. Therefore let us recall the following definitions.

Definition 1. $R(P, Q) = \text{co}\{(P(A), Q(A^c)) \forall A \in \mathfrak{A}: P(A) + Q(A^c) \leqq 1\}$ is called risk set of the tp (P, Q). The lower boundary $r_{(P,Q)}(\alpha) := \min\{y: (\alpha, y) \in R(P, Q)\}$ $\alpha \in [0, 1]$ of this set is called risk function (rf) of (P, Q).

Remark 1. $r_{(P,Q)}(\alpha)$ represents the probability of type II error for an optimal test. In the case of strict convexity of $r_{(P,Q)}(\alpha)$ in α

$$r_{(P,Q)}(\alpha) = Q(A_t^c) \quad \text{with} \quad A_t := \{\omega \in \Omega: q(\omega) > tp(\omega)\}$$

and $t = D_+ r_{(P,Q)}(\alpha)$. ($D^+ r$ $(D_+ r)$ resp. $D^- r$ $(D_- r)$ denotes (the absolute value of) the right-hand-side resp. left-hand-side derivative of the convex function r and p, q the Radon—Nikodym derivatives of P resp. Q with respect to a dominating σ-finite measure μ.) An obvious property of the rf is

$$0 \leqq r_{(P,Q)}(\alpha) \leqq 1 - \alpha \qquad \forall \alpha \in [0, 1],$$

1980 *Mathematics Subject Classification.* Primary 62F35; Secondary 62F03.
Key words and phrases. Robust hypotheses testing, least favourable distributions, Prochorov neighbourhoods.

where the first equality holds true for all $\alpha\in[0,1]$ iff $P\perp Q$. The second inequality holds true for one $\alpha\in(0,1)$ and hence $\forall\alpha\in[0,1]$ iff $P=Q$.

DEFINITION 2. $$R(P,\mathcal{Q}):=\bigcap_{Q'\in\mathcal{Q}} R(P,Q')$$

respectively $$r_{(P,\mathcal{Q})}(\alpha):=\sup\{r_{(P,Q')}(\alpha)\colon Q'\in\mathcal{Q}\}$$

is called the risk set resp. risk function of the composite testing problem $(P,\mathcal{Q})$.

DEFINITION 3. A pm $Q^*\in\mathcal{Q}$ is said to be a least favourable distribution (lfd) for the tp $(P,\mathcal{Q})$ iff $r_{(P,Q^*)}\equiv r_{(P,\mathcal{Q})}$.

REMARK 2. Let $\mathrm{Per}\,(R(P,Q))$ resp. $\mathrm{Per}\,(R(P,\mathcal{Q}))$ denote the perimeter of the risk set $R(P,Q)$ resp. $R(P,\mathcal{Q})$, e.g.

$$\mathrm{Per}\,(R(P,Q))=1-r_{(P,Q)}(0)+\int_0^1\sqrt{1+D_-^2 r_{(P,Q)}(\alpha)}\,d\alpha+\sqrt{2},$$

then obviously

$$\mathrm{Per}\,(R(P,Q^*))=\mathrm{Per}\,(R(P,\mathcal{Q}))=\min\{\mathrm{Per}\,(R(P,Q')),\ Q'\in\mathcal{Q}\}.$$

As a matter of fact, $\mathrm{Per}\,(R(P,Q))$ can be expressed in terms of an information divergence measure of probability distributions

$$H_\Phi(P,Q)=\int\Phi\left(\frac{q}{p+q}\right)(p+q)\,d\mu$$

as used in Theorem 6.1 of [2] in order to characterize least favourable pairs of distributions. Take $\Phi(u)=\sqrt{u^2+(1-u)^2}$ then

$$H_\Phi(P,Q)=\int\sqrt{p^2+q^2}\,d\mu=\int_{\{p=0\}}q\,d\mu+\int_{\{p>0\}}\sqrt{1+(q/p)^2}\,dP.$$

For the easy proof that the latter can be written in terms of the risk function by $1-r_{(P,Q)}(0)+\int_0^1\sqrt{1+D_-^2 r_{(P,Q)}(\alpha)}\,d\alpha$ and for further insights from the integralgeometric point of view we refer to [7].

For the existence of least favourable distributions and, more generally, least favourable pairs of distributions cf., eg. Theorem 4.1 in [2] and Satz 6.4 in [3]. However, in our special cases, both the existence and the construction is traced back to a problem of elementary geometry (see page 13).

The main interest of this paper is actually concentrated on the construction of the lfd Q^* for the case when $(\Omega,\mathfrak{A})=(\mathbf{R},\mathfrak{B})$ is the real line $\mathbf{R}$ equipped with its Borel σ-field $\mathfrak{B}$ and $\mathcal{Q}$ is a *Prochorov—neighbourhood* $\mathcal{Q}(Q;\varepsilon,\delta)$ of $Q\in\mathcal{M}_1(\mathbf{R},\mathfrak{B})$; $0\leq\varepsilon<\infty$, $0\leq\delta<1$

$$\mathcal{Q}(Q;\varepsilon,\delta)=\{Q'\in\mathcal{M}_1(\mathbf{R},\mathfrak{B})\colon Q'(B)\leq Q(B^\varepsilon)+\delta\quad\forall B\in\mathfrak{B}\}$$

where $B^\varepsilon:=\{x\in\mathbf{R}\colon \inf\{|x-y|,\ y\in B\}\leq\varepsilon\}$.

The case $\delta=0$ $\mathcal{Q}_{lv}(Q;\varepsilon)=(Q;\varepsilon,0)$ is called a *local-variation neighbourhood* of Q. (An application of a related distance in discrete geometry may be found in [4].)

In the case $\varepsilon=0$ we can even forget about the restriction to $(\mathbf{R}, \mathfrak{B})$ when constructing Q^*. In this case

$$\mathcal{Q}_{\mathrm{tv}}(Q;\delta) = \{Q' \in \mathcal{M}_1(\Omega, \mathfrak{A}) \colon Q'(A) \leqq Q(A)+\delta \quad \forall A \in \mathfrak{A}\}$$

is called a *total-variation neighbourhood.*

2. Construction of the least favourable distribution

The construction of the lfd for general $(\varepsilon, \delta > 0)$ Prochorov-neighbourhoods will be managed by a super-position of the solutions for local-variation and total-variation neighbourhoods. Let us start with the easiest case, namely total-variation neighbourhoods, and recall the following theorem from [5]. Notice, however, that for both the total-variation model and the ε-contamination model the construction of a least favourable pair was given already by Huber [1] and later by Rieder [8] and Österreicher [5].

THEOREM 1. *Let $(P, \mathcal{Q})$ be a composite testing problem, where $\mathcal{Q}=\mathcal{Q}_{\mathrm{tv}}(Q;\delta)$ is a total-variation neighbourhood of the pm Q. Then the least favourable distribution Q^* is given by*

$$q^*(\omega) = \begin{cases} \bar{t}p(\omega) & \omega \in A_{\bar{t}} \\ q(\omega) & \omega \in A_{\underline{t}}/A_{\bar{t}} \\ \underline{t}p(\omega) & \omega \in A^c_{\underline{t}} \end{cases}$$

if $r_{(P,Q)}(0)+\delta > 1$. Thereby $\bar{t}$ and $\underline{t}$ are chosen such that $1-\bar{t}\alpha$ and $\underline{t}(1-\alpha)$ are supporting lines relative to $r_{(P,Q)}+\delta$. If $r_{(P,Q)}(0)+\delta<1$ we have to replace $\bar{t}$ by $t_{\max}= =D_+ r_{(P,Q)}(0)$ and to redefine

$$q^*(\omega) = \bar{t}p(\omega) \quad \textit{for} \quad \omega \in A_{\bar{t}}$$

by

$$q^*(\omega) = \big(1-\delta/Q(A_{t_{\max}})\big)q(\omega) \quad \textit{for} \quad \omega \in A_{t_{\max}}.$$

REMARK 3. The idea of the proof, for which we refer to [5], is very vivid. It arise from the conjecture that, owing to

$$Q'(A^c_t) \leqq T(A^c_t) := Q(A^c_t)+\delta \qquad \forall t \geqq 0,\ Q' \in \mathcal{Q}$$

$$\operatorname{Per}\big(\operatorname{co}\{(0,1),(1,0)\} \cup \{(P(A_t), T(A^c_t)) \colon P(A_t)+T(A^c_t) \leqq 1,\ t \geqq 0\}\big) = \\ = \min\{\operatorname{Per}\big(R(P,Q')\big) \colon Q' \in \mathcal{Q}\}.$$

In order to tackle the construction of the lfd in the case of local-variation neighbourhoods, let us provide some further auxiliary means.

Let $F(x)=P\big((-\infty, x)\big)$ and $G(x)=Q\big((-\infty, x)\big)$, resp. $F_+(x)=P\big((-\infty, x]\big)$ and $G_+(x)=Q\big((-\infty, x]\big)$ be the distribution functions of the pm's P and Q on $(\mathbf{R}, \mathfrak{B})$. Let furthermore

$$\mathcal{A} := \{\alpha \in [0,1] \colon \exists x \in \mathbf{R} \colon F(x)=\alpha\},$$

$$\alpha_- = \max\{\alpha' \in \mathcal{A} \colon \alpha' \leqq \alpha\} \quad \text{and} \quad \alpha_+ = \inf\{\alpha' \in \mathcal{A} \colon \alpha' \geqq \alpha\}$$

and $F^{-1}(\alpha)=\max\{x \in \mathbf{R} \colon F(x) \leqq \alpha\}$, then

Definition 4. The function ψ: $[0, 1]\to[0, 1]$, given by

$$\psi_{(F,G)}(\alpha) \;(=\psi_{(P,Q)}(\alpha)) = \begin{cases} 1-G_+(F^{-1}(\alpha)) & \text{for } \alpha\in\mathscr{A}: F(F^{-1}(\alpha)) = F_+(F^{-1}(\alpha)) \\ 1-G(F^{-1}(\alpha_-))-\dfrac{\alpha-\alpha_-}{\alpha_+-\alpha_-}(G_+(F^{-1}(\alpha_-))-G(F^{-1}(\alpha_-)))\times & \\ \times 1_{(0,1]}(\alpha_+-\alpha_-) & \text{otherwise} \end{cases}$$

is called pre-risk function (prf) coordinated with (F, G) (the testing problem (P, Q)).

Remark 4. Obviously, $\psi_{(F,G)}=:\psi$ is monotone decreasing (and right-hand-side continuous). Hence it has a finite derivative $\lambda_{[0,1]}$-a.e. Setting

$$\mathscr{A}_t^c(\psi) = \{\alpha\in(0, 1]: D_-\psi(\alpha) \leqq t\} \quad (\cup\{0\} \text{ if } \psi(0) = 1) \quad t \geqq 0$$

$$\mathscr{A}_\infty(\psi) = \cap\{\mathscr{A}_t,\, t\in[0, \infty)\} \quad \text{and abbreviating} \quad \mathscr{A}_t = \mathscr{A}_t(\psi)$$

it can be readily seen

$$r_{(P,Q)}(\alpha) \equiv r_\psi(\alpha) := \min\{y\colon (\alpha, y)\in R\},$$

where

$$R = \mathrm{co}(\{(0, 1), (1, 0)\}\cup\{(\lambda(\mathscr{A}_t), \lambda(\psi(\mathscr{A}_t^c))),\ t \geqq 0\}),$$

which justifies the notation 'pre-risk function' for $\psi_{(P,Q)}$. Notice furthermore that $r_{(P,Q)}\equiv\psi_{(P,Q)}$ iff the likelihood ratio $\frac{q}{p}$ is monotone decreasing.

Let us explain the above statements for the case, when both distribution functions F and G are strictly monotone on the whole real line in order to avoid any unnecessary effort in writing. Then with $F(x)=\alpha$ we get $\psi_{(F,G)}(\alpha)=\psi(\alpha)=1-G(F^{-1}(\alpha))$ which is strictly monotone decreasing in $[0, 1]$ and differentiable in $(0, 1)$. Furthermore

$$\frac{d\psi(\alpha)}{d\alpha} = -\frac{dG(F^{-1}(\alpha))}{d(F^{-1}(\alpha))}\frac{dF^{-1}(\alpha)}{d\alpha} = -\frac{\dfrac{dG(x)}{dx}}{\dfrac{dF(x)}{dx}} = -\frac{q(x)}{p(x)}.$$

Thus a best test 1_{A_t} with $A_t=\{x\colon \frac{q}{p}(x)>t\}$ is transformed into $1_{\mathscr{A}_t}$ with

$$\mathscr{A}_t = \left\{\alpha\colon \left|\frac{d\psi(\alpha)}{d\alpha}\right|>t\right\}$$

and

$$P(A_t) = P(\{x\in\mathbf{R}\colon F(x)\in\mathscr{A}_t\}) = \lambda(\mathscr{A}_t)$$

and also

$$Q(A_t^c) = Q(\{x\in\mathbf{R}\colon F(x)\in\mathscr{A}_t^c\}) = Q(\{x\in\mathbf{R}\colon 1-G(x)\in\psi(\mathscr{A}_t^c)\}) = \lambda(\psi(\mathscr{A}_t^c)).$$

Hence we get in view of the definition of $r_{(P,Q)}$ for

$$\alpha_t = P(A_t) = \lambda(\mathscr{A}_t) \qquad r_{(P,Q)}(\alpha_t) = Q(A_t^c) = \lambda(\psi(\mathscr{A}_t^c)).$$

Suppose now that $\frac{q}{p}(x)$ is monotone decreasing, then so is $\left|\frac{d\psi(\alpha)}{d\alpha}\right|$ and

$A_t=(-\infty, x_t)$, $\mathscr{A}_t=(0, \alpha_t=F^{-1}(x_t))$ and $\lambda(\mathscr{A}_t)=\alpha_t$. Furthermore $\psi(\mathscr{A}_t^c)=[0, \psi(\alpha_t)]$ and hence $\lambda(\psi(\mathscr{A}_t^c))=\psi(\alpha_t)$ and finally $r_{(P,Q)}(\alpha_t)=\psi(\alpha_t)$.

Now let $\varepsilon>0$, $G_\varepsilon(x)=G(x-\varepsilon)$, $G^\varepsilon(x)=G(x+\varepsilon)$ and $\psi_\varepsilon=\psi_{(F,G^\varepsilon)}$ resp. $\psi^\varepsilon=\psi_{(F,G^\varepsilon)}$. Then

DEFINITION 5. The set $\Psi(\psi_{(P,Q)};\varepsilon)$ of all monotone decreasing right-hand-side continuous functions ψ': $[0,1]\to[0,1]$, which are linear in each interval $(a,b)\subset\subset[0,1]\setminus\mathscr{A}$ and satisfy

$$\psi_\varepsilon(\alpha)\leqq\psi'(\alpha)\leqq\lim_-\psi'(\alpha)\leqq\lim_-\psi^\varepsilon(\alpha)\quad\text{for all}\quad\alpha\in[0,1]$$

and $\psi'(1)=0$, is called the ε-corridor of $\psi_{(P,Q)}$ (whereby $\lim_-\psi'(\alpha)$ denotes the left-hand-side limit of ψ' in $\alpha\in(0,1]$. $\lim_-\psi'(0):=1$).

Let in addition to $F^{-1}(\alpha)=\max\{x\in\mathbf{R}\colon F(x)\leqq\alpha\}$ denote

$$F^{-1}_{\min}(\alpha)=\inf\{x\in\mathbf{R}\colon F(x)=\alpha\}\qquad\alpha\in[0,1].$$

Now we are able to state

LEMMA 1. *For each element ψ' of the ε-corridor $\Psi(\psi_{(P,Q)};\varepsilon)$ there exists a probability measure $Q'\in\mathscr{Q}_{\mathrm{lv}}(Q;\varepsilon)$ such that*

$$\psi_{(P,Q')}\equiv\psi'.$$

This pm is unique modulo P and is characterized by its distribution function

$$G_{\psi'}(x)=G'(x)=\begin{cases}1-\lim_-\psi'(F(x)) & \text{if}\quad F^{-1}_{\min}(F(x))+2\varepsilon>F^{-1}(F(x))\\ 1-\lim_-\psi'(F(x)) & \text{for}\quad x\in(F^{-1}_{\min}(F(x)),F^{-1}_{\min}(F(x))+\varepsilon]\\ G(x) & \text{for}\quad x\in(F^{-1}_{\min}(F(x))+\varepsilon, F^{-1}(F(x))-\varepsilon]\\ 1-\psi'(F(x)) & \text{for}\quad x\in(F^{-1}(F(x))-\varepsilon, F^{-1}(F(x))]\\ & \text{if}\quad F^{-1}_{\min}(F(x))+2\varepsilon\leqq F^{-1}(F(x)).\end{cases}$$

PROOF. $\psi_{(P,Q')}\equiv\psi'$ holds true owing to the definition of $Q'\in\mathscr{M}_1(\mathbf{R},\mathfrak{A})$. Thus it remains to check $Q'\in\mathscr{Q}_{\mathrm{lv}}(Q;\varepsilon)$, i.e.

$$Q(B)\leqq Q(B^\varepsilon)\qquad\forall B\in\mathfrak{B}.\tag{1}$$

We are going to do this by providing (1) successively for

(a) sets $B=[x',x)\quad\forall x>x'$

(b) sets $B=\bigcup_{i=1}^{n}[x_i',x_i)\quad\forall x_n>x_n'>x_{n-1}>\ldots>x_1>x_1'$.

Hence, by the well-known monotone class argument we get (1) for all

$$B\in\mathfrak{A}_\sigma(\{(-\infty,x),x\in\mathbf{R}\})=\mathfrak{B}.$$

At present we can observe from $\lim_-\psi_\varepsilon(\alpha)\leqq\lim_-\psi'(\alpha)\leqq\lim_-\psi^\varepsilon(\alpha)$ $\forall\alpha\in[0,1]$ and from the definition of G'

$$(1-G^\varepsilon)(x)\leqq(1-G')(x)\leqq(1-G_\varepsilon)(x)\qquad\forall x\in\mathbf{R}.$$

Applying these inequalities we have

ad (a): $Q(B^\varepsilon)-Q'(B)\geqq Q([x'-\varepsilon, x+\varepsilon))-Q'([x', x))=$
$=[(1-G_\varepsilon)(x')-(1-G')(x')]+[(1-G')(x)-(1-G^\varepsilon)(x)]\geqq 0.$

ad (b): If $[x'_j, x_j)^\varepsilon\cap[x'_{j+1}, x_{j+1})^\varepsilon\neq\emptyset$ let us abbreviate $C=[x'_j, x_j)\cup[x'_{j+1}, x_{j+1})$ and $C_1=[x'_j, x_{j+1})$.

Then obviously

$$Q(C^\varepsilon)-Q'(C) = Q(C_1^\varepsilon)-Q'(C_1)+Q'([x_j, x'_{j+1})) \geqq Q(C_1^\varepsilon)-Q'(C_1).$$

Hence we can assume that the intervals $[x'_i, x_i)^\varepsilon$ $i=1(1)n$ are disjoint. Then, however, $Q(B^\varepsilon)-Q'(B)=\sum_{i=1}^{n} Q([x'_i, x_i)^\varepsilon)-Q'([x'_i, x_i))$ and (1) for sets of type (b) follows immediately from the validity of (1) for sets of type (a). ∎

DEFINITION 6. Let ψ' be the pre-risk function of a testing problem (P, Q') on $(\mathbf{R}, \mathfrak{B})$ and $r_{\psi'}$ the associated risk function, then the arc length $l(\psi')$ resp. $l(r_{\psi'})$ of the curve

$$\{(\alpha, y)\colon 0\leqq\alpha\leqq 1,\ \psi'(\alpha)\leqq y\leqq \lim_{-}\psi'(\alpha)\}\quad \text{resp.}$$

$$\{(0, y)\colon r_{\psi'}(0)\leqq y\leqq 1\}\cup\{(\alpha, y)\colon 0<\alpha\leqq 1,\ y=r_{\psi'}(\alpha)\}$$

is called the arc length of the pre-risk function ψ' resp. of the associated risk function.

LEMMA 2. $l(\psi')=l(r_{\psi'})$.

PROOF. In view of Remark 4

$$\sum_{\alpha\in\mathscr{A}_\infty(\psi')} \lim_{-}\psi'(\alpha)-\psi'(\alpha) = 1-r_{\psi'}(0).$$

Hence we can assume ψ' to be a continuous mapping onto $[0, r_{\psi'}(0)]$ and get

$$l(\psi') = \int_{(0,1]} \sqrt{1+D_-^2\psi'(\alpha)}\, d\lambda(\alpha) = \int_0^\infty \lambda(\sqrt{1+D_-^2\psi'(\alpha)}>t)\,dt =$$

$$= \int_0^\infty \lambda(\mathscr{A}_{\sqrt{t^2-1}}(\psi'))\,dt = \int_0^\infty \lambda(\mathscr{A}_{\sqrt{t^2-1}}(r_{\psi'}))\,dt = l(r_{\psi'}).\quad ∎$$

THEOREM 2. *Let $(P, \mathscr{Q})$ be a composite testing problem, where $\mathscr{Q}=\mathscr{Q}_{lv}(Q;\varepsilon)$ is a local-variation neighbourhood of the pm Q on $(\mathbf{R}, \mathfrak{B})$. Let furthermore be ψ^* that element of $\Psi(\psi_{(P,Q)};\varepsilon)$ with minimal arc length, i.e. which fulfils*

$$l(\psi^*) = \min\{l(\psi')\colon \psi'\in\Psi(\psi_{(P,Q)};\varepsilon)\}.$$

Then the probability measure Q^* which is characterized by G_{ψ^*} (in the sense of Lemma 1) is the least favourable distribution for the testing problem $(P, \mathscr{Q})$.

PROOF. According to the definition of $\Psi(\psi_{(P,Q)};\varepsilon)$ we have $\psi_{(P,Q')}\in$ $\in\Psi(\psi_{(P,Q)};\varepsilon)$ if $Q'\in\mathscr{Q}_{lv}(Q;\varepsilon)$. On the other hand we have Lemma 1. Hence there is a one-to-one correspondence $\mathcal{Q}_{lv}(Q;\varepsilon)\leftrightarrow\Psi(\psi_{(P,Q)};\varepsilon)$ modulo P and therefore

$$\inf\{l(\psi'),\ \psi'\in\Psi(\psi_{(P,Q)};\varepsilon)\}=\inf\{l(\psi_{(P,Q')}),\ Q'\in\mathcal{Q}\}.$$

Moreover, we can replace 'infimum' by 'minimum', since an element $\psi^*\in\Psi(\psi_{(P,Q)};\varepsilon)$ with minimal arc length exists due to our construction (see page 349 ff.). Because the arc length is not altered by switching over to the associated risk functions $r_{(P,Q')}$ (see Lemma 2) we have

$$l(\psi^*)=\min\{l(\psi'),\ \psi'\in\Psi(\psi_{(P,Q)};\varepsilon)\}=\min\{l(r_{(P,Q')}),\ Q'\in\mathcal{Q}\}=l(r_{(P,Q^*)}).$$

In view of Remark 2 $Q^*\in\mathcal{Q}$ is the desired least favourable distribution. The latter is unique (modulo P) since ψ^* is unique. ∎

Let us illustrate the above theorem by a simple

EXAMPLE. Let $P=N(\xi,\sigma)$, $\xi>0$, $0<\sigma\leq 1$; $Q=N(0,1)$ and $\mathcal{Q}_{lv}(Q;\varepsilon)$ an ε-local-variation neighbourhood of Q with $0<\varepsilon<\xi$. Then we get with $\alpha=F(x)=$ $=\Phi((x-\xi)/\sigma)$, $G_\varepsilon(x)=\Phi(x-\varepsilon)$ and $G^\varepsilon(x)=\Phi(x+\varepsilon)$, $\psi^\varepsilon(\alpha)=1-\Phi(\sigma\Phi^{-1}(\alpha)+\xi-\varepsilon)$ and $\psi_\varepsilon(\alpha)=1-\Phi(\sigma\Phi^{-1}(\alpha)+\xi+\varepsilon)$. Hence

$$\frac{d^2\Phi(\alpha)}{d\alpha^2}=-\sqrt{2\pi}\,\sigma\big(\Phi^{-1}(\alpha)(1-\sigma^2)-\xi\sigma\big)\exp\Big((\Phi^{-1}(\alpha))^2-\frac{1}{2}(\sigma\Phi^{-1}(\alpha)+\xi)^2\Big).$$

Case $\sigma=1$. Then both ψ^ε and ψ_ε are convex and decreasing from 1 to 0. Thus $\psi^\varepsilon\equiv r_{(F,G_\varepsilon)}$ and $\psi_\varepsilon\equiv r_{(F,G^\varepsilon)}$. Hence the element $\psi^*\in\Psi(\psi_{(F,G)};\varepsilon)$ with minimal arc length equals $\psi^*\equiv\psi^\varepsilon$ and so $Q^*=N(\varepsilon,1)$ is the least favourable distribution.

Case $\sigma<1$. ψ^ε and ψ_ε are piecewise convex and concave. Hence in order to find the element $\psi^*\in\Psi(\psi_{(P,Q)};\varepsilon)$ with minimal arc length we have to find a straight line $d-t\alpha$ which is both tangent to ψ^ε and ψ_ε, i.e. which satisfies

$$\psi^\varepsilon(\alpha_1)=d-t\alpha_1,\quad \psi_\varepsilon(\alpha_2)=d-t\alpha_2$$

and

$$\left.\frac{d\psi^\varepsilon(\alpha)}{d\alpha}\right|_{\alpha=\alpha_1}=-t=\left.\frac{d\psi_\varepsilon(\alpha)}{d\alpha}\right|_{\alpha=\alpha_2}.$$

(Further details are omitted.) Then

$$\psi^*(\alpha)=\begin{cases}\psi^\varepsilon(\alpha) & \alpha\leq\alpha_1\\ d-t\alpha & \alpha\in(\alpha_1,\alpha_2]\quad 0<\alpha_1<\alpha_2<1\\ \psi_\varepsilon(\alpha) & \alpha>\alpha_2\end{cases}$$

and finally

$$q^*(x)=\begin{cases}q(x-\varepsilon) & x\leq\sigma\Phi^{-1}(\alpha_1)+\xi\\ tp(x) & x\in(\sigma\Phi^{-1}(\alpha_1)+\xi,\ \sigma\Phi^{-1}(\alpha_2)+\xi]\\ q(x+\varepsilon) & x>\sigma\Phi^{-1}(\alpha_2)+\xi\end{cases}$$

is the density of the least favourable distribution, whereby p, q and q^* are the densities of P, Q and Q^* with respect to λ. Note that for none of the cases q^*/p fits the family of censored versions in the total-variation- resp. ε-contamination sense.

REMARK 5. $Q'\left(\left\{\frac{q^*}{p}\leqq t\right\}\right)\leqq Q^*\left(\left\{\frac{q^*}{p}\leqq t\right\}\right)=Q\left(\left\{\frac{q^*}{p}\leqq t\right\}^\varepsilon\right)$ $\forall t\geqq 0$. To see this, observe that

$$\mathscr{A}_t^c(\psi^*)=\sum_{i\in I}\langle\alpha_i,\alpha_{i+1}],\quad I\subset\mathbf{N},$$

where

$$\psi^*(\alpha_i)=\psi^\varepsilon(\alpha_i);\quad \lim_-\psi^*(\alpha_{i+1})=\lim_-\psi_\varepsilon(\alpha_{i+1}).$$

Then in view of the already discussed one-to-one correspondence $\mathscr{A}_t^c(\psi^*)\leftrightarrow\left\{\frac{q^*}{p}\leqq t\right\}$ we have $\left\{\frac{q^*}{p}\leqq t\right\}=\sum_{i\in I}[x_i,x_{i+1})$, where the sets $[x_i,x_{i+1})^\varepsilon$ are disjoint,

$$\psi^*(\alpha_i)=Q^*([x_i,\infty))=Q([x_i-\varepsilon,\infty))=\psi^\varepsilon(\alpha_i)$$

and

$$\lim_-\psi^*(\alpha_{i+1})=Q^*(x_{i+1},\infty))=Q([x_{i+1}+\varepsilon,\infty))=\lim_-\psi_\varepsilon(\alpha_{i+1}).$$

Hence owing to Lemma 1

$$Q'\left(\left\{\frac{q^*}{p}\leqq t\right\}\right)=\sum_{i\in I}Q'([x_i,\infty))-Q'([x_{i+1},\infty))\leqq$$

$$\leqq\sum_{i\in I}Q([x_i-\varepsilon,\infty))-Q([x_{i+1}+\varepsilon,\infty))=\sum_{i\in I}Q^*([x_i,\infty))-Q^*([x_{i+1},\infty))=$$

$$=Q\left(\left\{\frac{q^*}{p}\leqq t\right\}^\varepsilon\right)=Q^*\left(\left\{\frac{q^*}{p}\leqq t\right\}\right).$$

The following generalization of [4; Theorem 2.4] is an immediate consequence of Theorem 2.

THEOREM 3. *Let $(P,\mathscr{Q})$ be a composite testing problem, where $\mathscr{Q}=\mathscr{Q}(Q;\varepsilon,\delta)$ is a Prochorov-neighbourhood of the pm Q on $(\mathbf{R},\mathfrak{B})$. Let furthermore $\mathscr{Q}_{lv}(Q;\varepsilon)$ be the corresponding local-variation neighbourhood and let Q^*_{lv} be the lfd for the tp $(P,\mathscr{Q}_{lv}(Q;\varepsilon))$.*

Then the least favourable distribution Q^* with respect to $(P,\mathscr{Q})$ equals the least favourable distribution Q^*_{tv} for $(P,\mathscr{Q}_{tv}(Q^*_{lv};\delta))$, where $\mathscr{Q}_{tv}(Q^*_{lv};\delta)$ is the total-variation neighbourhood of Q^*_{lv}.

PROOF. Let $A_{lv}(t)=\{q^*_{lv}>tp\}$ resp. $A_{tv}(t)=\{q^*_{tv}>tp\}$ $t\geqq 0$ be the net of sets corresponding to the nonrandomized optimal tests for the tp (P,Q^*_{lv}) resp. the tp (P,Q^*_{tv}). Let further ψ^*_{lv} be the element of the ε-corridor $\Psi(\psi_{(P,Q)};\varepsilon)$ with minimal arc length.

Then, owing to Remark 5,

$$Q^*_{lv}(A^c_{lv}(t))=Q(A^c_{lv}(t)^\varepsilon)\qquad\forall t\geqq 0$$

and hence

$$Q'(A^c_{lv}(t))\leqq Q^*_{lv}(A^c_{lv}(t))+\delta\qquad\forall t\geqq 0\quad\text{and}\quad\forall Q'\in\mathscr{Q}.$$

Applying Theorem 1 to the tp $(P,\mathscr{Q}_{tv}(Q^*_{lv},\delta))$ and choosing t, $\bar t$ resp. t_{max} accordingly, we have

$$A_{lv}(t) = A_{tv}(t)$$

and therefore

$$Q'(A^c_{tv}(t)) \leqq Q^*_{tv}(A^c_{tv}(t)) \qquad \forall t \in [\underline{t}, \bar{t}] \text{ resp. } [\underline{t}, t_{\max}].$$

This yields

$$r_{(P,\mathcal{Q})} \leqq r_{(P,Q^*_{tv})}.$$

On the other hand we have

$$Q^*_{tv}(B) \leqq Q^*_{lv}(B) + \delta \leqq Q(B^\varepsilon) + \delta \qquad \forall B \in \mathfrak{B},$$

i.e. $Q^*_{tv} \in \mathcal{Q}$, and thus $r_{(P,\mathcal{Q})} \equiv r_{(P,Q^*_{tv})}$. ∎

Obviously, the existence of a least favourable distribution reduces in our case to a vivid geometric problem. Namely the existence of an element ψ^* of the ε-corridor of $\psi_{(P,Q)}$ with shortest arc length.

We are going to solve the mentioned extremal problem constructively.

Construction of the element ψ^* of $\Psi(\psi_{(P,Q)}; \varepsilon)$ with minimal arc length.

For the sake of simplicity let us apply the coordinate transform governed by

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix},$$

i.e.

$$A \circ [0, 1]^2 = \{(x, y): -1/\sqrt{2} \leqq x \leqq 1/\sqrt{2}, \max(-x, x) \leqq y \leqq \min(\sqrt{2}+x, \sqrt{2}-x)\}.$$

This allows the following representation of an element $\psi \in \Psi(\psi_{(P,Q)}; \varepsilon)$ by a continuous function

$$\varphi_\psi = \varphi: [-1/\sqrt{2}, 1/\sqrt{2}] \to [0, \sqrt{2}]$$

which satisfies $\varphi_\psi(-1/\sqrt{2}) = \varphi_\psi(1/\sqrt{2}) = 1/\sqrt{2}$ and

$$-1 \leqq D^+\varphi(x), \; D^-\varphi(x) \leqq 1,$$

namely

$$\varphi(x) = x + \sqrt{(\alpha_x - x\sqrt{2})^2 + y_x^2}$$

where

$$(\alpha_x, y_x) = \{(\alpha, y): 0 \leqq \alpha \leqq 1, \quad y = \alpha - x\sqrt{2}, \quad \psi(\alpha) \leqq y \leqq \lim_- \psi(\alpha)\}.$$

Let Φ_ε be the set of all functions $\varphi = \varphi_\psi: \psi \in \Psi(\psi_{(P,Q)}; \varepsilon)$ and $\varphi_\varepsilon = \varphi_{\psi_\varepsilon}$ resp. $\varphi^\varepsilon = \varphi_{\psi_\varepsilon}$. Then a function $\varphi^* \in \Phi_\varepsilon$ of minimal arc length — if any exists — has to satisfy a number of conditions which we now are going to point out:

Of course, for $x \in X_= = \{x \in [-1/\sqrt{2}, 1/\sqrt{2}]: \varphi_\varepsilon(x) = \varphi^\varepsilon(x)\}$ $\varphi^*(x) = \varphi_\varepsilon(x) = \varphi^\varepsilon(x)$ must be fulfilled.

Beside this trivial set $X_=$ there are two further subsets X_ε resp. X^ε of "contact", i.e. sets of points $x\in[-1/\sqrt{2},\ 1/\sqrt{2}]\setminus X_=$ for which $\varphi_\varepsilon(x)=\varphi^*(x)$ resp. $\varphi^\varepsilon(x)=\varphi^*(x)$ holds true — otherwise φ^* could not be of minimal arc length. So necessary conditions for the minimum arc length of φ^* imply further constraints for φ^*.

Let X_ε be the set of all $x\in\left[-\frac{1}{\sqrt{2}},\ \frac{1}{\sqrt{2}}\right]\setminus X_=$ such that there exists a line $d+t\xi$ for which the following conditions hold true:

(1_ε) $\varphi_\varepsilon(x)=d+tx$ and
$\varphi_\varepsilon(\xi)\leqq d+t\xi$ for all $\xi\in(x-\delta,\ x+\delta)$, $\delta>0$ sufficient small;

(2_ε) the component (in the sense of connectedness with respect to the usual topology on $A\circ[0,1]^2$) $C_{(x,\varphi_\varepsilon(x))}$ of $(x,\varphi_\varepsilon(x))$ in $\{(\xi,y)\in A\circ[0,1]^2:\ d+t\xi\leqq y\leqq\varphi^\varepsilon(\xi)\}$ satisfies

$$C_{(x,\varphi_\varepsilon(x))}\subset\{(\xi,y)\in A\circ[0,1]^2:\ \varphi_\varepsilon(\xi)\leqq y\}.$$

Then due to obvious geometric reasons $\varphi^*(x)=\varphi_\varepsilon(x)$ and

$(*_\varepsilon)$ $\qquad \varphi^*(\xi)\leqq d+t\xi$ for all $\xi:\ (\xi,y)\in C_{(x,\varphi_\varepsilon(x))}$.

Let furthermore X^ε be the set of all $x\in[-1/\sqrt{2},\ 1/\sqrt{2}]\setminus X_=$ such that there exists a line $d+t\xi$ for which

(1^ε) $\varphi^\varepsilon(x)=d+tx$ and
$\varphi^\varepsilon(\xi)\geqq d+t\xi$ for all $\xi\in(x-\delta,\ x+\delta)$, $\delta>0$ sufficient small;

(2^ε) the component $C_{(x,\varphi^\varepsilon(x))}$ of $(x,\varphi^\varepsilon(x))$ in $\{(\xi,y)\in A\circ[0,1]^2:\ \varphi_\varepsilon(\xi)\leqq y\leqq d+t\xi\}$ satisfies

$$C_{(x,\varphi^\varepsilon(x))}\subset\{(\xi,y)\in A\circ[0,1]^2:\ y\leqq\varphi^\varepsilon(\xi)\}.$$

Then also $\varphi^*(x)=\varphi^\varepsilon(x)$ and

$(*^\varepsilon)$ $\qquad \varphi^*(\xi)\geqq d+t\xi$ for all $\xi:\ (\xi,y)\in C_{(x,\varphi^\varepsilon(x))}$.

Obviously, $X_\varepsilon\cap X^\varepsilon=\emptyset$. Note that owing to the continuity of φ_ε and φ^ε both $X_=$, $X_\varepsilon\cup X_=$ and $X^\varepsilon\cup X_=$ are closed. The latter in view of $(*_\varepsilon)$ and $(*^\varepsilon)$. Hence $X=X_=\cup X_\varepsilon\cup X^\varepsilon$ is compact. Now let $x\in[-1/\sqrt{2},\ 1/\sqrt{2}]\setminus X$. Then both $x_=(x)$ and $x^=(x)$ with $x_=(x)=\max\{\xi\in X,\ \xi<x\}<x<x^=(x)=\min\{\xi\in X,\ \xi>x\}$ exist.

Suppose that the straight line

$$g(\xi)=\varphi^*(x_=(x))+\frac{\varphi^*(x^=(x))-\varphi^*(x_=(x))}{x^=(x)-x_=(x)}(\xi-x_=(x))$$

does not fulfil $\varphi_\varepsilon(\xi)\leqq g(\xi)\leqq\varphi^\varepsilon(\xi)$, $\xi\in[x_=(x),\ x^=(x)]$, then there are two possibilities: 1) There is a point $\xi_0\in(x_=(x),\ x^=(x))$ satisfying $\varphi_\varepsilon(\xi_0)=\varphi^\varepsilon(\xi_0)$ $(\neq g(\xi_0))$. This is in contradiction to the definition of $X_=$. 2) There is a point $\xi_0\in(x_=(x),\ x^=(x))$ and a straight line $d+t\xi\neq g(\xi)$ parallel to g and satisfying (1_ε) and (2_ε) or (1^ε) and (2^ε) for ξ_0. This is in contradiction to the definition of X_ε resp. X^ε. Hence g fulfils $\varphi_\varepsilon(\xi)\leqq g(\xi)\leqq\varphi^\varepsilon(\xi)$ and, obviously, φ^* has to satisfy $\varphi^*(\xi)=g(\xi)$, $\xi\in[x_=(x),\ x^=(x)]$.

Thus the element $\varphi^* \in \Phi_\varepsilon$ with minimal arc length is completely determined, existent by construction and unique.

Acknowledgement. The author is indebted to Prof. P. J. Huber for a stimulating discussion when entering robustness and to his Salzburg colleagues, especially to Doz. J. Linhart for valuable conversations concerning integral geometry.

REFERENCES

[1] HUBER, P. J., A robust version of the probability ratio test, *Ann. Math. Statist.* **36** (1965), 1753—1758. *MR* **32** # 3209.

[2] HUBER, P. J. and STRASSEN, V., Minimax tests and the Neyman—Pearson lemma for capacities, *Ann. Statist.* **1** (1973), 251—263 and *Ann. Statist.* **2** (1974), 223—224. *MR* **50** # 8776 and **50** # 15027.

[3] HUMMITSCH, W., Ungünstigste Verteilungen und Approximation von Minimax Tests, Dissertation, Ruhr-Universität, Bochum, 1978.

[4] LINHART, J. and ÖSTERREICHER, F., Gleichmäßige Verteilung von Punkten in gewissen metrischen Räumen, speziell auf der Kugel, *Monatsh. Math.* **89** (1980), 111—120. *MR* **81h**: 10073.

[5] ÖSTERREICHER, F., On the construction of least favourable pairs of distributions, *Z. Wahrsch. Verw. Gebiete* **43** (1978), 49—55. *MR* **80c**: 62048.

[6] ÖSTERREICHER, F., Least favorable distributions. Entry from KOTZ, S. and JOHNSON, L. (Editors), *Encyclopedia of Statistical Sciences*, Vol. 4, John Wiley & Sons, New York, 1983, 588—592.

[7] ÖSTERREICHER, F. und FELDMAN, D., Divergenzen von Wahrscheinlichkeitsverteilungen — integralgeometrisch betrachtet, *Acta Math. Acad. Sci. Hungar.* **37** (1981), 329—337.

[8] RIEDER, H., Least favourable pairs for special capacities, *Ann. Statist.* **5** (1977), 909—921. *MR* **57** # 7846.

(Received July 17, 1981)

UNIVERSITY OF SALZBURG
PETERSBRUNNSTRASSE 19
A—5020 SALZBURG
AUSTRIA

and (visiting)

DEPARTMENT OF STATISTICS AND PROBABILITY
MICHIGAN STATE UNIVERSITY
WELLS HALL
EAST LANSING, MI 48824
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 353—366

AUSFÜLLUNG EINES KREISES DURCH KONGRUENTE KREISE IN DER HYPERBOLISCHEN EBENE

K. BEZDEK

Wir betrachten in der euklidischen Ebene ein konvexes Gebiet G. Sind in G mindestens zwei kongruente Kreise eingepackt, so ist die Kreisdichte in G bekanntlich [1] kleiner als $\pi/\sqrt{12}$. Molnár [2] hat gezeigt, daß ein entsprechender Satz auch in der sphärischen Geometrie gilt in der hyperbolischen Geometrie dagegen nicht. In diesem Aufsatz beweisen wir den folgenden Satz, den vorher L. Fejes Tóth als Vermutung ausgesprochen hat:

SATZ. *Sind in der hyperbolischen Ebene in einem Kreis mindestens zwei kongruente Kreise eingelagert, so ist die Kreisdichte in dem Kreis kleiner als* $\pi/\sqrt{12}$.

Dieser Satz ist deshalb bemerkenswert, weil in der Kreispackung, die aus den Flächeninkreisen eines regulären Mosaiks $\{p, 3\}$ mit $p>6$ besteht, die Dichte eines Kreises in der betreffenden Fläche größer ist als $\pi/\sqrt{12}$.

BEWEIS. Es sei K ein Kreis mit dem Mittelpunkt O und dem Radius R in der hyperbolischen Ebene, in dem die nicht übereinandergreifenden Kreise k_i mit dem Mittelpunkt O_i $(i=1, 2, \ldots, n;\ n\geqq 2)$ und dem Radius r liegen. Wir bezeichnen die Dichte der Kreise $k_1, k_2, \ldots, k_n$ bezüglich des Kreises K mit δ $\left(\text{d.h. } \delta=\frac{\sum_{i=1}^{n} k_i}{K}\right)$. Hier und im folgenden bezeichnen wir einen Bereich und seinen Flächeninhalt mit demselben Symbol.

BEMERKUNG 1. Wir können voraussetzen, daß $R>2r$ gilt. Wir haben nämlich $R\geqq 2r$, und im Falle von $R=2r$ ist $n=2$, wann

$$\delta = \frac{1}{2\mathrm{ch}^2\frac{r}{2}} < \frac{1}{2} < \frac{\pi}{\sqrt{12}} \quad \text{gilt.}$$

BEMERKUNG 2. Wir können voraussetzen, daß $n\geqq 3$ gilt. Wenn nämlich $n=2$, so ist $\delta<\frac{1}{2}<\frac{\pi}{\sqrt{12}}$, woraus die Gültigkeit des Satzes folgt.

1980 *Mathematics Subject Classification.* Primary 52A45; Secondary 51M10.
Key words and phrases. Packing, hyperbolic plane.

Es sei $H = \text{conv}\,\{O_1, O_2, \ldots, O_n\}$.

Bemerkung 3. Wir können voraussetzen, daß

$$\dim H = 2.$$

Wegen $n \geqq 3$ ist diese Bemerkung trivial.

Wir nennen einen Mittelpunkt O_i eine Ecke des konvexen Polygons H, wenn O_i am Rande von H liegt.

Bemerkung 4. Wir können ohne Beschränkung der Allgemeinheit voraussetzen, daß die Ecken von H am Rande des Kreises K' mit dem Mittelpunkt O und dem Radius $R-r$ liegen. Offensichtlich ist $H \subset K'$. Liegt eine Ecke O_i im Inneren des Kreises K', so bringt eine Translation in der Richtung der äußeren Winkelhalbierenden der Ecke O_i von H den Kreis k_i in eine solche Lage, daß der Mittelpunkt von k_i schon zum Rand des Kreises K' gehört.

Bemerkung 5. Wir können voraussetzen, daß $O \in \text{int}\, H$. Haben wir nämlich $O \in \text{ext}\, H$, so sind H und O durch eine Seite O_iO_j von H getrennt. Die Punkte O_i, O_j sind Ecken von H, deshalb liegen sie am Rande des Kreises K'. Eine Halbdrehung um den Punkt O überführt die Kreise k_i, k_j in die Kreise k_i^*, k_j^* mit den Mittelpunkten O_i^*, O_j^*. Dann bilden die Kreise $\{k_1, \ldots, k_i^*, \ldots, k_j^*, \ldots, k_n\}$ eine Packung in K, und $O \in \text{int}\,(\text{conv}\,\{O_1, \ldots, O_i^*, \ldots, O_j^*, \ldots, O_n\})$. Ist aber $O \in \text{mar}\, H$, so finden wir eine Seite O_iO_j des konvexen Polygons H, die den Punkt O enthält. Jetzt kann eine geeignete kleine Drehung um O den Kreis k_i in eine solche Lage übertragen, die unserem Zweck entspricht.

Bemerkung 6. Wir können voraussetzen, daß $k_i \setminus H$ zusammenhängend ist, wenn $i \in \{1, 2, \ldots, n\}$. Zuerst nehmen wir an, daß O_i eine Ecke des Polygons H ist (d.h. O_i am Rande des Kreises K' liegt) und k_i eine Strecke der Seite O_jO_l $(j \neq i \neq l)$ von H bedeckt. So enthält die Halbgerade $\overrightarrow{OO_j}$ oder die Halbgerade $\overrightarrow{OO_l}$ einen inneren Punkt von k_i. ($O \in \text{int}\, H$!) Also, wir können ohne Beschränkung der Allgemeinheit voraussetzen, daß die Halbgerade $\overrightarrow{OO_j}$ auch solche Punkte enthält, von denen einige zum Inneren von k_i andere zum Inneren von k_j gehören. Dies ist aber unmöglich, weil die Mittelpunkte O_i, O_j am Rande des Kreises K' liegen und die Kreise k_i, k_j keinen gemeinsamen inneren Punkt haben also $\text{int}\,(\text{conv}\,\{k_i \cup O\}) \cap \text{int}\,(\text{conv}\,\{k_j \cup O\}) = \emptyset$ gelten muß. Schließlich nehmen wir an, daß $O_i \in \text{int}\, H$ und der Kreis k_i die Seiten $O_lO_{l^*}$, $O_jO_{j^*}$ von H schneidet (Abb. 1). Es sei n_l (bzw. n_j) eine Halbgerade mit dem Anfangspunkt O_i, die die Seite $O_lO_{l^*}$ (bzw. $O_jO_{j^*}$) senkrecht schneidet. Es ist leicht einzusehen, daß $n_l \cap (OO_l \cup OO_{l^*}) \neq \emptyset$ oder $n_j \cap (OO_j \cup OO_{j^*}) \neq \emptyset$ gilt. Ohne Beschränkung der Allgemeinheit können wir voraussetzen, daß die Strecke OO_l und die Halbgerade n_l einen gemeinsamen Punkt P haben. Dann überführt die Spiegelung an der Geraden $O_lO_{l^*}$ den Kreis k_i in k_i^l. Ferner sei O_i^l das Spiegelbild von O_i. So ist $O_iO_i^l < 2r \leqq O_iO_l$ also $PO_i^l < PO_l$ woraus $OO_i^l < OO_l = R-r$ folgt. Folglich bilden die Kreise $\{k_1, \ldots, k_{i-1}, k_i^l, k_{i+1}, \ldots, k_n\}$ eine Packung in K, wo wir durch eine Translation erreichen können, daß der Mittelpunkt O_i^l schon am Rande des Kreises K' liegt. Für den Kreis k_i^l ist dann $k_i^l \setminus \text{conv}\,\{O_1, \ldots, O_{i-1}, O_i^l, O_{i+1}, \ldots, O_n\}$ zusammenhängend.

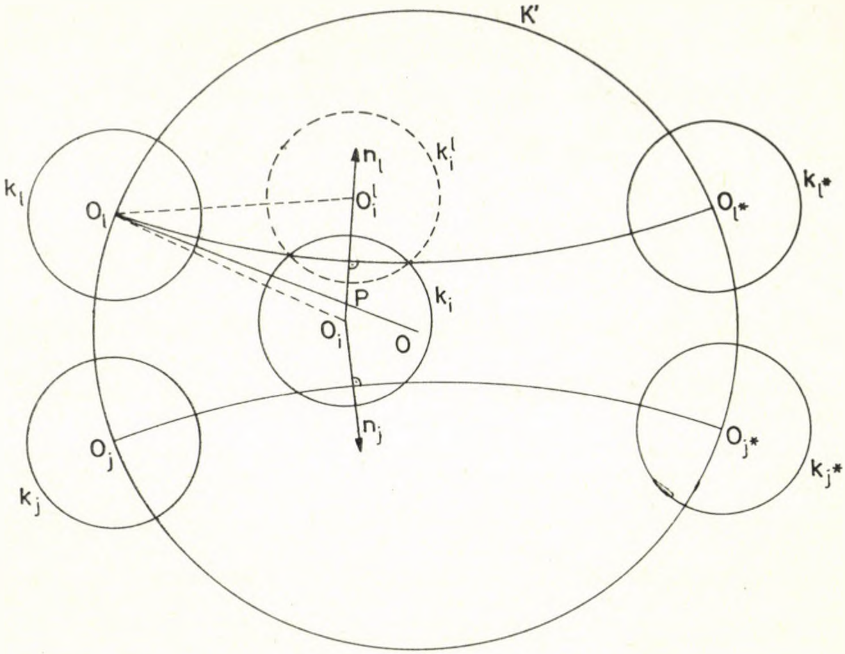


Abb. 1

Wir bezeichnen mit $\frac{2\pi}{a}$ die Winkel eines gleichseitigen Dreiecks mit der Seitenlänge $2r$. So gilt $\operatorname{cosec}\frac{\pi}{a}=2\operatorname{ch} r$.

Im weiteren unterscheiden wir zwei Fälle im Beweis des Satzes.

I. Fall $r \geqq r_0 = \operatorname{arch}\left(\frac{1}{2}\operatorname{cosec}\frac{\pi}{6{,}7}\right)=0{,}457388\ldots$.

Es sei $G=\overline{K\setminus K'}$ ein geschlossener Kreisring. (Hier bedeutet $\overline{K\setminus K'}$ die abgeschlossene Hülle von $K\setminus K'$.) Wir werden jetzt die Dichte der Kreise $\{k_1, k_2, \ldots \ldots, k_n\}$ bezüglich G abschätzen. Zu diesem Zwecke ordnen wir zu jedem Kreis k_l ($l\in\{1, 2, \ldots, n\}$), der mit G gemeinsamen Innenpunkt hat, eine Zelle T_l zu (Abb. 2). Es seien $I_1=\{i|O_i$ liegt am Rande des Kreises $K'\}$ und $I_2=\{j|O_j$ liegt im Inneren des Kreises K' und k_j hat eine gemeinsame Sehne mit $K'\}$.

(i) Im Falle $i\in I_1$ ist die Zelle T_i der Durchschnitt von G und vom kleinsten Winkelbereich mit dem Eckpunkt O, der den Kreis k_i enthält.

(ii) Im Falle $j\in I_2$ ist A_jB_j die gemeinsame Sehne der Kreise K', k_j so gilt $A_jB_j<2r$ und liegen die Punkte O_j, O in einer durch A_jB_j begrenzten offenen Halbebene, nachdem die Zelle als Durchschnitt des Ringes G und des Winkelbereichs $\sphericalangle A_jOB_j(<\pi)$ definiert wird.

Es ist leicht einzusehen, daß die so konstruierten Zellen paarweise keinen gemeinsamen inneren Punkt haben. Ferner enthält jede Zelle den zu G gehörigen Teil des

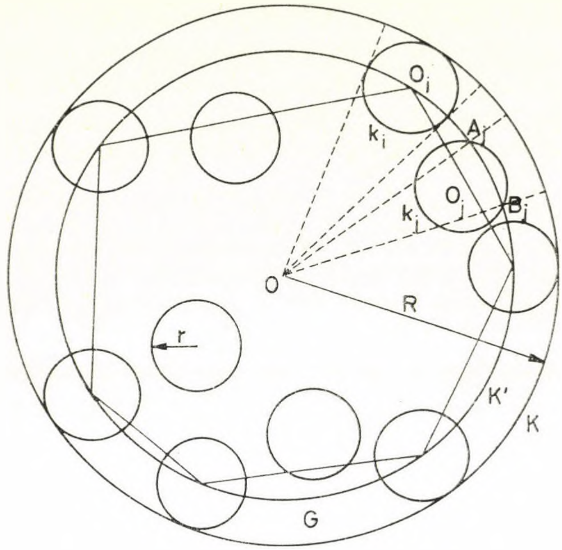


Abb. 2

entsprechenden Kreises. Darum genügt es die Dichte $\frac{\Delta_l}{T_l}$, $\Delta_l = k_l \cap G$, des Kreises k_l bezüglich T_l von oben abzuschätzen.

Für einen Kreis k_j $(j \in I_2)$ sei k'_j ein Kreis mit dem Mittelpunkt O'_j und dem Radius r, wo der Punkt O'_j auf der Halbgeraden $\overrightarrow{OO_j}$ liegt und die gemeinsame Sehne $A'_jB'_j$ der Kreise K', k'_j die Länge $2r$ hat (Abb. 3). Ferner sei $T'_j = \sphericalangle A'_jOB'_j \cap G$ und $\Delta'_j = k'_j \cap G$ (hier ist $\sphericalangle A'_jOB'_j < \pi$).

HILFSSATZ 1.

$$\frac{\Delta_j}{T_j} < \frac{\Delta'_j}{T'_j} \qquad (j \in I_2).$$

BEWEIS. Es sei $\widehat{A_jB_j}$ (bzw. $\widehat{A'_jB'_j}$) der Kreisbogen des Kreises K', der durch Δ_j (bzw. Δ'_j) bedeckt ist. Wir nehmen jetzt einen, mit dem Kreis K', konzentrischen Kreis, dessen Kreisbogen $\widehat{X_jY_j}$ (bzw. $\widehat{X'_jY'_j}$) durch Δ_j (bzw. Δ'_j) bedeckt ist (Abb. 3). Es gilt

$$\frac{\widehat{A'_jB'_j}}{\widehat{A_jB_j}} = \frac{\sphericalangle A'_jOB'_j}{\sphericalangle A_jOB_j}; \qquad \frac{\widehat{X'_jY'_j}}{\widehat{X_jY_j}} = \frac{\sphericalangle X'_jOY'_j}{\sphericalangle X_jOY_j}$$

wo $\sphericalangle A'_jOB'_j > \sphericalangle X'_jOY'_j$ und $\sphericalangle A_jOB_j > \sphericalangle X_jOY_j$. Sogar ist es leicht einzusehen: $\sphericalangle A_jOB_j - \sphericalangle X_jOY_j > \sphericalangle A'_jOB'_j - \sphericalangle X'_jOY'_j$. So ist

$$1 < \frac{\sphericalangle A'_jOB'_j}{\sphericalangle A_jOB_j} < \frac{\sphericalangle X'_jOY'_j}{\sphericalangle X_jOY_j}$$

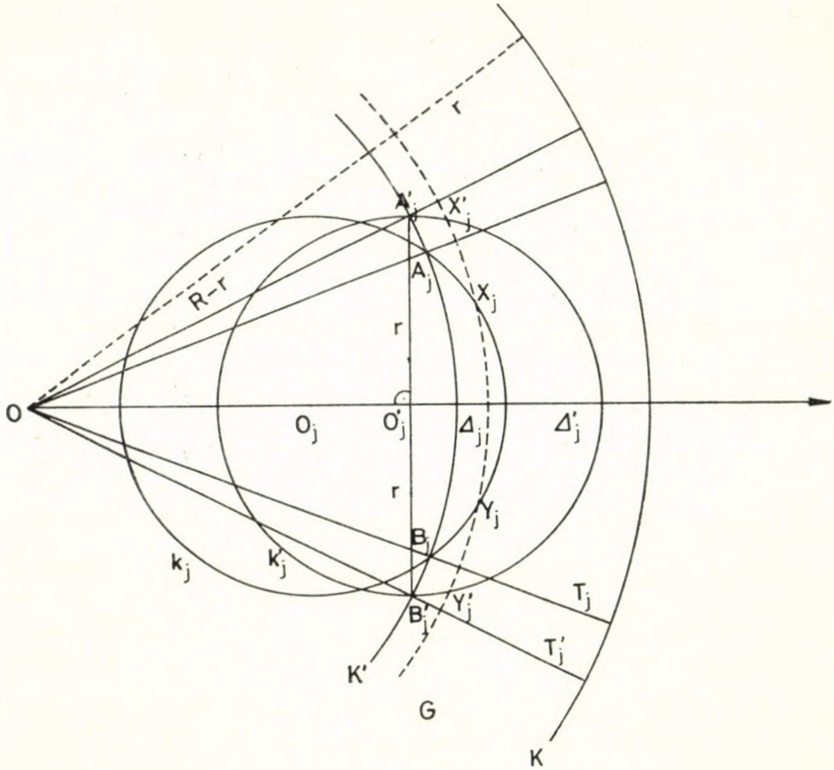


Abb. 3

das heißt

$$\frac{\widehat{A'_j B'_j}}{\widehat{A_j B_j}} < \frac{\widehat{X'_j Y'_j}}{\widehat{X_j Y_j}}.$$

Schließlich erhalten wir ganz einfach

$$\frac{\Delta'_j}{\Delta_j} > \frac{\widehat{A'_j B'_j}}{\widehat{A_j B_j}} = \frac{T'_j}{T_j} \qquad \text{q.e.d.}$$

Nun betrachten wir einen Kreis k_i mit $i \in I_1$ (Abb. 4). Wir bemerken daß $T_i = T'_j$ und $\Delta_i > \Delta'_j$ für beliebigen Index $j \in I_2$. Folglich gilt auf Grund des Hilfssatzes 1:

(1) $$\frac{\Delta_j}{T_j} < \frac{\Delta_i}{T_i}, \qquad j \in I_2 \quad \text{und} \quad i \in I_1$$

Wir führen jetzt noch einige Bezeichnungen bezüglich der Abb. 4 ein. Es sei K'' ein Kreis mit dem Mittelpunkt O und dem Radius $R-2r$ und ferner $G' = \overline{K' \setminus K''}$

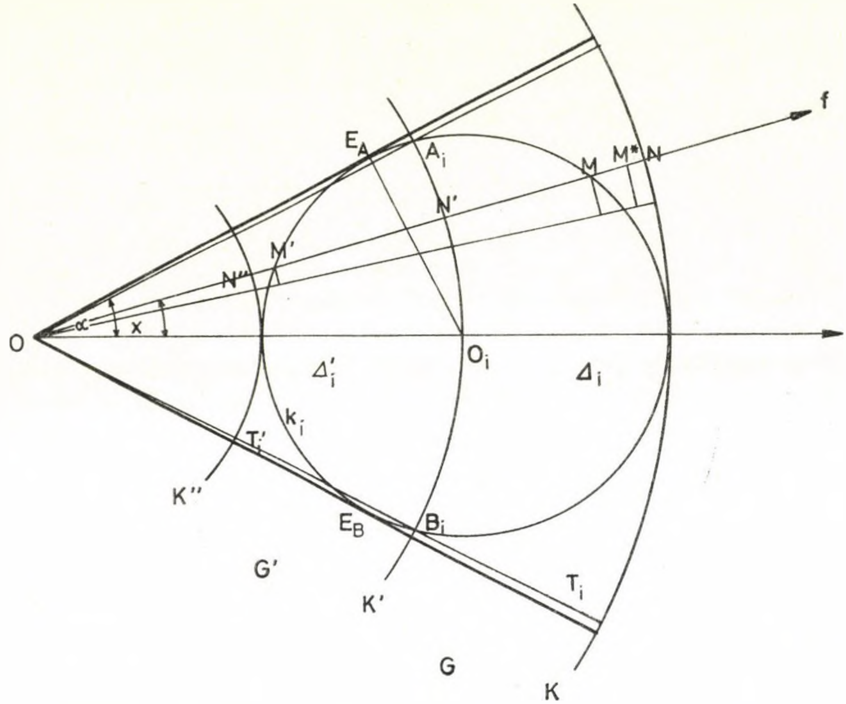


Abb. 4

ein geschlossener Kreisring. Berühren die von O ausgehenden Geraden den Kreis k_i in den Punkten E_A, E_B, so bezeichnen wir den Durchschnitt $\sphericalangle E_A O E_B \cap G'$ (hier ist $\sphericalangle E_A O E_B < \pi$) mit T_i' bzw. $k_i \cap G'$ mit Δ_i'.

HILFSSATZ 2.

$$\frac{\Delta_i}{T_i} < \frac{\Delta_i'}{T_i'}, \qquad i \in I_1$$

BEWEIS. Es sei $A_i B_i$ die gemeinsame Sehne der Kreise K', k_i (Abb. 4). Ferner sei $\hat{T}_i = \sphericalangle A_i O B_i \cap G$; $\hat{T}_i' = \sphericalangle A_i O B_i \cap G'$ und $\hat{\Delta}_i' = k_i \cap \hat{T}_i'$, wo Winkelbereich $\sphericalangle A_i O B_i$ kleiner als π ist. Offenbar haben wir nur die folgende Ungleichung zu zeigen:

$$\text{(2)} \qquad \frac{\Delta_i}{\hat{T}_i} < \frac{\hat{\Delta}_i'}{\hat{T}_i'}.$$

Zu diesem Zwecke sei $\sphericalangle A_i O O_i = \alpha \left(< \sphericalangle E_A O O_i < \frac{\pi}{2} \right)$. Im weiteren bedeutet x einen beliebigen Winkelwert, für den $0 \leqq x \leqq \alpha$ gilt. Dementsprechend nehmen wir die Halbgerade f mit dem Anfangspunkt O im Winkelbereich $\sphericalangle A_i O O_i$, die mit der Halbgeraden $\overrightarrow{OO_i}$ den Winkel x einschließt. Die Halbgerade f schneidet den Rand der Kreise K, K', K'' in den Punkten N, N', N''. Ferner seien die Punkte M, M', der Halbgeraden f zugleich die Punkte des Randes von k_i, wo aber $M \in \hat{T}_i$ bzw. $M' \in \hat{T}_i'$. Wir führen die Bezeichnung $\odot(x)$ für den Flächeninhalt eines Kreises mit dem Radius x ein und definieren die folgenden Funktionen:

S: $[0, \alpha] \to \mathbf{R}$

$$x \mapsto S(x) = \frac{\odot(OM) - \odot(ON')}{\odot(ON) - \odot(ON')}$$

S^*: $[0, \alpha] \to \mathbf{R}$

$$x \mapsto S^*(x) = \frac{\odot(ON') - \odot(OM')}{\odot(ON') - \odot(ON'')}.$$

Es ist leicht einzusehen, daß die Funktionen S, S^* stetig und beschränkt sind, sogar

$$\frac{\Delta_i}{\hat{T}_i} = \frac{1}{\alpha}\int_0^\alpha S(x)\,dx \qquad \frac{\hat{\Delta}_i'}{\hat{T}_i'} = \frac{1}{\alpha}\int_0^\alpha S^*(x)\,dx.$$

Gleichzeitig gilt für $x \in (0, \alpha]$ $S(x) < S^*(x)$, denn $N'M < N'M'$ und eine Halbdrehung um den Punkt N' überführt M' in einen Punkt M^*, wobei

$$\frac{\odot(OM) - \odot(ON')}{\odot(ON) - \odot(ON')} < \frac{\odot(OM^*) - \odot(ON')}{\odot(ON) - \odot(ON')} < \frac{\odot(ON') - \odot(OM')}{\odot(ON') - \odot(ON'')}$$

gilt. So erhalten wir

$$\frac{1}{\alpha}\int_0^\alpha S(x)\,dx < \frac{1}{\alpha}\int_0^\alpha S^*(x)\,dx,$$

womit (2) bewiesen ist. Q.e.d.

Mit Rücksicht auf den Hilfssatz 2 haben wir

$$\frac{\Delta_i}{T_i} < \frac{\Delta_i + \Delta_i'}{T_i + T_i'}.$$

Ist im rechtwinkligen Dreieck $E_A O O_i$ $\sphericalangle E_A O O_i = \beta$, so ergibt sich durch einfache Rechnungen

$$\frac{\Delta_i + \Delta_i'}{T_i + T_i'} = \frac{\pi}{4\,\mathrm{ch}^2 \frac{r}{2}} \frac{\sin\beta}{\beta} < \frac{\pi}{4\,\mathrm{ch}^2 \frac{r}{2}}.$$

Mit Rücksicht auf (1) folgt also

$$\frac{\Delta_j}{T_j} < \frac{\Delta_i}{T_i} < \frac{\pi}{4\,\mathrm{ch}^2 \frac{r}{2}},$$

woraus sich die Ungleichung

$$\delta < \frac{K' + G \dfrac{\pi}{4\,\mathrm{ch}^2 \frac{r}{2}}}{K' + G}$$

ergibt. Wegen

$$\frac{G}{K'}=\left(\operatorname{ch}\frac{r}{2}+\operatorname{cth}\frac{R-r}{2}\operatorname{sh}\frac{r}{2}\right)^2-1>e^r-1$$

erhalten wir dann

$$\delta<\frac{1+(e^r-1)\dfrac{\pi}{4\operatorname{ch}^2\dfrac{r}{2}}}{e^r}=\mathscr{F}(r).$$

Offenbar nimmt die Funktion $\mathscr{F}(r)$ in $(0,+\infty)$ monoton ab, daher gilt $\delta<\mathscr{F}(r)\leqq\leqq\mathscr{F}(r_0)=0{,}906\,659\ldots<\frac{\pi}{\sqrt{12}}=0{,}906\,899\ldots$ was wir beweisen wollten.

II. Fall $r<r_0=\operatorname{arch}\left(\frac{1}{2}\operatorname{cosec}\frac{\pi}{6{,}7}\right)=0{,}457\,388\ldots.$

Wir ordnen jedem Kreis k_i die Menge D_i aller Punkte von H zu, die nicht weiter bei O_i liegen als bei jedem anderen Mittelpunkt O_j $(i\neq j)$. Es ist leicht einzusehen, daß die konvexen Polygone D_i ein Mosaik bilden, daß sie also das konvexe Polygon H, ohne übereinanderzugreifen, vollständig überdecken (Abb. 5).

Liegt k_i in H, so gilt $k_i\subset D_i$. Nach einem Satz von J. Molnár (siehe dazu [3] S. 234) haben wir jetzt

$$\frac{k_i}{D_i}\leqq d(a)=\frac{3\operatorname{cosec}\dfrac{\pi}{a}-6}{[a]-3-\dfrac{6}{\pi}\operatorname{arctg}\left\{\sqrt{3}\operatorname{tg}\dfrac{\pi}{a}\operatorname{ctg}\left(1-\dfrac{[a]}{a}\right)\pi\right\}}.$$

Nach einigen Rechnungen sehen wir, daß die Funktion $d(a)$ in $(6{,}7)$ zuerst streng monoton abnimmt dann zunimmt. Wegen $r<r_0$ gilt $6<a<6{,}7$.

Also wegen $d(6{,}7)=0{,}903\,199\ldots<\frac{\pi}{\sqrt{12}}$ und $d(6)=\lim_{a\to+6}d(a)=\frac{\pi}{\sqrt{12}}$ ergibt sich $d(a)<\frac{\pi}{\sqrt{12}}$ und folglich $\frac{k_i}{D_i}<\frac{\pi}{\sqrt{12}}$.

Es sei $I=\{i|k_i\not\subset H\}$ und $i^*\in I$. Wir betrachten den Streckenzug $L_{i^*}=\overline{\operatorname{int}k_{i^*}\cap\cap\operatorname{mar}H}$. (Wegen der Bemerkung 6 besteht L_{i^*} aus zwei Strecken, wenn O_{i^*} eine Ecke von H ist, übrigens aus einer Strecke.) Es existieren solche eindeutig bestimmte Streckenzüge L_{j^*}; L_{l^*} $(j^*, l^*\in I)$ am Rande des konvexen Polygons H, die im beliebigen Umlaufssinn die Nachbarn des Streckenzugs L_{i^*} sind (Abb. 5). Jetzt definieren wir den Winkelbereich $\sphericalangle O_{l^*}O_{i^*}O_{j^*}$ so, daß $\sphericalangle O_{l^*}O_{i^*}O_{j^*}\supset(k_{i^*}\setminus H)$. Ferner nehmen wir die Tangente e_{l^*} (bzw. e_{j^*}) des Kreises k_{i^*}, deren Berührungspunkt zur Halbgeraden $\overline{O_{i^*}O_{l^*}}$ (bzw. $\overline{O_{i^*}O_{j^*}}$) gehört. Bezeichnen wir die durch e_{l^*} (bzw. e_{j^*}) begrenzte und den Kreis k_{i^*} enthaltene geschlossene Halbebene mit S_{l^*} (bzw. S_{j^*}), so kann man das folgende Vieleck konstruieren

$$\hat{D}_{i^*}=\sphericalangle O_{l^*}O_{i^*}O_{j^*}\cap S_{l^*}\cap S_{j^*}\cap H.$$

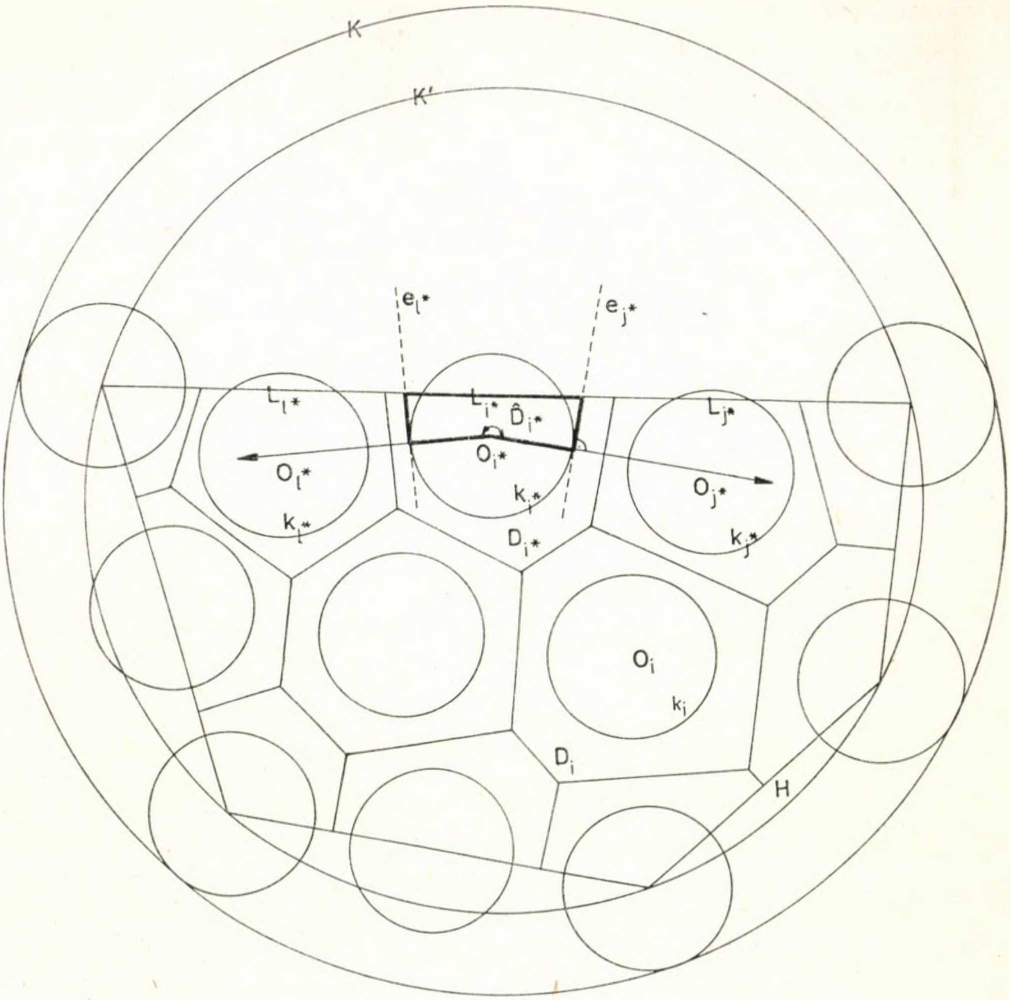


Abb. 5

Es läßt sich leicht beweisen, daß $\hat{D}_{i^*} \subset D_{i^*}$. Und nun betrachten wir die Zelle $\tilde{D}_{i^*} = \hat{D}_{i^*} \cup (k_{i^*} \cap H)$, woraus $k_{i^*} \cap H \subseteq \tilde{D}_{i^*} \subset D_{i^*}$ folgt. Dann nehmen wir die Menge $H_r = A_r \setminus \text{int}\, H$, wo A_r den äußeren Parallelbereich von H vom Abstand r bedeutet. (H_r besteht einerseits aus einigen Sektoren von einem Hyperzyklus vom Abstand r andererseits aus einigen Kreissektoren eines Kreises mit dem Radius r; offenbar ist H_r in K enthalten.) Unser Beweis ist beendet, wenn die folgende Ungleichung gezeigt wird:

$$\frac{\sum_{i \in I} k_i}{H_r + \sum_{i \in I} \tilde{D}_i} < \frac{\pi}{\sqrt{12}}. \tag{3}$$

Wir betrachten gewisse Teilmengen der Menge $H_r \cup (\bigcup_{i \in I} D_i)$ (Abb. 6). Sind die Strekkenzüge L_i, L_j am Rande des konvexen Polygons H benachbart, so bezeichnen wir die Seite von H mit $\mathcal{O}$, die mindestens eine Teilstrecke sowohl von L_i als auch von L_j enthält. Ferner sei t_i (bzw. t_j) die den Punkt O_i (bzw. O_j) enthaltene Gerade, die die Seite $\mathcal{O}$ senkrecht schneidet. Dann können wir den durch t_i, t_j begrenzten Streifen S^{ij} nehmen. Sind jetzt $V_i = t_i \cap \mathcal{O}$ und $V_j = t_j \cap \mathcal{O}$, so wird der zur Strecke V_iV_j gehörige Hyperzyklussektor vom Abstand r mit $H_r^{V_iV_j}$ bezeichnet. Schließlich enthält die durch Gerade O_iO_j begrenzte geschlossene Halbebene S_{ij} den Hyperzyklussektor $H_r^{V_iV_j}$. Also kann die folgende Menge definiert werden:

$$D_{ij} = H_r^{V_iV_j} \cup (\hat{D}_i \cap S^{ij} \cap S_{ij}) \cup (\hat{D}_j \cap S^{ij} \cap S_{ij}) \cup (k_i \cap S^{ij}) \cup (k_j \cap S^{ij}).$$

Der Einfachheit halber sei $k_{ij} = (k_i \cap S^{ij}) \cup (\cap k_j S^{ij})$.

Da offenbar $(\sum k_{ij}) - (\sum_{l \in I} k_l) = (\sum D_{ij}) - (H_r + \sum_{l \in I} \tilde{D}_l) = 0$, folgt die Ungleichung (3) unmittelbar aus der Ungleichung

$$\frac{k_{ij}}{D_{ij}} < \frac{\pi}{\sqrt{12}}. \tag{4}$$

Deshalb benötigen wir nur den Beweis von (4).

Wir können voraussetzen, daß $O_iO_j = 2r$ gilt. Ist nämlich $O_iO_j > 2r$, so können die Kreise k_i, k_j in der Richtung der Geraden O_iO_j verschoben werden, um einander zu berühren. Also während einer geeigneten Translation der Kreise k_i, k_j nimmt D_{ij} ab, d.h. nimmt $\frac{k_{ij}}{D_{ij}}$ zu.

Wir können voraussetzen, daß $O_i = V_i$ oder $O_j = V_j$. Im entgegengesetzten Falle sind nämlich $O_i \neq V_i$; $O_j \neq V_j$. Ohne Beschränkung der Allgemeinheit können wir voraussetzen, daß $O_iV_i \geqq O_jV_j$ ist (Abb. 7).

Zuerst beweisen wir, daß $O_iV_j > 2r$ gilt. Ist nämlich $\pi > \sphericalangle O_iO_jV_j \geqq \frac{\pi}{2}$, so ist unsere Behauptung evident. Daher nehmen wir an, daß $\sphericalangle O_iO_jV_j < \frac{\pi}{2}$ gilt. Im Saccherischen Viereck $V_iV_jO_i'O_i$ gilt dann $O_iO_i' \geqq 2r$. Mit den Bezeichnungen $O_iO_i' = 2x$; $V_iV_j = 2y$; $O_iV_j = V_iO_i' = z$; $O_iV_i = O_i'V_j = m$ erhalten wir $\operatorname{sh} x = \operatorname{sh} y \operatorname{ch} m$ und $\operatorname{ch} z = \operatorname{ch} 2y \operatorname{ch} m$, woraus $\operatorname{ch} z = \operatorname{ch} m + 2\frac{\operatorname{sh}^2 x}{\operatorname{ch} m}$ folgt. Nun ist aber $r < r_0$, deshalb ist $\operatorname{sh} r < \frac{\sqrt{2}}{2}$, woraus sich die Ungleichung $\operatorname{ch} 2r < \operatorname{ch} m + 2\frac{\operatorname{sh}^2 r}{\operatorname{ch} m} \leqq \operatorname{ch} z$, d.h. $2r < z = O_iV_j$ ergibt. Also können wir den Kreis k_j um den Punkt O_i um einen solchen Winkel drehen, der den Mittelpunkt O_j von k_j in einen inneren Punkt der Strecke V_iV_j überführt. Dadurch nimmt D_{ij} ab, d.h. $\frac{k_{ij}}{D_{ij}}$ nimmt zu.

Wegen der obigen Überlegungen genügt es den Fall $O_iO_j = 2r$ und $O_i = V_i$ zu untersuchen. Jetzt werden die Endpunkte des Hyperzyklusbogens von $H_r^{V_iV_j}$ mit W_i, W_j bezeichnet ($W_i \in t_i$, $W_j \in t_j$). Ferner sei $V_jO_j = x$ (Abb. 8).

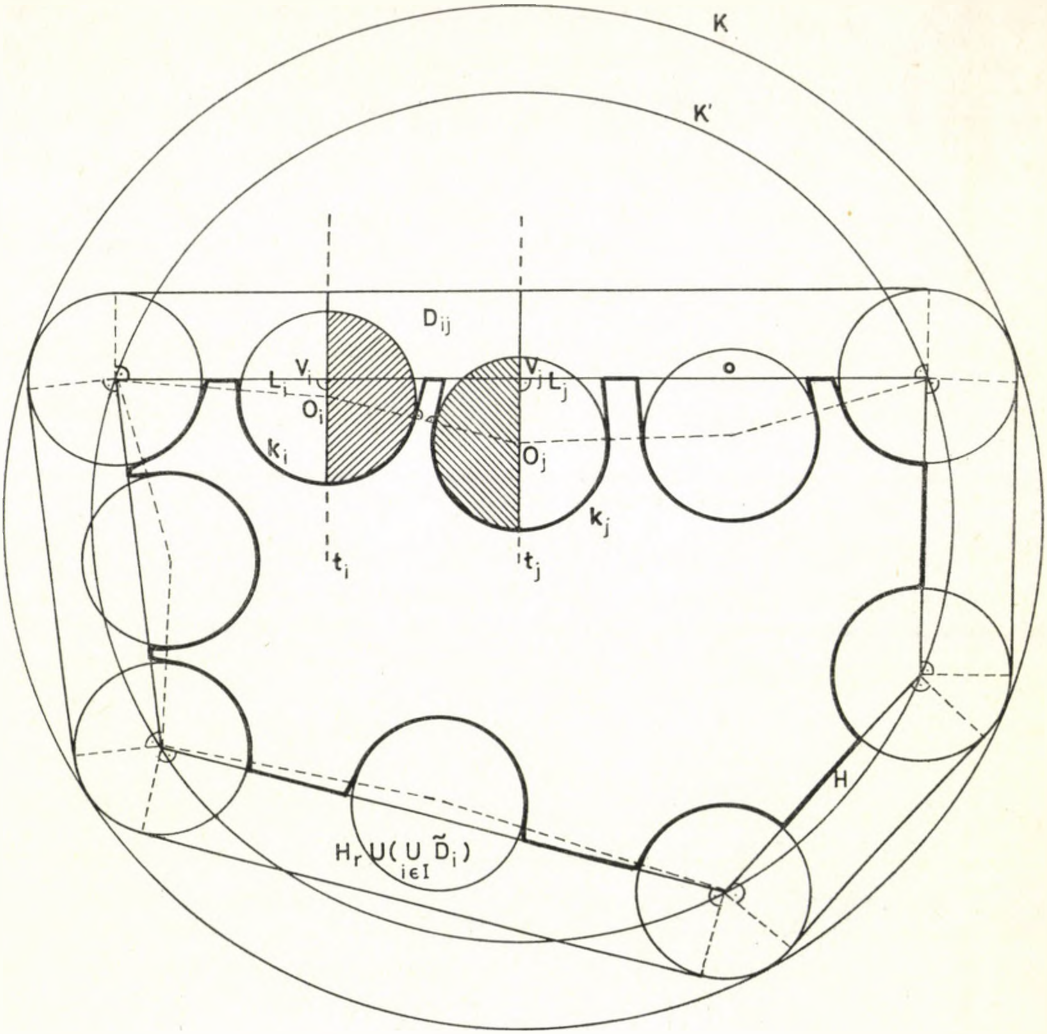

K
K'
D_{ij}
V_i
L_i
O_i
k_i
V_j
L_j
O_j
k_j
o
t_i
t_j
H
$H_r \cup (\bigcup_{i \in I} \tilde{D}_i)$

Abb. 6

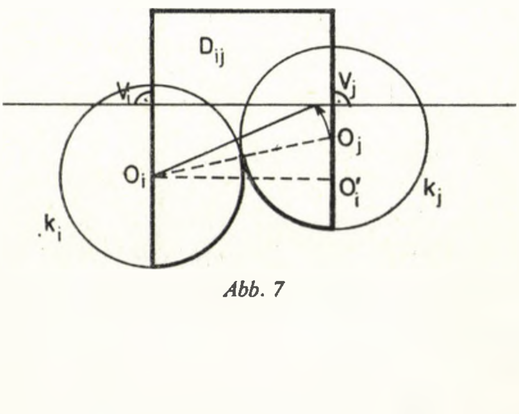

D_{ij}
V_i
V_j
O_j
O_i
O_i'
k_j
k_i

Abb. 7

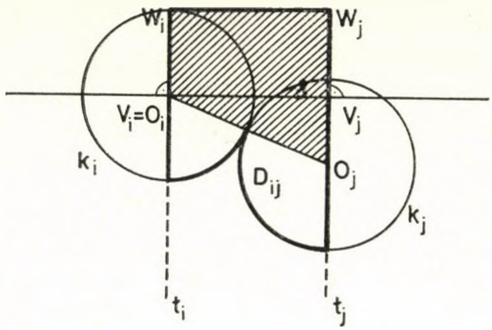


Abb. 8

Durch einfache Rechnung ergibt sich für den Flächeninhalt der Menge $T_{O_iO_jW_jW_i}=$ $=H_r^{V_iV_j}\cup O_iO_jV_j\Delta$:

$$T_{O_iO_jW_jW_i}=\operatorname{sh} r\cdot\operatorname{arch}\left(\frac{\operatorname{ch} 2r}{\operatorname{ch} x}\right)+2\operatorname{arctg}\left(\operatorname{th}\frac{x}{2}\cdot\sqrt{\frac{\operatorname{ch} 2r-\operatorname{ch} x}{\operatorname{ch} 2r+\operatorname{ch} x}}\right)=T(x).$$

Wir werden zeigen, daß $T(x)$ eine konkave Funktion im Intervall $[0, r]$ ist. Die Ableitung von $T(x)$ kann folgenderweise dargestellt werden.

$$T'(x)=\frac{-\operatorname{sh} r\operatorname{ch} 2r\operatorname{th} x}{\sqrt{\operatorname{ch}^2 2r-\operatorname{ch}^2 x}}+\frac{2\sqrt{\dfrac{\operatorname{ch} 2r-\operatorname{ch} x}{\operatorname{ch} 2r+\operatorname{ch} x}}}{1+\operatorname{th}^2\left(\dfrac{x}{2}\right)\dfrac{\operatorname{ch} 2r-\operatorname{ch} x}{\operatorname{ch} 2r+\operatorname{ch} x}}\left(\frac{1}{2\operatorname{ch}^2\left(\dfrac{x}{2}\right)}-\frac{2\operatorname{sh}^2\left(\dfrac{x}{2}\right)\operatorname{ch} 2r}{\operatorname{ch}^2 2r-\operatorname{ch}^2 x}\right).$$

Nach einiger Rechnung folgt für $0<x<r$:

$$\frac{1}{2\operatorname{ch}^2\left(\dfrac{x}{2}\right)}-\frac{2\operatorname{sh}^2\left(\dfrac{x}{2}\right)\operatorname{ch} 2r}{\operatorname{ch}^2 2r-\operatorname{ch}^2 x}>0;\qquad \frac{2\sqrt{\dfrac{\operatorname{ch} 2r-\operatorname{ch} x}{\operatorname{ch} 2r+\operatorname{ch} x}}}{1+\operatorname{th}^2\left(\dfrac{x}{2}\right)\dfrac{\operatorname{ch} 2r-\operatorname{ch} x}{\operatorname{ch} 2r+\operatorname{ch} x}}>0$$

$$\left(\frac{1}{2\operatorname{ch}^2\left(\dfrac{x}{2}\right)}-\frac{2\operatorname{sh}^2\left(\dfrac{x}{2}\right)\operatorname{ch} 2r}{\operatorname{ch}^2 2r-\operatorname{ch}^2 x}\right)'<0;\qquad \left(\frac{2\sqrt{\dfrac{\operatorname{ch} 2r-\operatorname{ch} x}{\operatorname{ch} 2r+\operatorname{ch} x}}}{1+\operatorname{th}^2\left(\dfrac{x}{2}\right)\dfrac{\operatorname{ch} 2r-\operatorname{ch} x}{\operatorname{ch} 2r+\operatorname{ch} x}}\right)'<0$$

$$\left(\frac{-\operatorname{sh} r\operatorname{ch} 2r\operatorname{th} x}{\sqrt{\operatorname{ch}^2 2r-\operatorname{ch}^2 x}}\right)'<0.$$

Deshalb gilt $T''(x)<0$ für $0<x<r$. Folglich ist $\inf\limits_{x\in[0,r]} T(x)=\min\{T(0);\ T(r)\}$.

HILFSSATZ 3. $T(0)<T(r)$.

BEWEIS. Der Einfachheit halber seien $T_{O_i^0O_j^0W_j^0W_i^0}=T(0)$ und $T_{O_i^rO_j^rW_j^rW_i^r}=T(r)$. Weiter werden wir annehmen, daß $O_i^0=O_i^r$, $W_i^0=W_i^r$ und der Hyperzyklusbogen

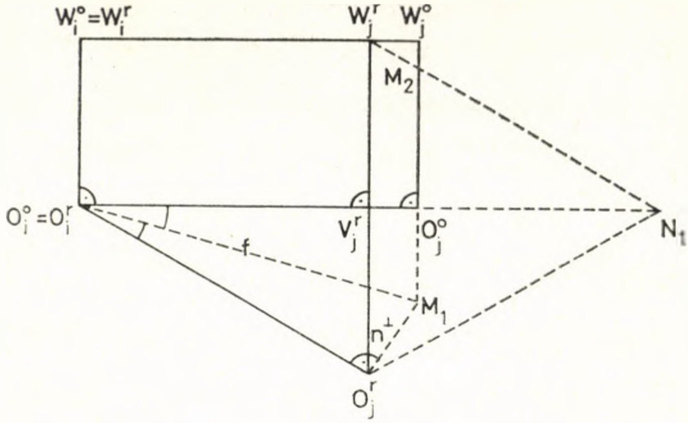


Abb. 9

$\widehat{W_i^r W_j^r}$ ein Teil des Hyperzyklusbogens $\widehat{W_i^0 W_j^0}$ ist (Abb. 9). Dann ist $V_j^r = O_j^r W_j^r \cap \cap O_i^r O_j^0$ und eine Halbdrehung um den Punkt V_j^r führt den Punkt O_i^r in einen Punkt N_1 über. Hier ist $\sphericalangle O_i^r O_j^r V_j^r = \frac{2\pi}{a}$, so gilt $\sphericalangle O_i^r O_j^r N_1 = \frac{4\pi}{a} > \frac{\pi}{2}$, denn wegen $r < r_0$ ist $a < 6{,}7$. Ferner sei f die innere Winkelhalbierende von $\sphericalangle O_j^0 O_i^0 O_j^r$ $(<\pi)$ und $n^\perp$ die Gerade, die den Punkt O_j^r enthält und die Gerade $O_i^r O_j^r$ senkrecht schneidet. Offenbar ist $f \cap n^\perp \neq \emptyset$, daher sei $M_1 = f \cap n^\perp$. Selbstverständlich liegt der Punkt M_1 im Inneren des Dreiecks $N_1 V_j^r O_j^r$ und gehört zur Geraden $O_j^0 W_j^0$. Schließlich sei $M_2 = O_j^0 W_j^0 \cap N_1 W_j^r$. Also gilt $\frac{r}{2} = \frac{O_j^r V_j^r}{2} < \frac{O_j^r O_j^0}{2} < O_j^0 M_1 < O_j^0 M_2$, woraus $O_j^0 M_2 > M_2 W_j^0$ folgt. So können wir aber den durch den Hyperzyklusbogen $\widehat{W_j^0 W_j^r}$ und die Strecken $W_j^0 M_2$, $W_j^r M_2$ begrenzten Bereich durch das rechtwinklige Dreieck $M_2 O_j^0 N_1$ überdecken. Somit ist der Beweis der Ungleichung $T(0) < T(r)$ beendet. Q.e.d.

Aus den obigen Überlegungen folgt also, daß $T_{O_i O_j W_j W_i} = T(x) \geq T(0) = 2r \operatorname{sh} r$ gilt. Anderseits ist aber

$$D_{ij} \geq \frac{k_{ij}}{2} + T_{O_i O_j W_j W_i}$$

folglich

$$\frac{k_{ij}}{D_{ij}} \leq \frac{4\pi \operatorname{sh}^2\left(\frac{r}{2}\right)}{2r \operatorname{sh} r + 2\pi \operatorname{sh}^2\left(\frac{r}{2}\right)} = \frac{2\pi}{4\frac{\frac{r}{2}}{\operatorname{th}\frac{r}{2}} + \pi} < \frac{2\pi}{4+\pi} < \frac{\pi}{\sqrt{12}}$$

was wir beweisen wollten.

LITERATURVERZEICHNIS

[1] FEJES TÓTH, L., *Lagerungen in der Ebene, auf der Kugel und im Raum*, Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete, Bd. 65, Springer-Verlag, Berlin—Göttingen—Heidelberg, 1953. *MR* **15**—248.

[2] MOLNÁR, J., Ausfüllung und Überdeckung eines konvexen sphärischen Gebietes durch Kreise I., *Publ. Math. Debrecen* **2** (1952), 266—275. *MR* **14**—788.

[3] MOLNÁR, J., Körelhelyezések állandó görbületű felületeken (Location of circles on surfaces of constant curvature), *Magyar Tud. Akad. Mat. Fiz. Oszt. Közl.* **12** (1962), 223—263 (in Hungarian). *MR* **28**#1535.

(Eingegangen am 11. September 1981)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
GEOMETRIAI TANSZÉK
RÁKÓCZI ÚT 5
H—1088 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 367—378

STABILITY OF SOME PROBLEMS OF CHARACTERIZATION OF THE NORMAL AND RELATED DISTRIBUTIONS

L. B. KLEBANOV and J. A. MELAMED

1. Let $\mathscr{F}_\mu$ be a class of distribution functions (d.f.) $F(x)$, $x=(x_1, \ldots, x_p)\in R^p$, satisfying the following conditions:

(a_1) $F(x)$ has a density function $f(x)$ which is absolutely continuous in each variable;

(b_1) $\int_{R^p} x_p f(x)\,dx=0$, $\int_{R^p} x_l x_j f(x)\,dx=\alpha_{lj}$, $\alpha_{ll}=\alpha_l^2<\infty$, $l,j=1, \ldots, p$ (here and further on $dx=dx_1\ldots dx_p$), and $\mu=\|\alpha_{lj}\|$ is a non-singular matrix;

(c_1) $J_l=J_l(x)=-\dfrac{\partial f(x)}{\partial x_l}\Big/ f(x)\in L_f^2$, $l=1, \ldots, p$ (as usually, L_f^2 denotes the Hilbert space of functions which are square integrable with respect to the measure of density function $f(x)$).

Put $I_{lj}=\int_{R^p} J_l(x)J_j(x)f(x)\,dx$, $l,j=1, \ldots, p$. Let $M_1\subset L_f^2$ be a subspace of linear functions of x. Define the values $J_l^{(1)}$, $l=1, \ldots, p$, by

$$J_l^{(1)} = J_l^{(1)}(x) = \hat{E}(J_l\,|\,M_1),$$

where $\hat{E}$ is a projection operator in L_f^2, and put

$$I_{lj}^{(1)} = \int_{R^p} J_l^{(1)}(x)J_j^{(1)}(x)f(x)\,dx, \qquad l,j = 1, \ldots, p.$$

It is well-known that the matrix $\mathbf{J}=\|I_{lj}\|$ represents the matrix of Fisher information on the vector parameter $\theta=(\theta_j, \ldots, \theta_p)\in R^p$, contained in an observation over the population. $\{f(x-\theta), \theta\in R^p\}$. Similarly, $\mathbf{J}^{(1)}=\|I_{lj}^{(1)}\|$ is the matrix of Fisher information contained in linear functions of an observation over the population $\{f(x-\theta), \theta\in R^p\}$.

It is easy to see that $\mathbf{J}-\mathbf{J}^{(1)}$ is a positive semidefinite matrix. We are going to investigate for which distributions this matrix is the zero matrix:

$$\mathbf{J}-\mathbf{J}^{(1)} = \mathbf{O}. \tag{1}$$

THEOREM 1. *Let $F\in\mathscr{F}_\mu$, then* (1) *holds if and only if $F(x)$ is a d.f. of the normal law with zero vector of mean values and covariance matrix μ.*

1980 *Mathematics Subject Classification.* Primary 62E10; Secondary 62B10, 62F12.

Key words and phrases. Characterization, stability of characterization, normal and related distributions, location and scale parameters, equivariant estimator, asymptotic optimality.

PROOF. Equality (1) is equivalent to the relations

$$I_{lj} = I_{lj}^{(1)}, \qquad l, j = 1, \ldots, p,$$

which, in turn, are equivalent to the equalities (in L_f^2)

$$J_l(x) = J_l^{(1)}(x), \qquad l = 1, \ldots, p,$$

i.e.

$$-\frac{\partial f(x)}{\partial x_l} \Big/ f(x) = \sum_{j=1}^{p} a_{lj} x_j, \qquad l = 1, \ldots, p. \tag{2}$$

The solution of the system of equations (2) is the density function of the normal distribution with zero vector of mean values and covariance matrix μ. Direct calculation shows that (1) holds for the mentioned normal distribution. The proof is complete.

Let us see now when the matrices of information $\mathbf{J}$ and $\mathbf{J}^{(1)}$ are ε^2-close, i.e.

$$|I_{lj} - I_{lj}^{(1)}| \leqq \varepsilon^2, \qquad l, j = 1, \ldots, p. \tag{3}$$

Introduce the class $\mathscr{F}_{\varepsilon, \mu}$ of all d.f. $F(x)$, $x \in R^p$, satisfying the conditions

(d_1) $F \in \mathscr{F}_\mu$ and F satisfies (3);

(e_1) let $\Delta(x) = [x_1, x_1 + y_1, \ldots, x_p, x_p + y_p]$ be a rectangle in R^p, then for sufficiently small $|y_j|$, $j = 1, \ldots, p$, the functions $\sup_{z \in \Delta(x)} \left| \frac{\partial f(z)}{\partial z_l} \right|$, $l = 1, \ldots, p$, are integrable and

$$\int_{R^p} \sup_{z \in \Delta(x)} \left| \frac{\partial f(x)}{\partial z_l} \right| dx \leqq M$$

(where M may depend on μ).

For every d.f. F and G let $\varrho_1(F, G)$ be the distance in variation between F and G. Denote by $\Phi_\mu(x)$ the d.f. of the normal law with zero mean vector and covariance matrix μ.

THEOREM 2. *There is a constant* $C = C(\mu, p, M)$ *depending only on* μ, p *and* M *such that for sufficiently small* $\varepsilon > 0$ *the inequality*

$$\sup_{F \in \mathscr{F}_\mu} \varrho_1(F, \Phi_\mu) \leqq C \left(\log \frac{1}{\varepsilon} \right)^{-\frac{1}{2p+1}} \tag{4}$$

holds.

PROOF. Relations (3) imply

$$J_l(x) - J_l^{(1)}(x) = R_l(x), \qquad l = 1, \ldots, p, \tag{5}$$

where

$$\int_{R^p} R_l(x) f(x)\, dx \leqq \varepsilon^2, \qquad l = 1, \ldots, p. \tag{6}$$

Rewrite relations (5) in the form

$$-\frac{\partial f(x)}{\partial x_l} = \sum_{j=1}^{p} a_{lj} x_j f(x) + R_l(x) f(x), \qquad l = 1, \ldots, p, \tag{7}$$

where $\mathscr{A}=\|a_{lj}\|=\mu^{-1}$. Denote by $\varphi(t)$ the characteristic function (c.f.) of $F(x)$, i.e.

$$\varphi(t)=\int_{R^p}\exp\left\{i\sum_{j=1}^{p}t_jx_j\right\}f(x)\,dx \qquad (t=(t_1,\ldots,t_p)\in R^p),$$

and put

$$\varrho_j(t)=-i\int_{R^p}\exp\left\{i\sum_{l=1}^{p}t_lx_l\right\}R_j(x)f(x)\,dx, \qquad j=1,\ldots,p.$$

Then relations (6) and (7) imply

$$\sum_{j=1}^{p}a_{lj}\frac{\partial\varphi(t)}{\partial t_j}+t_l\varphi(t)=\varrho_l(t), \qquad l=1,\ldots,p,$$

or

$$\frac{\partial\varphi(t)}{\partial t_l}+\sum_{j=1}^{p}\alpha_{lj}t_j\varphi(t)=\varrho_l^*(t), \qquad l=1,\ldots,p, \tag{8}$$

where

$$|\varrho_l^*(t)|\leqq A_l\varepsilon, \qquad l=1,\ldots,p,$$

and $A_l=\left|\sum_{j=1}^{p}a_{lj}\right|$. If we fix all variables t_j with $j\neq l$ in the l-th $(l=1,\ldots,p)$ equation of system (8) we obtain an ordinary linear differential equation with respect to $\varphi(t_1,\ldots,t_l,\ldots,t_p)$, considered as a function of t_l. Integrating each of these equations and making use of the probability meaning of the function φ $(\varphi(0,\ldots,0,\ldots,0)=1)$, we find that the solution of system (8) is the c.f.

$$\begin{aligned}\varphi(t)={}&\exp\left\{-\frac{1}{2}\alpha_{11}t_1^2-\sum_{j=2}^{p}\alpha_{1j}t_1t_j\right\}\int_0^{t_1}\exp\left\{\frac{1}{2}\alpha_{11}\tau^2+\sum_{j=2}^{p}\alpha_{1j}t_j\tau\right\}\varrho_1^*(\tau,t_2,\ldots,t_p)d\tau+\\&+\exp\left\{-\frac{1}{2}\alpha_{11}t_1^2-\sum_{j=1}^{p}\alpha_{1j}t_1t_j\right\}\exp\left\{-\frac{1}{2}\alpha_{22}t_2^2-\sum_{j=3}^{p}\alpha_{2j}t_2t_j\right\}\times\\&\int_0^{t_2}\exp\left\{\frac{1}{2}\alpha_{22}\tau^2+\sum_{j=3}^{p}\alpha_{2j}t_j\tau\right\}\varrho_2^*(0,\tau,t_3,\ldots,t_p)\,d\tau+\ldots\\&\ldots+\exp\left\{-\frac{1}{2}\alpha_{11}t_1^2-\sum_{j=1}^{p}\alpha_{1j}t_1t_j\right\}\exp\left\{-\frac{1}{2}\alpha_{22}t_2^2-\sum_{j=3}^{p}\alpha_{2j}t_2t_j\right\}\times\ldots\times\exp\left\{-\frac{1}{2}\alpha_{pp}t_p^2\right\}\times\\&\int_0^{t_p}\exp\left\{\frac{1}{2}\alpha_{pp}\tau^2\right\}\varrho_p^*(0,\ldots,0,\tau)\,d\tau+\exp\left\{-\frac{1}{2}\sum_{1}^{p}\alpha_{jj}t_j^2-\sum_{l>j=1}^{p}\alpha_{lj}t_lt_j\right\}.\end{aligned}\tag{9}$$

It is well-known that the c.f. of the normal d.f. $\Phi_\mu(x)$ has the form

$$\psi(t)=\exp\left\{-\frac{1}{2}\sum_{1}^{p}\alpha_{jj}t_j^2-\sum_{l>j=1}^{p}\alpha_{lj}t_lt_j\right\}.$$

Hence, from (9) we obtain

$$|\varphi(t)-\psi(t)| \leqq \varepsilon\Big[A_1 \exp\Big\{-\frac{1}{2}\alpha_{11}t_1^2-\sum_{j=2}^{p}\alpha_{1j}t_1t_j\Big\}\int_0^{t_1}\exp\Big\{\frac{1}{2}\alpha_{11}\tau^2+\sum_{j=2}^{p}\alpha_{1j}t_j\tau\Big\}d\tau+$$

(10)

$$+\ldots+A_p\exp\Big\{-\frac{1}{2}\sum_{1}^{p}\alpha_{jj}t_j^2-\sum_{l>j=1}^{p}\alpha_{lj}t_lt_j\Big\}\int_0^{t_p}\exp\Big\{\frac{1}{2}\alpha_{pp}\tau^2\Big\}d\tau\Big] = \varepsilon g(t_1, \ldots, t_p).$$

On the other hand, the c.f. $\psi(t)$ represents the solution of the homogeneous system of differential equations corresponding to system (8). Therefore

$$\frac{\partial}{\partial t_l}(\varphi(t)-\psi(t)) = \sum_{j=1}^{p}\alpha_{lj}t_j(\varphi(t)-\psi(t))+\varrho_l^*(t), \qquad l = 1, \ldots, p.$$

In view of (10) we find

$$(11)\qquad \Big|\frac{\partial}{\partial t_l}(\varphi(t)-\psi(t))\Big| \leqq \Big[\sum_{j=1}^{p}|\alpha_{lj}|\,|t_j|\,g(t_1, \ldots, t_p)+A_l\Big]\varepsilon, \qquad l = 1, \ldots, p.$$

Consider the value

$$\omega(F;\delta) = \sup_{A\in\mathfrak{A}}\ \sup_{|y_j|\leqq\delta_j,\, j=1,\ldots,p}\Big|\int_A dF(x)-\int_{\{A-y\}}dF(x)\Big|,$$

where $\mathfrak{A}$ is the σ-algebra of Borel subsets of R^p,

$$\delta = (\delta_1, \ldots, \delta_p),\quad \delta_j > 0,\quad j = 1, \ldots, p,\quad \text{and}\quad \{A-y\} = \{x\colon x = z-y,\ z\in A\}.$$

We have

$$\Big|\int_A dF(x)-\int_{\{A-y\}}dF(x)\Big| = \Big|\int_A[f(x-y)-f(x)]\,dx\Big| \leqq \sum_{j=1}^{p}|y_j|\int_A \sup_{z\in\Delta(x)}\Big|\frac{\partial f(z)}{\partial z_j}\Big|\,dx.$$

From here and from condition (e_1) it follows that

$$(12)\qquad \omega(F;\delta) \leqq M\sum_{j=1}^{p}\delta_j.$$

Similarly, we conclude that

$$(13)\qquad \omega(\Phi_\mu;\delta) \leqq C(\mu)\sum_{j=1}^{p}\delta_j.$$

To estimate the closeness of the d.f. F and Φ_μ, we use Theorem 1 from [1]. According to this theorem, for any $T_j>1$, $\delta=(\delta_1, \ldots, \delta_p)$, $\delta_j>0$, $j=1, \ldots, p$, $r>p$ ($r\geqq 2$) the inequality

$$\varrho_1(F, \Phi_\mu) \leqq \omega(F;\delta)+\omega(\Phi_\mu;\delta)+C_{p,r}\Big[1+\sum_{j=1}^{p}\Big(\frac{2}{T_j}+\delta_j\Big)\Big]\times$$

$$\times\Big[\int_{T^{(p)}}|\varphi(t)-\psi(t)|^r\,dt\Big]^{\frac{1}{r}}+C_{p,r}\sum_{j=1}^{p}\Big[\int_{T^{(p)}}\Big|\frac{\partial}{\partial t_j}[\varphi(t)-\psi(t)]\Big|^r dt\Big]^{\frac{1}{r}}+$$

$$+\hat{C}(1+3^p)/\Big(\prod_{j=1}^{p}\delta_j^2\Big)\sum_{1}^{p}T_j^{-2}$$

holds, where $\hat{C}$ is an absolute constant, $T^{(p)}=\{t\colon |t_j|\leqq T_j, j=1, \ldots, p\}$, and

$$C_{p,r}=(2p)^{(2-s-ps)/(2s)}\left(\int\limits_{R^p}\frac{dx}{\left|1+i\left(\sum\limits_1^p x_j^2\right)^{1/2}\right|}\right)^{1/2}, \qquad \frac{1}{s}+\frac{1}{r}=1.$$

Put $T_j=T$, $\delta_j=\delta_0$, $j=1, \ldots, p$. Then the inequalities (10)—(14) imply for any fixed $r>p$ $(r\geqq 2)$

$$\varrho_1(F, \Phi_\mu)\leqq C[\delta_0+e^{\lambda T^2}\varepsilon T^2+e^{\lambda T^2}\varepsilon T^{3+1/r}+\delta_0^{-2p}T^{-2}],$$

where $C=C(\mu, M, r, p)$ and $\lambda>0$ is some constant depending on μ.

If we choose here $T=\gamma\left(\log\frac{1}{\varepsilon}\right)^{\frac{1}{2}}$ with $\gamma>0$ and such that $\lambda\gamma^2<1$ and $\delta_0=\left(2p\Big/\left(\gamma\log\frac{1}{\varepsilon}\right)\right)^{\frac{1}{2p+1}}$, we obtain

$$\varrho_1(F, \Phi_\mu)\leqq C\left(\log\frac{1}{\varepsilon}\right)^{-\frac{1}{2p+1}}.$$

The proof of Theorem 2 is complete.

REMARK 1. If we consider Levi's distance $L(F, \Phi_\mu)$ instead of the distance in variation, then the abovementioned argument together with Theorem 2 of [1] shows that for sufficiently small $\varepsilon>0$

$$\sup_{F\in\tilde{\mathscr{F}}_{\varepsilon,\mu}} L(F, \Phi_\mu)\leqq C\left(\log\frac{1}{\varepsilon}\right)^{-\frac{1}{4}},$$

where $\tilde{\mathscr{F}}_{\varepsilon,\mu}$ is a class of d.f. $F(x)$ satisfying conditions $(\mathrm{a}_1)-(\mathrm{d}_1)$.

REMARK 2. Under the additional assumptions of the existence of all moments and the completeness of polynomials in L_f^2, property (1) is equivalent to the asymptotic optimality (with respect to the matrix quadratic loss function[1]) in the class of equivariant estimators of the estimator $\bar{X}=\left(\frac{1}{n}\sum\limits_1^n X_j^{(1)}, \ldots, \frac{1}{n}\sum\limits_1^n X_j^{(p)}\right)$ (based on the random sample $X_j=(X_j^{(1)}, \ldots, X_j^{(p)}), j=1, \ldots, n$, from a population with the d.f. $F(x-\theta)$) of the vector parameter θ of location. Property (3) under the same additional assumptions is equivalent to the asymptotic ε^2-optimality of $\bar{X}$ in the class of equivariant estimators.

2. Consider now a set of univariate d.f. $F(x)$ $(x\in R^1)$ satisfying the following conditions:

[1] I. e. the risk of an estimator $\tilde{\theta}$ of θ is defined as $E_\theta(\tilde{\theta}-\theta)^T(\tilde{\theta}-\theta)$, where E_θ is the sign of mathematical expectation under the value θ of the parameter, and T is the transposition sign.

(a_2) $F(x)$ has a symmetric absolutely continuous density function $f(x)$;

(b_2) $$\int_{-\infty}^{\infty} x^2 f(x)\,dx = \alpha_2; \quad \int_{-\infty}^{\infty} x^4 f(x)\,dx = \alpha_4 < \infty;$$

(c_2) $$J_1 = J_1(x) = -\frac{f'(x)}{f(x)} \in L_f^2, \quad J_2 = J_2(x) = -\left(1 + x\frac{f'(x)}{f(x)}\right) \in L_f^2.$$

Put $\tilde{J}_2 = J_2 - \hat{E}(J_2|J_1)$. The value $\tilde{I}_2 = \int_{-\infty}^{\infty} \tilde{J}_2^2(x) f(x)\,dx$ is the Fisher information on the scale parameter in presence of the nuisance location parameter, contained in an observation over the population $\left\{\frac{1}{\sigma} f\left(\frac{x-\theta}{\sigma}\right),\ \theta \in R^1,\ \sigma \in R_+^1\right\}$.

Let $M_2 \subset L_f^2$ be a space of polynomials of degree $\leqq 2$ and $M_2^{(1)} \subset M_2$ be a subspace of M_2 orthogonal to J_1. Put $\tilde{J}_2^{(2)} = \hat{E}(\tilde{J}_2|M_2^{(1)})$. The value $\tilde{I}_2^{(2)} = \int_{-\infty}^{\infty} \tilde{J}_2^{(1)2}(x) f(x)\,dx$ measures Fisher's information on the parameter σ in presence of the nuisance parameter θ, contained in polynomials of degree $\leqq 2$ of an observation over the population $\left\{\frac{1}{\sigma} f\left(\frac{x-\theta}{\sigma}\right),\ \theta \in R^1,\ \sigma \in R_+^1\right\}$. For further details of the values $\tilde{I}_2$ and $\tilde{I}_2^{(2)}$ see [2].

Let $x_1, \ldots, x_n$ be a random sample from a population with a d.f. $F\left(\frac{x-\theta}{\sigma}\right)$. Denote $\bar{x} = \frac{1}{n}\sum_1^n x_j$ and $s^2 = \frac{1}{n}\sum_1^n (x_j - \bar{x})^2$. It was shown in [3] that, under certain assumptions on the d.f. $F(x)$, asymptotic optimality in the class of equivariant estimators of the estimator s^2 of a parameter function $\alpha_2\sigma^2$ is equivalent to the equality

(15) $$\tilde{I}_2 = \tilde{I}_2^{(2)}.$$

It was shown in [3] that for a d.f. satisfying the conditions (a_2)—(c_2) relation (15) holds iff

(16) $$f(x) = g(x; \alpha_2, \alpha_4) = \frac{a^{p/2}}{\Gamma\left(\frac{p}{2}\right)} |x|^{p-1} e^{-ax^2}, \qquad x \in R^1,$$

with

$$a = \frac{\alpha_2}{\alpha_4 - \alpha_2^2}, \quad p = \frac{2\alpha_2^2}{\alpha_4 - \alpha_2^2}.$$

In [3] the question was raised when the quantities of Fisher information $\tilde{I}_2$ and $\tilde{I}_2^{(2)}$ are ε-close. However, there is an error in the proof of Theorem 3 in [3]. We bring here a correct proof which, at the same time, improves the order of closeness of $F(x)$ and the d.f. $G_{\alpha_2,\alpha_4}(x)$ given by the density function (16).

Consider a class $\mathcal{F}_\varepsilon = \mathcal{F}_\varepsilon(\alpha_2, \alpha_4)$ of d.f. $F(x)$ satisfying conditions (a_2)–(c_2) and also

(17) $$\tilde{I}_2 - \tilde{I}_2^{(2)} \leqq \varepsilon^2.$$

THEOREM 3. *There is a constant* $C=C(\alpha_2, \alpha_4)$ *depending only on the moments* α_2 *and* α_4 *such that for* $\varepsilon\in(0, 1)$ *the inequality*

(18) $$\sup_{F\in\mathcal{F}_\varepsilon} \varrho_1(F, G_{\alpha_2,\alpha_4}) \leqq C\varepsilon$$

holds.

PROOF. It was shown in [3] that for $F\in\mathcal{F}_\varepsilon$ the relation (17) implies

(19) $$xf'(x)+[2ax^2-(p-1)]f(x) = r(x)f(x)$$

with

$$a = \frac{\alpha_2}{\alpha_4-\alpha_2^2}, \quad p = \frac{2\alpha_2^2}{\alpha_4-\alpha_2^2}$$

and

$$\int_{-\infty}^{\infty} r^2(x)f(x)\,dx \leqq \varepsilon^2.$$

Let $\varphi(t)$ and $\psi(t)$ be the c.f. of the distributions $F(x)$ and $G_{\alpha_2,\alpha_4}(x)$, respectively. Then (19) implies

(20) $$2a\varphi''(t)+t\varphi'(t)+p\varphi(t) = R(t; \varepsilon),$$

where $|R(t; \varepsilon)|\leqq\varepsilon$. By the definition of $g(x; \alpha_2, \alpha_4)$ we have

(21) $$2a\psi''(t)+t\psi'(t)+p\psi(t) = 0.$$

Put $\xi(t)=\varphi(t)-\psi(t)$. Relations (20) and (21) define a differential equation for $\xi(t)$:

(22) $$2a\xi''(t)+t\xi'(t)+p\xi(t) = R(t; \varepsilon)$$

with initial conditions $\xi(0)=0$, $\xi'(0)=0$. Let $\xi_1^{(0)}(t)$ and $\xi_2^{(0)}(t)$ be linearly independent solutions of the homogeneous differential equation corresponding to (22). Then the solution $\xi(t)$ $(\xi(0)=0,\ \xi'(0)=0)$ of (22) can be written in the form of

(23) $$\xi(t) = A(t)\xi_1^{(0)}(t)+B(t)\xi_2^{(0)}(t),$$

where

$$A(t) = -\int_0^t \frac{R(\tau; \varepsilon)\xi_2^{(0)}(\tau)}{W(\tau)}\,d\tau, \quad B(t) = \int_0^t \frac{R(\tau; \varepsilon)\xi_1^{(0)}(t)}{W(\tau)}\,d\tau,$$

(24)

$$W(t) = \begin{vmatrix} \xi_1^{(0)}(t) & \xi_2^{(0)}(t) \\ \xi_1^{(0)\prime}(t) & \xi_2^{(0)\prime}(t) \end{vmatrix} = Ce^{-\frac{t^2}{4a}}$$

(see e.g. [4, pp. 144, 95]). In order to analyse the behaviour of the solutions $\xi_1^{(0)}(t)$ and $\xi_2^{(0)}(t)$ as $t\to\infty$, substitute

$$\xi_i^{(0)}(t) = e^{-\frac{t^2}{8a}}\,y_i(t), \qquad i = 1, 2.$$

It is clear that the functions $y_i(t)$, $i=1, 2$, represent linearly independent solutions of the differential equation

$$y''(t) = \left(\frac{t^2}{16a^2}-\frac{2p-1}{4a}\right)y(t).$$

For such solutions the following asymptotics holds as $t\to\infty$:

$$y_{1,2}(t)\sim\frac{C_{1,2}}{\sqrt{t}}\exp\left\{\pm\frac{t^2}{8a}\right\} \tag{25}$$

(see, e.g., item 1) of Proposition 1 of [5]), so that

$$\xi_1^{(0)}(t)\underset{t\to\infty}{\simeq}\frac{C_1}{\sqrt{t}}e^{-\frac{t^2}{4a}},\quad \xi_2^{(0)}(t)\underset{t\to\infty}{\simeq}\frac{C_2}{\sqrt{t}}, \tag{26}$$

and also these asymptotic formulas can be differentiated any number of times (see [5]). The relations (23), (24), (26) and the inequality $|R(t;\varepsilon)|\leqq\varepsilon$ imply

$$|\varphi(t)-\psi(t)|\leqq C\varepsilon. \tag{27}$$

By differentiating the relations (23) and (26), it is easy to see that

$$|\varphi'(t)-\psi'(t)|\leqq C\varepsilon. \tag{28}$$

It follows from [1] that

$$\varrho_1(F,G_{\alpha_2,\alpha_4})\leqq\omega(F;\delta)+\omega(G_{\alpha_2,\alpha_4};\delta)+\hat{C}\delta^{-2}T^{-2}+$$
$$+C_r(1+2T^{-1}+\delta)\Big(\int_{-T}^{T}|\varphi(t)-\psi(t)|^r\,dt\Big)^{1/r}+C_r\Big(\int_{-T}^{T}|\varphi'(t)-\psi'(t)|^r\,dt\Big)^{1/r},$$

where

$$C_r=(2\pi)^{\frac{1-s}{s}}\Big(\int_{-\infty}^{\infty}\frac{dx}{|1+i|x||^r}\Big)^{\frac{1}{r}},\quad \frac{1}{r}+\frac{1}{s}=1,\quad r\geqq 2,\quad T>1,\quad \delta>0.$$

In view of estimates (27), (28), the inequality for $\varrho_1(F,G_{\alpha_2,\alpha_4})$ can be written in the form

$$\varrho_1(F,G_{\alpha_2,\alpha_4})\leqq\omega(F;\delta)+\omega(G_{\alpha_2,\alpha_4};\delta)+\hat{C}\delta^{-2}T^{-2}+C\varepsilon C_r(2T)^{1/r}(2+2T^{-1}+\delta).$$

It is easy to see that $C_r\xrightarrow[r\to\infty]{}1$. Therefore

$$\varrho_1(F,G_{\alpha_2,\alpha_4})\leqq\omega(F;\delta)+\omega(G_{\alpha_2,\alpha_4};\delta)+\hat{C}\delta^{-2}T^{-2}+C\varepsilon(2+2T^{-1}+\delta).$$

Passing to the limit in this inequality as $T\to\infty$ and afterwards as $\delta\to 0$, we find that

$$\varrho_1(F,G_{\alpha_2,\alpha_4})\leqq C\varepsilon.$$

In this connection we make use of the relations $\omega(F;\delta)\xrightarrow[\delta\to 0]{}0$ and $\omega(G_{\alpha_2,\alpha_4};\delta)\xrightarrow[\delta\to 0]{}0$, which follow from global continuity in L of the density functions $f(x)$ and $g(x;\alpha_2,\alpha_4)$ (see, e.g. [6, p. 14]).

REMARK 3. The obtained result allows to make Theorem 4 of [3] more precise. This theorem establishes the stability of characterizations of the d.f. $G_{\alpha_2,\alpha_4}(x)$ by the property of asymptotic optimality of s^2 in the class of equivariant estimators of a parametric function $\alpha_2\sigma^2$.

Under the assumptions of this theorem the asymptotic ε^2-optimality of s^2 implies

$$\varrho_1(F,G_{\alpha_2,\alpha_4})\leqq C\varepsilon.$$

REMARK 4. Using the method of [7] it is easy to show that the inequality (18) cannot be improved in the order, i.e. there are constants C_1 and C_2 (depending only on α_2 and α_4) for which

$$C_1\varepsilon \leqq \sup_{F\in\mathcal{F}_2} \varrho_1(F, G_{\alpha_2,\alpha_4}) \leqq C_2\varepsilon.$$

3. Now let a family of univariate d.f. $F(x)$ be given, satisfying the following conditions:

(a_3) $F(x)$ has an absolutely continuous density function $f(x)$;

(b_3) $$J_1\in L_f^2,\quad J_2\in L_f^2\left(J_1=-\frac{f'(x)}{f(x)},\quad J_2=-\left(1+x\frac{f'(x)}{f(x)}\right)\right);$$

(c_3) $$\int_{-\infty}^{\infty} xf(x)\,dx=0,\qquad \int_{-\infty}^{\infty} x^2f(x)\,dx=\alpha_2,\qquad \int_{-\infty}^{\infty} x^4f(x)\,dx=\alpha_4<\infty.$$

Put $\tilde{J}_1=J_1-\hat{E}(J_1|J_2)$. The value $\tilde{I}_1=\int_{-\infty}^{\infty}\tilde{J}_1^2(x)f(x)\,dx$ is the Fisher information on the location parameter θ in presence of the nuisance scale parameter σ, contained in an observation over the population $\left\{\frac{1}{\sigma}f\left(\frac{x-\theta}{\sigma}\right),\ \theta\in R^1,\ \sigma\in R_+^1\right\}$. Let $M_1\subset L_f^2$ be a space of linear functions of x and let $\tilde{J}_1^{(1)}=\hat{E}(\tilde{J}_1|M_1)$. The value $\tilde{I}_1^{(1)}=$ $=\int_{-\infty}^{\infty}\tilde{J}_1^{(1)2}(x)f(x)\,dx$ measures Fisher information on the location parameter θ in presence of the nuisance scale parameter σ, contained in linear functions of an observation over the mentioned population. For further details on these information characteristics see [2].

Characterization of the normal and gamma distributions by the property $\tilde{I}_1=\tilde{I}_1^{(1)}$ was obtained in [8]. Here we investigate the question of coincidence of the information quantities

$$\tilde{I}_1=\tilde{I}_1^{(1)},\quad \tilde{I}_2=\tilde{I}_2^{(2)}. \tag{29}$$

THEOREM 4. *Let the d.f. $F(x)$ satisfy the conditions* (a_3)—(c_3). *Equalities* (29) *hold iff $F(x)$ is a d.f. of the normal law with parameters* 0 *and* α_2.

PROOF. Assume that the relations (29) are true. Then we easily deduce that the density function $f(x)$ satisfies the equations (see [8], [3])

$$\begin{cases} f'(x)-\dfrac{Ax-a}{ax-1}f(x)=0, \\ f'(x)-\dfrac{B_1x^2+B_2+1}{b-x}f(x)=0, \end{cases} \tag{30}$$

where

$$A=\frac{1}{\alpha_2},\quad a=\frac{I_{12}}{I_{22}},\quad b=\frac{I_{12}}{I_{11}}\Big(I_{lj}=\int_{-\infty}^{\infty}J_l(g)J_j(x)f(x)\,dx,\quad l,j=1,2\Big),$$

$$B_1=\frac{2\alpha_2}{\alpha_4-\alpha_2^2},\quad B_2=-\frac{2\alpha_2^2}{\alpha_4-\alpha_2^2}.$$

Comparing the equations of system (30) we conclude that

$$a=0,\quad b=0,\quad A=B_1,\quad B_2=-1.$$

Hence,

$$f'(x)=-\frac{x}{\alpha_2}f(x),$$

i.e. $f(x)$ is a density function of the normal distribution with parameters 0 and α_2.

A direct calculation shows that for the normal density function with parameters 0 and α_2

$$\tilde{J}_1=\frac{x}{\alpha_2},\quad \tilde{J}_2=\frac{x^2}{\alpha_2}-1,$$

so that the relations (29) are fulfilled. The proof is complete.

Next we study the stability of the obtained characterization of the normal law. Denote by $\Phi_{\alpha_2}(x)$ a d.f. of the normal law with zero mean and variance α_2. Let $\mathscr{H}_\varepsilon=\mathscr{H}_\varepsilon(\alpha_2,\alpha_4)$ be a set of d.f. $F(x)$ satisfying conditions (a_3)—(c_3) and, in addition, the conditions

$$\tilde{I}_1-\tilde{I}_1^{(1)}<\varepsilon^2,\quad \tilde{I}_2-\tilde{I}_2^{(2)}<\varepsilon^2. \tag{31}$$

THEOREM 5. *There is a constant $C=C(\alpha_2)$ depending only on α_2 such that for any* $\varepsilon\in(0,1)$

$$\sup_{F\in\mathscr{H}_\varepsilon}\sup_x|F(x)-\Phi_{\alpha_2}(x)|\leq C\varepsilon.$$

PROOF. Inequalities (31) imply the validity of the system of differential equations

$$\begin{aligned}&-\frac{f'(x)}{f(x)}+a\Big(1+x\frac{f'(x)}{f(x)}\Big)-Ax=r_1(x),\\ &-\Big(1+x\frac{f'(x)}{f(x)}\Big)+b\frac{f'(x)}{f(x)}-B_1x^2-B_2=r_2(x),\end{aligned} \tag{32}$$

with the same a, b, A, B_1 and B_2 as in (30) and

$$\int_{-\infty}^{\infty}r_j^2(x)f(x)\,dx<\varepsilon^2,\qquad j=1,2. \tag{33}$$

We easily deduce from the relations (32) and (33) that

$$-\frac{f'(x)}{f(x)}-\frac{1}{\alpha_2}x=r_1(x). \tag{34}$$

Equation (34) was investigated in the proof of Theorem 2 of [8]. It was shown that (34) implies the relation

$$\sup_x |F(x)-\Phi_{\alpha_2}(x)| \leqq C(\alpha_2)\varepsilon.$$

The proof is complete.

The theorem just proved can be applied for investigation of the asymptotic properties of the estimators of location and scale parameters.

Let $x_1, \ldots, x_n$ be a random sample from a population with the d.f. $F\left(\frac{x-\theta}{\sigma}\right)$. Assume that $F(x)$ satisfies the conditions (a_3)—(c_3). Denote by $(\tilde{\theta}_n, \tilde{\sigma}_n)$ some equivariant estimator of (θ, σ) and define the risk of this estimator as

$$\mathsf{R}(\tilde{\theta}_n, \tilde{\sigma}_n) = \frac{1}{\sigma^2}\mathsf{E}_{\theta,\sigma}[(\tilde{\theta}_n-\theta)^2+(\tilde{\sigma}_n-\sigma)^2],$$

where $\mathsf{E}_{\theta,\sigma}$ is the mathematical expectation corresponding to the d.f. $F\left(\frac{x-\theta}{\sigma}\right)$.

We say that the estimator $\left(\bar{x}, \frac{s}{\alpha_2^{1/2}}\right)$ is asymptotically ε^2-optimal in the class of equivariant estimators, if for any equivariant estimator $(\tilde{\theta}_n, \tilde{\sigma}_n)$ the limit relation

$$\varlimsup_{n\to\infty} n\left[\mathsf{R}\left(\bar{x}, \frac{s}{\alpha_2^{1/2}}\right)-\mathsf{R}(\tilde{\theta}_n, \tilde{\sigma}_n)\right] \leqq \varepsilon^2$$

holds. Denote by M_∞ the closure of the space of all polynomials of an observation in the L_f^2 metric.

Theorem 6. *Let the d.f. $F(x)$ have moments of all orders, satisfy the conditions (a_3)—(c_3) and, in addition, $J_j \in M_\infty$, $\bar{I}_j \leqq T$, $j=1, 2$. Then the asymptotic ε^2-optimality of the estimator $\left(\bar{x}, \frac{s}{\alpha_2^{1/2}}\right)$ in the class of equivariant estimators implies the relation*

$$\sup_x |F(x)-\Phi_{\alpha_2}(x)| \leqq C(\alpha_2, \alpha_4, T)\varepsilon,$$

where the constant $C(\alpha_2, \alpha_4, T)$ depends only on α_2, α_4 and T.

Proof. Under the assumptions of Theorem 6 it follows from the result of [2] that asymptotic ε^2-optimality of $\left(\bar{x}, \frac{s}{\alpha_2^{1/2}}\right)$ is equivalent to the inequality

$$\frac{1}{\bar{I}_1^{(1)}}-\frac{1}{\bar{I}_1}+\frac{1}{\bar{I}_2^{(2)}}-\frac{1}{\bar{I}_2} \leqq \varepsilon^2,$$

which implies the inequalities

(35) $$\bar{I}_1-\bar{I}_1^{(1)} \leqq \bar{I}_1^{(1)}\bar{I}_1\varepsilon^2, \quad \bar{I}_2-\bar{I}_2^{(2)} \leqq \bar{I}_2^{(2)}\bar{I}_2\varepsilon^2.$$

Since $\bar{I}_1^{(1)}$ and $\bar{I}_2^{(2)}$ depend only on α_2, α_4 and $\bar{I}_j \leqq T, j=1, 2$, the assertion of Theorem 6 is a corollary of Theorem 5. The proof is complete.

Acknowledgement. We are grateful to the referees for their constructive suggestions to improve the presentation of this paper.

REFERENCES

[1] KLEBANOV, L. B., The estimate of closeness of distributions in terms of the characteristic functions, *Teor. Verojatnost. i Primenen.* **22** (1977) (in Russian).

[2] KLEBANOV, L. B. and MELAMED, J. A., The asymptotic behaviour of certain estimates of location and scale parameters, *Problemy Peredači Informacii* **12** (1976), 41—56 (in Russian). *MR* **57** # 1745.

[3] KLEBANOV, L. B. and MELAMED, J. A., Asymptotic ε-admissibility of the sample variance as an estimator of the populátion variance, *Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov (LOMI)* **61** (1976), 75—83, 137 (in Russian). *MR* **57** # 4396.

[4] KAMKE, E., *Differentialgleichungen, Lösungsmethoden und Lösungen. I. Gewöhnliche Differentialgleichungen*, Neunte Auflage, B. G. Teubner, Stuttgart, 1977. *MR* **57** # 6549 (Russian translations: Izdat. Nauka, Moscow, 1971 and 1976. *MR* **54** # 3061 and *MR* **54** # 3062).

[5] KLEBANOV, L.B., Local behaviour of the solutions of ordinary differential equations, *Differencial'nye Uravnenija* **7** (1971), 1393—1397, 1539. *MR* **44** # 5560; ibid **8** (in Russian).

[6] BESOV, O. V., IL'IN, V. P. and NIKOL'SKI, S.M., *Integral representations of functions, and embedding theorems*, Izdat. Nauka, Moscow, 1975 (in Russian). *MR* **55** # 3776.

[7] KAGAN, A.M. and KLEBANOV, L.B., On the conditions of asymptotic ε-admissibility of the polynomial Pitman estimators of a location parameter and certain properties of information measures of closeness, Statistical Distributions in Sci. Work, D. Reidel, 1975.

[8] KLEBANOV, L.B. and MELAMED, J.A., Characterization of the normal and gamma distributions by properties of the Fisher information, *Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov (LOMI)* **43** (1974), 53—58, 170 (in Russian). *MR* **52** # 12157.

(Received September 22, 1981)

DEPARTMENT OF PROBABILITY THEORY AND
MATHEMATICAL STATISTICS AT THE
TBILISI MATHEMATICAL INSTITUTE
GEORGIAN ACADEMY OF SCIENCES
150A PLEKHANOV AVENUE
SU—380 012 TBILISI 12
SOVIET UNION

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 379—391

ON QUASI-UNIFORM DISTRIBUTION OF SEQUENCES OF INTEGERS

A. SÁRKÖZY

1. Throughout this paper, we use the following notation:

We write $e(\alpha)=e^{2\pi i\alpha}$. We put $\{x\}=x-[x]$ and $\|x\|=\min(\{x\}, 1-\{x\})$. $\mathcal{A}=\{a_1, a_2, \ldots\}$ (where $a_1<a_2<\ldots$) denotes a sequence of non-negative integers. We associate a sequence $\mathcal{D}=\{d_1, d_2, \ldots\}$ of non-negative real numbers with the sequence $\mathcal{A}$. For $u\geqq 0$, let

$$S_u(\alpha)=\sum_{a_i\leqq u} d_i e(a_i\alpha). \tag{1}$$

If $u\geqq 0$, and $q\geqq 1$ and h are arbitrary integers, we write

$$D(u, q, h)=\sum_{\substack{a_i\leqq u\\ a_i\equiv h \pmod q}} d_i \tag{2}$$

and

$$D(u)=D(u, 1, 0)=\sum_{a_i\leqq u} d_i.$$

The Hardy—Littlewood method is often used in order to show that a sequence $\mathcal{A}$ is an additive basis of finite order. In all these applications, we have to give upper bounds for the absolute values of sums of the form $S_u(b/q)$. (Note that we usually have $d_1=d_2=\ldots=1$; however, e.g. if $\mathcal{A}$ denotes the sequence of the prime numbers, i.e., $a_i=p_i$ then it is more convenient to put $d_i=\log p_i$. This is the reason of that that we associate a sequence $\mathcal{D}$ with the sequence $\mathcal{A}$.) In order to estimate $|S_u(b/q)|$, we have to study the numbers $D(u, q, h)$ (for $h=0, 1, \ldots, q-1$), roughly speaking, the *$\mathcal{D}$-distribution of the sequence $\mathcal{A}$ in the residue classes* (modulo q). (See [3]—[6].) The simplest case is when the sequence $\mathcal{A}$ is *$\mathcal{D}$-uniformly distributed in the residue classes* modulo q, i.e.,

$$\left|D(u, q, h)-\frac{D(u)}{q}\right|$$

is small (in terms of u and q) for all h. However, in the most applications, the sequence $\mathcal{A}$ is *not* uniformly distributed in this sense. However, the numbers $|S_u(b/q)|$ are small also in all these cases. The aim of this paper is to study the arithmetic background of this fact.

1980 *Mathematics Subject Classification.* Primary 10L99.

Key words and phrases. Sequences, uniform distribution, Hardy—Littlewood method, exponential sums, residue classes.

In fact, for $u \geqq 0$, $q=1, 2, \ldots, 0<\delta<1$, let

(3)
$$\gamma(u, q, \delta) \overset{\text{def}}{=} \frac{\max\limits_{\substack{0 \leqq h<q \\ (h,q)=1}} \sum\limits_{t=0}^{q-1} \left| \sum\limits_{t-\delta q<l \leqq t} D(u, q, lh) - \delta D(u) \right|}{\delta q}$$

and

(4)
$$\Gamma(u, q) \overset{\text{def}}{=} \inf_{0<\delta \leqq 1/2} \gamma(u, q, \delta).$$

If the numbers $\Gamma(u, q)$ are small (for large but fixed q and $u \to +\infty$) then we may say that the sequence $\mathcal{A}$ is *$\mathcal{D}$-quasi-uniformly distributed* (or $\mathcal{D}$-uniformly distributed on average) in the residue classes; if $\mathcal{D}=\{1, 1, \ldots\}$ then we may say briefly that $\mathcal{A}$ is *quasi-uniformly distributed* in the residue classes.

In Section 3, we show that the estimate of $\max\limits_{\substack{0 \leqq b<q \\ (b,q)=1}} |S_u(b/q)|$ is near equivalent to the estimate of $\Gamma(u, q)$. In particular, it can be shown by combining Theorem 4 with well-known results that, e.g., the sequences $\{1^k, 2^k, \ldots, n^k, \ldots\}$ and $\{2^k, 3^k, \ldots$ $\ldots, p_n^k, \ldots\}$ (k is a fixed positive integer) are quasi-uniformly distributed.

Finally, in Section 4 we prove a theorem which can be used for the estimation of $\Gamma(u, q)$. (This theorem can be applied, e.g. in order to study Goldbach's problem, in particular, to derive Vinogradov's fundamental lemma; see [4].)

2. In this section, we prove some preliminary lemmas.

LEMMA 1. *For any real number α we have*

$$|1-e(\alpha)| \leqq 2\pi |\alpha|.$$

PROOF.

$$|1-e(\alpha)| = |(1-\cos 2\pi\alpha) - i \sin 2\pi\alpha| =$$

$$= ((1-\cos 2\pi\alpha)^2 + \sin^2 2\pi\alpha)^{1/2} = (1-2\cos 2\pi\alpha + \cos^2 2\pi\alpha + \sin^2 2\pi\alpha)^{1/2} =$$

$$= (2-2\cos 2\pi\alpha)^{1/2} = (2(1-\cos 2\pi\alpha))^{1/2} = (2 \cdot 2 \sin^2 \pi\alpha)^{1/2} =$$

$$= 2 |\sin \pi\alpha| \leqq 2\pi |\alpha|.$$

LEMMA 2. *For any real number α we have*

$$|1-e(\alpha)| \geqq 4 \|\alpha\|.$$

This lemma is identical with Lemma 4 in [5].

LEMMA 3. *If t is a positive integer and α is any real number then we have*

$$\left| \sum_{j=0}^{t-1} e(j\alpha) \right| \leqq \min\left(t, \frac{1}{2\|\alpha\|}\right)$$

$\left(\textit{where } \min\left(t, \frac{1}{0}\right) \overset{\text{def}}{=} t\right)$.

PROOF.

$$\left|\sum_{j=0}^{t-1} e(j\alpha)\right| \leqq t$$

is trivial while

$$\left|\sum_{j=0}^{t-1} e(j\alpha)\right| \leqq \frac{1}{2\|\alpha\|}$$

holds by Lemma 5 in [5].

LEMMA 4. *Let q be a positive integer, $z_0, z_1, \ldots, z_{q-1}$ any complex numbers and put*

$$A_b = \sum_{j=0}^{q-1} z_j e(jb/q) \qquad (\textit{for } b = 0, \pm 1, \pm 2, \ldots). \tag{5}$$

Then we have

$$\sum_{b=0}^{q-1} |A_b|^2 = q \sum_{j=0}^{q-1} |z_j|^2.$$

PROOF.

$$\sum_{b=0}^{q-1} |A_b|^2 = \sum_{b=0}^{q-1} A_b \bar{A}_b = \sum_{b=0}^{q-1} \left(\sum_{j=0}^{q-1} z_j e(jb/q)\right)\left(\sum_{k=0}^{q-1} \bar{z}_k e(-kb/q)\right) =$$

$$= \sum_{j=0}^{q-1} \sum_{k=0}^{q-1} z_j \bar{z}_k \left(\sum_{b=0}^{q-1} e((j-k)b/q) = \sum_{q=0}^{q-1} z_j \bar{z}_j q = q \sum_{j=0}^{q-1} |z_j|^2.$$

LEMMA 5. *Let q be a positive integer satisfying*

$$q > 1, \tag{6}$$

$\ldots, z_{-2}, z_{-1}, z_0, z_1, z_2, \ldots$ any complex numbers such that $z_{n+q} = z_n$ for all n, δ a real number such that

$$0 < \delta \leqq 1/2. \tag{7}$$

Put

$$\sum_{j=0}^{q-1} z_j = Z \tag{8}$$

and

$$\sum_{\substack{0 \leqq h < q \\ (h,q)=1}} \sum_{t=0}^{q-1} \left| \sum_{t-\delta q < l \leqq t} z_{lh} - \delta Z \right| = K. \tag{9}$$

If b is any integer satisfying $(b, q) = 1$ and A_b is defined by (5) then we have

$$|A_b| \leqq 3K/\delta q.$$

PROOF. This lemma is a slightly modified form of Theorem 7 in [4], and it can be proved similarly but some modifications are needed also in the proof.

Let m, n be integers such that $q \nmid m$ and $(n, q) = 1$. For $h = 0, 1, \ldots, q-1$, define the integer h' by

$$h' \equiv hn \pmod{q}, \qquad 0 \leqq h' \leqq q-1. \tag{10}$$

By $(n, q)=1$, there exists a uniquely determined integer n^* for which

$$nn^* \equiv 1 \pmod q, \qquad 0 \leqq n^* \leqq q-1.$$

Then, obviously, $(n^*, q)=1$. (10) implies that

$$h \equiv h'n^* \pmod q.$$

Furthermore, if h runs over the numbers $0, 1, \dots, q-1$, then again by $(n, q)=1$, also h' runs over these numbers. Thus we have

$$A_{mn} = \sum_{h=0}^{q-1} z_h e(hmn/q) = \sum_{h'=0}^{q-1} z_{h'n^*} e(h'm/q),$$

hence (with respect to (6) and $q \nmid m$)

(11)
$$\begin{aligned}
&\Big|\sum_{0\leqq j<\delta q} e(mj/q)\Big| |A_{mn}| = \\
&= \Big|\sum_{0\leqq j<\delta q} e(mj/q) \sum_{h=0}^{q-1} z_h e(hmn/q)\Big| = \\
&= \Big|\sum_{0\leqq j<\delta q} e(mj/q) \sum_{h'=0}^{q-1} z_{h'n^*} e(h'm/q)\Big| = \\
&= \Big|\sum_{t=0}^{q-1} \Big(\sum_{t-\delta q<l\leqq t} z_{ln^*}\Big) e(tm/q)\Big| = \\
&= \Big|\sum_{t=0}^{q-1} \delta Z e(tm/q) + \sum_{t=0}^{q-1}\Big(\sum_{t-\delta q<l\leqq t} z_{ln^*} - \delta Z\Big) e(tm/q)\Big| = \\
&= \Big|\sum_{t=0}^{q-1}\Big(\sum_{t-\delta q<l\leqq t} z_{ln^*} - \delta Z\Big) e(tm/q)\Big| \quad (\text{where } q \nmid m \text{ and } (n, q) = 1).
\end{aligned}$$

By using this identity with $m=1$ and $n=b$, we obtain with respect to (9) that

(12)
$$\begin{aligned}
&\Big|\sum_{0\leqq j<\delta q} e(j/q)\Big| |A_b| = \\
&= \Big|\sum_{t=0}^{q-1}\Big(\sum_{t-\delta q<l\leqq t} z_{lb^*} - \delta Z\Big) e(t/q)\Big| \leqq \sum_{t=0}^{q-1} \Big|\sum_{t-\delta q<l\leqq t} z_{lb^*} - \delta Z\Big| \leqq K.
\end{aligned}$$

Define the positive integer j_0 by

(13)
$$j_0 - 1 < \delta q \leqq j_0.$$

Then by Lemmas 1 and 2, and with respect to (6), we have

(14)
$$\begin{aligned}
&\Big|\sum_{0\leqq j<\delta q} e(j/q)\Big| = \Big|\sum_{j=0}^{j_0-1} e(j/q)\Big| = \frac{|1-e(j_0/q)|}{|1-e(1/q)|} \geqq \\
&\geqq \frac{4\|j_0/q\|}{2\pi/q} = \frac{2}{\pi}\|j_0/q\| q.
\end{aligned}$$

If $j_0 \leqq q/2$ then by (13) we have

(15) $$\|j_0/q\| = j_0/q \geqq \delta q/q = \delta \qquad (\text{for } j_0 \leqq q/2).$$

Assume now that $j_0 > q/2$. Then (7) and (13) yield that q must be odd and $j_0 = (q+1)/2$, hence in view of (6) and (7)

(16) $$\|j_0/q\| = 1 - j_0/q = 1-(q+1)/2q = \frac{1}{2} - \frac{1}{2q} \geqq \frac{1}{2} - \frac{1}{6} = \\ = \frac{1}{3} = \frac{2}{3}\cdot\frac{1}{2} \geqq \frac{2}{3} \qquad (\text{for } j_0+1 > q/2).$$

From (14), (15) and (16) we obtain that

(17) $$\Big|\sum_{0 \leqq j < \delta q} e(j/q)\Big| \geqq \frac{2}{\pi}\cdot\frac{2}{3}\delta q = \frac{4}{3\pi}\delta q > \delta q/3.$$

(12) and (17) yield that

$$|A_b| \leqq \frac{K}{\Big|\sum\limits_{0 \leqq j < \delta q} e(j/q)\Big|} \leqq \frac{K}{\delta q/3} = 3K/\delta q$$

which completes the proof of the lemma.

Lemma 6. *Let us define $q, \ldots, z_{-2}, z_{-1}, z_0, z_1, z_2, \ldots, \delta$, Z and K in the same way as in Lemma 5, and define A_b by* (5), *j_0 by* (13). *For $d|q$, put*

$$\max_{\substack{0 \leqq r < q \\ (r,q)=d}} |A_r| = M_d.$$

Then we have

(18) $$K \leqq |Z| + \Big(3qj_0 \sum_{\substack{d|q \\ d \leqq q/2j_0}} \frac{M_d^2}{d} + q^2 \sum_{\substack{d|q \\ q/2j_0 < d \leqq q/2}} \frac{M_d^2}{d^2}\Big)^{1/2}.$$

Corollary 1. *Define $q, \ldots, z_{-2}, z_{-1}, z_0, z_1, z_2, \ldots, Z, K$ (with $1/2$ in place of δ) and A_b in the same way as in Lemma 6. Then we have*

(19) $$K \leqq |Z| + \sqrt{3}\, q \Big(\sum_{\substack{d|q \\ d \leqq q/2}} \frac{M_d^2}{d^2}\Big)^{1/2}.$$

Proof of Lemma 6. By using Lemma 3 and the identity (11) with h^* in place of n (so that $n^*=(h^*)^*=h$), we obtain for $q\nmid m$ and $(h,q)=1$ that

$$\left|\sum_{t=0}^{q-1}\Big(\sum_{t-\delta q<l\leq t} z_{lh}-\delta Z\Big)e(tm/q)\right| =$$

(20)
$$= |A_{mh^*}|\left|\sum_{0\leq j<\delta q} e(mj/q)\right| \leq |A_{mh^*}|\min\left(\sum_{0\leq j<\delta q} 1,\ \frac{1}{2\left\|\frac{m}{q}\right\|}\right) =$$

$$= |A_{mh^*}|\min\left(j_0,\ \frac{1}{2\left\|\frac{m}{q}\right\|}\right) \qquad (\text{for } q\nmid m \text{ and } (h,q)=1)$$

where j_0 is defined by (13). By $(h,q)=1$, we have $(h^*,q)=1$ so that $(mh^*,q)= =(m,q)$. Thus we obtain from (20) that

$$\left|\sum_{t=0}^{q-1}\Big(\sum_{t-\delta q<l\leq t} z_{lh}-\delta Z\Big)e(tm/q)\right| \leq$$

(21)
$$\leq \Big(\max_{\substack{0\leq r<q\\ (r,q)=(m,q)}} |A_r|\Big)\min\left(j_0,\ \frac{1}{2\left\|\frac{m}{q}\right\|}\right) =$$

$$= M_{(m,q)}\min\left(j_0,\ \frac{1}{2\left\|\frac{m}{q}\right\|}\right) \qquad (\text{for } q\nmid m \text{ and } (h,q)=1).$$

Furthermore, by $(h,q)=1$ we have

$$\left|\sum_{t=0}^{q-1}\Big(\sum_{t-\delta q<l\leq t} z_{lh}-\delta Z\Big)\right| = \left|\sum_{t=0}^{q-1}\sum_{0\leq j<\delta q} z_{(t-j)h}-\delta qZ\right| =$$

(22)
$$= \left|\sum_{0\leq j<\delta q}\Big(\sum_{t=0}^{q-1} z_{(t-j)h}\Big)-\delta qZ\right| = \left|\sum_{0\leq j<\delta q}\sum_{r=0}^{q-1} z_r-\delta qZ\right| =$$

$$= \left|\sum_{0\leq j<\delta q} Z-\delta qZ\right| = |Z|\left|\sum_{0\leq j<\delta q} 1-\delta q\right| \leq |Z|.$$

By using Lemma 4 with

$$\sum_{j-\delta q<l\leq j} z_{lh}-\delta Z$$

in place of z_j, we obtain with respect to (21) and (22) that

$$q\sum_{t=0}^{q-1}\Big|\sum_{t-\delta q<l\leq t} z_{lh}-\delta Z\Big|^2 = \sum_{m=0}^{q-1}\left|\sum_{t=0}^{q-1}\Big(\sum_{t-\delta q<l\leq t} z_{lh}-\delta Z\Big)e(tm/q)\right|^2 =$$

$$= \left|\sum_{t=0}^{q-1}\Big(\sum_{t-\delta q<l\leq t} z_{lh}-\delta Z\Big)\right|^2 + \sum_{m=1}^{q-1}\left|\sum_{t=0}^{q-1}\Big(\sum_{t-\delta q<l\leq t} z_{lh}-\delta Z\Big)e(tm/q)\right|^2 \leq$$

$$\leqq |Z|^2+\sum_{m=1}^{q-1}\left(M_{(m,q)}\min\left(j_0,\frac{1}{2\left\|\frac{m}{q}\right\|}\right)\right)^2 \leqq$$

$$\leqq |Z|^2+2\sum_{1\leqq m\leqq q/2}\left(M_{(m,q)}\min\left(j_0,\frac{1}{2\left\|\frac{m}{q}\right\|}\right)\right)^2 =$$

$$= |Z|^2+2\sum_{1\leqq m\leqq q/2}\left(M_{(m,q)}\min\left(j_0,\frac{q}{2m}\right)\right)^2 =$$

$$= |Z|^2+2\sum_{\substack{d|q\\ d\leqq q/2}}\sum_{\substack{1\leqq k\leqq q/2d\\ (k,q/d)=1}}\left(M_d\min\left(j_0,\frac{q}{2dk}\right)\right)^2 \leqq$$

$$\leqq |Z|^2+2\sum_{\substack{d|q\\ d\leqq q/2}}M_d^2\sum_{1\leqq k\leqq q/2d}\left(\min\left(j_0,\frac{q}{2dk}\right)\right)^2 \leqq$$

$$\leqq |Z|^2+2\sum_{\substack{d|q\\ d\leqq q/2}}M_d^2\left(\sum_{1\leqq k\leqq q/2dj_0}j_0^2+\sum_{q/2dj_0<k\leqq q/2d}\frac{q^2}{4d^2k^2}\right) = \tag{23}$$

$$= |Z|^2+2\sum_{\substack{d|q\\ d\leqq q/2j_0}}M_d^2\left(j_0^2\sum_{1\leqq k\leqq q/2dj_0}1+\sum_{q/2dj_0<k\leqq q/2d}\frac{q^2}{4d^2k^2}\right)+$$

$$+2\sum_{\substack{d|q\\ q/2j_0<d\leqq q/2}}M_d^2\sum_{1\leqq k\leqq q/2d}\frac{q^2}{4d^2k^2} \leqq$$

$$\leqq |Z|^2+2\sum_{\substack{d|q\\ d\leqq q/2j_0}}M_d^2\left(\frac{qj_0}{2d}+\frac{q^2}{4d^2}\sum_{k=[q/2dj_0]+1}\frac{1}{(k-1)k}\right)+$$

$$+2\sum_{\substack{d|q\\ q/2j_0<d\leqq q/2}}M_d^2\frac{q^2}{4d^2}\sum_{k=1}^{+\infty}\frac{1}{k^2} \leqq$$

$$\leqq |Z|^2+\sum_{\substack{d|q\\ d\leqq q/2j_0}}M_d^2\left(\frac{qj_0}{d}+\frac{q^2}{2d^2}\frac{1}{[q/2dj_0]}\right)+\sum_{\substack{d|q\\ q/2j_0<d\leqq q/2}}M_d^2\frac{q^2}{d^2} \leqq$$

$$\leqq |Z|^2+\sum_{\substack{d|q\\ d\leqq q/2j_0}}M_d^2\left(\frac{qj_0}{d}+\frac{q^2}{d^2}\frac{2dj_0}{q}\right)+q^2\sum_{\substack{d|q\\ q/2j_0<d\leqq q/2}}\frac{M_d^2}{d^2} =$$

$$= |Z|^2+3qj_0\sum_{\substack{d|q\\ d\leqq q/2j_0}}\frac{M_d^2}{d}+q^2\sum_{\substack{d|q\\ q/2j_0<d\leqq q/2}}\frac{M_d^2}{d^2}.$$

Finally, by the Cauchy—Schwarz inequality we have

$$\left(\sum_{t=0}^{q-1}\Big|\sum_{t-\delta q<l\leqq t} z_{lh}-\delta Z\Big|\right)^2 \leqq q\sum_{t=0}^{q-1}\Big|\sum_{t-\delta q<l\leqq t} z_{lh}-\delta Z\Big|^2. \tag{24}$$

With respect to the inequality

$$|a|^2+|b|^2 \leqq (|a|+|b|)^2,$$

(23) and (24) yield that

$$\sum_{t=0}^{q-1}\Big|\sum_{t-\delta q<l\leqq t} z_{lh}-\delta Z\Big| \leqq \Bigg(|Z|^2+3qj_0\sum_{\substack{d|q\\ d\leqq q/2j_0}}\frac{M_d^2}{d}+q^2\sum_{\substack{d|q\\ q/2j_0<d\leqq q/2}}\frac{M_d^2}{d^2}\Bigg)^{1/2} \leqq$$

$$\leqq |Z|+\Bigg(3qj_0\sum_{\substack{d|q\\ d\leqq q/2j_0}}\frac{M_d^2}{d}+q^2\sum_{\substack{d|q\\ q/2j_0<d\leqq q/2}}\frac{M_d^2}{d^2}\Bigg)^{1/2}.$$

This holds for all $(h, q)=1$ which proves (18) and the proof of Lemma 6 is completed.

PROOF OF COROLLARY 1. We use Lemma 6 with $\delta=1/2$ so that

$$j_0-1<\frac{q}{2}\leqq j_0$$

hence

$$j_0<\frac{q}{2}+1\leqq q$$

and

$$\frac{q}{2j_0}\leqq 1.$$

Then we obtain that

$$K\leqq |Z|+\Bigg(3qj_0\sum_{\substack{d|q\\ d\leqq q/2j_0}}\frac{M_d^2}{d}+q^2\sum_{\substack{d|q\\ q/2j_0<d\leqq q/2}}\frac{M_d^2}{d^2}\Bigg)^{1/2} \leqq$$

$$\leqq |Z|+\Bigg(3q^2\sum_{\substack{d|q\\ d\leqq q/2j_0}}\frac{M_d^2}{d^2}+q^2\sum_{\substack{d|q\\ q/2j_0<d\leqq q/2}}\frac{M_d^2}{d^2}\Bigg)^{1/2} \leqq$$

$$\leqq |Z|+\Bigg(3q^2\sum_{\substack{d|q\\ d\leqq q/2}}\frac{M_d^2}{d_2}\Bigg)^{1/2} = |Z|+\sqrt{3}\,q\Bigg(\sum_{\substack{d|q\\ d\leqq q/2}}\frac{M_d^2}{d^2}\Bigg)^{1/2}$$

which completes the proof of Corollary 1.

3. In this section, we use Lemmas 5 and 6 and Corollary 1 in order to estimate $\gamma(u, q, \delta)$ and $\Gamma(u, q)$.

THEOREM 1. *If* $u\geqq 0$, $0<\delta\leqq 1/2$ *and* q, b *are integers such that* $q\geqq 2$ *and* $(b, q)=1$, *then we have*

$$|S_u(b/q)| \leqq 3\gamma(u, q, \delta). \tag{25}$$

THEOREM 2. *If* $u \geqq 0$, $0 < \delta \leqq 1/2$ *and* q *is an integer such that* $q \geqq 2$, *then we have*

$$\gamma(u, q, \delta) \leqq \frac{1}{\delta q}\Big\{D(u) + \Big(3j_0 \sum_{\substack{d|q \\ 2j_0 \leqq d}} d \max_{\substack{0 \leqq b < d \\ (b,d)=1}} |S_u(b/d)|^2 + \sum_{\substack{d|q \\ 2 \leqq d < 2j_0}} d^2 \max_{\substack{0 \leqq b < d \\ (b,d)=1}} |S_u(b/d)|^2\Big)^{1/2}\Big\}$$

where j_0 *is defined by* (13).

THEOREM 3. *If* $u \geqq 0$ *and* q, b *are integers such that* $q \geqq 2$ *and* $(b, q) = 1$, *then we have*

(26) $$|S_u(b/q)| \leqq 3\Gamma(u, q).$$

THEOREM 4. *If* $u \geqq 0$ *and* q *is an integer such that* $q \geqq 2$, *then we have*

(27) $$\Gamma(u, q) \leqq \frac{2}{q} D(u) + \frac{4}{q}\Big(\sum_{\substack{d|q \\ 2 \leqq d}} d^2 \max_{\substack{0 \leqq b < d \\ (b,d)=1}} |S_u(b/d)|^2\Big)^{1/2}.$$

COROLLARY 2. *If* $u \geqq 0$ *and* p *is a prime number, then we have*

$$\Gamma(u, p) \leqq \frac{2}{p} D(u) + 4 \max_{0 < b < p} |S_u(b/p)|.$$

(Note that Corollary 2 shows that Theorem 3 is nearly best possible.)

PROOF OF THEOREM 1. We have

(28) $$|S_u(b/q)| = \Big|\sum_{a_i \leqq u} d_i e(a_i b/q)\Big| = \\ = \Big|\sum_{h=0}^{q-1} \Big(\sum_{\substack{a_i \leqq u \\ a_i \equiv h \pmod q}} d_i\Big) e(hb/q)\Big| = \Big|\sum_{h=0}^{q-1} D(u, q, h) e(hb/q)\Big|.$$

Thus Lemma 5 yields (with $D(u, q, j)$ in place of z_j) that

$$|S_u(b/q)| = \Big|\sum_{h=0}^{q-1} D(u, q, h) e(hb/q)\Big| \leqq$$

$$\leqq \frac{3}{\delta q} \max_{\substack{0 \leqq h < q \\ (h,q)=1}} \sum_{t=0}^{q-1} \Big|\sum_{t - \delta q < l \leqq t} D(u, q, lh) - \delta \sum_{j=0}^{q-1} D(u, q, j)\Big| =$$

$$= \frac{3}{\delta q} \max_{\substack{0 \leqq h < q \\ (h,q)=1}} \sum_{t=0}^{q-1} \Big|\sum_{t - \delta q < l < t} D(u, q, lh) - \delta D(u)\Big| = 3\gamma(u, q, \delta)$$

which completes the proof of Theorem 1.

Proof of Theorem 2. In view of (28), Lemma 6 yields (with $D(u, q, j)$ in place of z_j) that

$$\gamma(u, q, \delta) = \frac{1}{\delta q} \max_{\substack{0 \le h < q \\ (h, q)=1}} \sum_{t=0}^{q-1} \Big| \sum_{t-\delta q < l \le t} D(u, q, lh) - \delta D(u) \Big| =$$

$$= \frac{1}{\delta q} \max_{\substack{0 \le h < q \\ (h, q)=1}} \sum_{t=0}^{q-1} \Big| \sum_{t-\delta q < l \le t} D(u, q, lh) - \delta \sum_{j=0}^{q-1} D(u, q, j) \Big| \le$$

$$\le \frac{1}{\delta q} \Big\{ \sum_{j=0}^{q-1} D(u, q, j) + \Big(3qj_0 \sum_{\substack{d' | q \\ d' \le q/2j_0}} \frac{1}{d'} \max_{\substack{0 \le r < q \\ (r, q)=d'}} \Big| \sum_{h=0}^{q-1} D(u, q, h) e(hr/q) \Big|^2 +$$

$$+ q^2 \sum_{\substack{d' | q \\ q/2j_0 < d' \le q/2}} \frac{1}{(d')^2} \max_{\substack{0 \le r < q \\ (r, q)=d'}} \Big| \sum_{h=0}^{q-1} D(u, q, h) e(hr/q) \Big|^2 \Big)^{1/2} \Big\} =$$

$$= \frac{1}{\delta q} \Big\{ D(u) + \Big(3qj_0 \sum_{\substack{d' | q \\ d' \le q/2j_0}} \frac{1}{d'} \max_{\substack{0 \le b < q/d' \\ (b, q/d')=1}} |S_u(d'b/q)|^2 +$$

$$+ q^2 \sum_{\substack{d' | q \\ q/2j_0 < d' \le q/2}} \frac{1}{(d')^2} \max_{\substack{0 \le b < q/d' \\ (b, q/d')=1}} |S_u(d'b/q)|^2 \Big)^{1/2} \Big\}.$$

Writing $d = q/d'$, we obtain that

$$\gamma(u, q, \delta) \le$$

$$\le \frac{1}{\delta q} \Big\{ D(u) + \Big(3qj_0 \sum_{\substack{d | q \\ 2j_0 \le d}} \frac{d}{q} \max_{\substack{0 \le b < d \\ (b, d)=1}} |S_u(b/d)|^2 + q^2 \sum_{\substack{d | q \\ 2 \le d < 2j_0}} \frac{d^2}{q^2} \max_{\substack{0 \le b < d \\ (b, d)=1}} |S_u(b/d)|^2 \Big)^{1/2} \Big\} =$$

$$= \frac{1}{\delta q} \{ D(u) + (3j_0 \sum_{\substack{d | q \\ 2j_0 \le d}} d \max_{\substack{0 \le b < d \\ (b, d)=1}} |S_u(b/d)|^2 + \sum_{\substack{d | q \\ 2 \le d < 2j_0}} d^2 \max_{\substack{0 \le b < d \\ (b, d)=1}} |S_u(b/d)|^2)^{1/2} \}$$

which completes the proof of Theorem 2.

Proof of Theorem 3. By Theorem 1, (25) holds for all $0 < \delta \le 1/2$ which implies (26).

Proof of Theorem 4. We have

$$\Gamma(u, q) = \inf_{0 < \delta \le 1/2} \gamma(u, q, \delta) \le \gamma(u, q, 1/2). \tag{29}$$

With respect to (28), Corollary 1 yields that

$$\gamma\left(u, q, \frac{1}{2}\right) = \frac{2}{q} \max_{\substack{0 \leqq h < q \\ (h,q)=1}} \sum_{t=0}^{q-1} \left| \sum_{t-q/2<l\leqq t} D(u, q, lh) - \frac{1}{2} D(u) \right| =$$

$$= \frac{2}{q} \max_{\substack{0 \leqq h < q \\ (h,q)=1}} \sum_{t=0}^{q-1} \left| \sum_{t-q/2<l\leqq t} D(u, q, lh) - \frac{1}{2} \sum_{j=0}^{q-1} D(u, q, j) \right| \leqq$$

$$\leqq \frac{2}{q} \left\{ \sum_{j=0}^{q-1} D(u, q, j) + \sqrt{3}\, q \left(\sum_{\substack{d'|q \\ d' \leqq q/2}} \frac{1}{(d')^2} \max_{\substack{0 \leqq r < q \\ (r,q)=d'}} \left| \sum_{h=0}^{q-1} D(u, q, h) e(hr/q) \right|^2 \right)^{1/2} \right\} =$$

$$= \frac{2}{q} D(u) + 2\sqrt{3} \left(\sum_{\substack{d'|q \\ d' \leqq q/2}} \frac{1}{(d')^2} \max_{\substack{0 \leqq b < q/d' \\ (b, q/d')=1}} |S_u(d'b/q)|^2 \right)^{1/2}.$$

Writing $d = q/d'$, we obtain that

$$\gamma(u, q, 1/2) \leqq \frac{2}{q} D(u) + 4 \left(\sum_{\substack{d|q \\ 2 \leqq d}} \left(\frac{d}{q}\right)^2 \max_{\substack{0 \leqq b < d \\ (b,d)=1}} |S_u(b/d)|^2 \right)^{1/2} =$$

(30)

$$= \frac{2}{q} D(u) + \frac{4}{q} \Big(\sum_{\substack{d|q \\ 2 \leqq d}} d^2 \max_{\substack{0 \leqq b < d \\ (b,d)=1}} |S_u(b/d)|^2 \Big)^{1/2}.$$

(29) and (30) yield (27).

PROOF OF COROLLARY 2. Corollary 2 is a consequence of Theorem 4. In fact, if $q=p$ is a prime number in (27) then we obtain that

$$\Gamma(u, p) \leqq \frac{2}{p} D(u) + \frac{4}{p} \Big(\sum_{\substack{d|q \\ 2 \leqq d}} d^2 \max_{\substack{0 \leqq b < d \\ (b,d)=1}} |S_u(b/d)|^2 \Big)^{1/2} =$$

$$= \frac{2}{p} D(u) + \frac{4}{p} \Big(p^2 \max_{\substack{0 \leqq b < p \\ (b,p)=1}} |S_u(b/p)|^2 \Big)^{1/2} = \frac{2}{p} D(u) + 4 \max_{0<b<p} |S_u(b/p)|.$$

4. In this section, we estimate $\Gamma(u, q)$.

LEMMA 7. *Let q, h, t be integers satisfying $q>2$ and $(q, h)=1$, ..., $z_{-2}, z_{-1}, z_0, z_1, z_2, \ldots$ any complex numbers such that $z_{n+q}=z_n$ for all n and*

(31) $$z_m = 0 \text{ for all } m \text{ satisfying } (q, m) > 1.$$

Define Z by (8). *Then for all $\delta>0$ we have*

(32) $$\Big| \sum_{t-\delta q<l\leqq t} z_{lh} - \delta Z \Big| \leqq c_1 q^{-1/2} \log\log q \left(\log q \sum_{\chi \neq \chi_0} \left| \sum_{j=0}^{q-1} \chi(j) z_j \right| + |Z| \right).$$

(Here and in what follows, in $\sum_{\chi\neq\chi_0}$, χ runs over all the characters χ modulo q, different from the principal character χ_0.)

Note that Lemma 7 can be used also in the more general case when (31) does not hold; in fact, in this case we have

$$\sum_{t-\delta q<l\leq t} z_{lh}-\delta Z = \sum_{d|q}\Big(\sum_{\substack{t/d-\delta q/d<l\leq t/d\\(l,q/d)=1}} z_{ldh}-\delta\sum_{\substack{1\leq j\leq q/d\\(j,q/d)=1}} z_{jd}\Big) \tag{33}$$

and here each term can be estimated by using Lemma 7.

Proof of Lemma 7. Lemma 7 is identical with Theorem 8 in [4]; in fact, (32) can be derived by using the Pólya—Vinogradov inequality (see [2] and [7]).

Theorem 5. *If $u\geq 0$ and q is an integer such that $q>2$ then we have*

$$\Gamma(u,q)\leq c_2\Big\{q^{-1/2}\log\log q\Big(\log q\sum_{\chi\neq\chi_0}\Big|\sum_{j=0}^{q-1}\chi(j)D(u,q,j)\Big|+D(u)\Big)+\sum_{\substack{a_i\leq u\\(a_i,q)>1}} d_i\Big\}.$$

Note that Theorem 5 covers the case when $(a_i, q)=1$ holds apart from a few exceptional a_i (like in the case of Goldbach's problem; see [4] and [1], Chapter 16). However, also the general case can be treated by combining Lemma 7 with (33).

Proof. By using Lemma 7 with

$$z_n=\begin{cases}D(u,q,n) & \text{for } (n,q)=1\\ 0 & \text{for } (n,q)>1\end{cases}$$

we obtain for $t=0,\pm1,\pm2,\dots,(h,q)=1$ and $0<\delta\leq 1/2$ that

$$\Big|\sum_{t-\delta q<l\leq t} D(u,q,lh)-\delta D(u)\Big| =$$

$$=\Big|\Big(\sum_{\substack{t-\delta q<l\leq t\\(l,q)=1}} D(u,q,lh)-\delta\sum_{\substack{a_i\leq u\\(a_i,q)=1}} d_i\Big)+$$

$$+\sum_{\substack{t-\delta q<l\leq t\\(l,q)<1}}\ \sum_{\substack{a_i\leq u\\a_i\equiv lh\,(\mathrm{mod}\,q)}} d_i-\delta\sum_{\substack{a_i\leq u\\(a_i,q)>1}} d_i\Big|\leq$$

$$\leq\Big|\sum_{\substack{t-\delta q<l\leq t\\(l,q)=1}} D(u,q,lh)-\delta\sum_{\substack{a_i\leq u\\(a_i,q)=1}} d_i\Big|+\sum_{\substack{a_i\leq u\\(a_i,q)>1}} d_i+\delta\sum_{\substack{a_i\leq u\\(a_i,q)<1}} d_i\leq$$

$$\leq c_1q^{-1/2}\log\log q\Big(\log q\sum_{\chi\neq\chi_0}\Big|\sum_{j=0}^{q-1}\chi(j)D(u,q,j)\Big|+\sum_{\substack{a_i\leq u\\(a_i,q)=1}} d_i\Big)+(1+\delta)\sum_{\substack{a_i\leq u\\(a_i,q)<1}} d_i\leq$$

$$\leq c_3\Big\{q^{-1/2}\log\log q\Big(\log q\sum_{\chi\neq\chi_0}\Big|\sum_{j=0}^{q-1}\chi(j)D(u,q,j)\Big|+D(u)\Big)+\sum_{\substack{a_i\leq u\\(a_i,q)>1}} d_i\Big\}$$

(where $c_3 = \max(c_1, 3/2)$). Thus we have

$$\Gamma(u, q) \leqq \gamma(u, q, 1/2) = \frac{2}{q} \max_{\substack{0 \leqq h < q \\ (h, q)=1}} \sum_{t=0}^{q-1} \Big| \sum_{t-\delta q < l \leqq t} D(u, q, lh) - \delta D(u) \Big| \leqq$$

$$\leqq \frac{2}{q} \max_{\substack{0 \leqq h < q \\ (h, q)=1}} \sum_{t=0}^{q-1} c_3 \left\{ q^{-1/2} \log\log q \left(\log q \sum_{\chi \neq \chi_0} \Big| \sum_{j=0}^{q-1} \chi(j) D(u, q, j) \Big| + D(u) \right) + \sum_{\substack{a_i \leqq u \\ (a_i, q) > 1}} d_i \right\} =$$

$$= c_4 \left\{ q^{-1/2} \log\log q \left(\log q \sum_{\chi \neq \chi_0} \Big| \sum_{j=0}^{q-1} \chi(j) D(u, q, j) \Big| + D(u) \right) + \sum_{\substack{a_i \leqq u \\ (a_i, q) > 1}} d_i \right\}$$

(where $c_4 = 2c_3$) which completes the proof of Theorem 5.

REFERENCES

[1] Montgomery, H. L., *Topics in multiplicative number theory*, Lecture Notes in Mathematics, Vol. 227, Springer-Verlag, Berlin—New York, 1971. *MR* 49 # 2616.

[2] Pólya, G., Über die Verteilung der quadratischen Reste und Nichtreste, *Göttinger Nachrichten* 1918, 21—29.

[3] Sárközy, A., On additive representations of integers, I, *Studia Sci. Math. Hungar.* **14** (1979), 145—167.

[4] Sárközy, A., On additive representations of integers, II, *Acta Math. Acad. Sci. Hungar.* **38** (1981), 157—181.

[5] Sárközy, A., On additive representations of integers, III, *Period. Math. Hungar.* **14** (1983), 7—30.

[6] Sárközy, A., On additive representations of integers, IV, *Topics in classical number theory*, ed. by G. Halász, Colloq. Math. Soc. J. Bolyai **34**, North-Holland Publ. Co., Amsterdam—New York.

[7] Vinogradov, I. M., Sur la distribution des résidus et des nonrésidus des puissances, *Perm. Ž. Fiz.-Mat. Obsč.* **1** (1918), 18—28 and 94—98.

(Received September 25, 1981)

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

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 393—401

TESTING FOR NORMALITY OF ERRORS IN LINEAR MODELS

WIESŁAW WAGNER

1. Introduction

1.1. *Literature review.* Many authors dealt with the problem of estimation of random errors in linear models. Among them it is worthwhile to mention: Theil [14—16], Koerts and Abrahamse [8], Golub and others [6], Benhelli [1] and Caussinus [3]. Theil introduced residuals of BLUS type for error estimation in regression analysis. Farebrother [5] published the program in ALGOL, in which he took advantage of Theil's ideas. Golub and others analized pseudo-residuals of outliers for detecting spurosity in linear model. Benhelli gave descriptions of many procedures useful for calculation of following residuals: BLUE, BLUS, BLUF, BAUS and NBAUS. Finally, Caussinus introduced the applications of the above named residuals for normality testing purposes.

1.2. *Some assumptions, definitions and notation.* By linear model, cf., e.g. Graybill ([7] p. 97), we mean an equation, in which the observable dependent variable is on the left side as for as the right side contains the observable known variables, the unknown parameters and the not observable random variables; the relation between the left and right sides being linear. In addition as an integral part of the model some information on the parameters (e.g. restriction) and on the random variables (e.g. assumption on the distribution) may be given. In the linear model, to be applied for analysis of experimental data, the observable variable y_i is assumed to be a linear function:

$$y_i = \sum_{j=1}^{q} x_{ij}\beta_j + e_i, \qquad i = 1, \dots, n,$$

where β_j, $j=1, \dots, q$ are unknown parameters, x_{ij}, $i=1, \dots, n$, $j=1, \dots, q$ are known constants, e_i, $i=1, \dots, n$ are not observable random variables with some distribution. The random variables e_i are called random errors and we assume usually that they are uncorrelated with expectation null and with identical, but unknown, variance σ^2.

The described model has the following matrix representation

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}. \tag{1.1}$$

This linear model is determined by the triplet $(\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$, where $\mathbf{X}$ is a known $n\times q$ matrix of any rank and $\mathbf{I}$ is the $n\times n$ unit matrix, $E(\mathbf{y})=\mathbf{X}\boldsymbol{\beta}$ and $D(\mathbf{y})=\sigma^2\mathbf{I}$. We use the symbol $\mathcal{M}_{m,n}$ for the linear space of $m\times n$ matrices of real elements.

1980 *Mathematics Subject Classification.* Primary 62G10.
Key words and phrases. Testing for departure from normality, residual vector with linear model.

In the statistical analysis of model (1.1) we are interested in the point estimation of the parameters β_j and σ^2. In some cases we test various parametric hypotheses under the assumption that the random error vector $\mathbf{e}$ has a multidimensional normal distribution. Such assumption is useful for the analysis of model (1.1), however, it is not always supported by the actual experimental data. It is necessary to decide whether the assumption on the normality of the vector $\mathbf{e}$ is consistent with the empirical data. Below procedures are described which assess the agreement between the empirical distribution and the normal theory.

Let $\mathscr{F}$ denote the set of normal cumulative distribution functions (c.d.f.) and Ω denote the set of c.d.f. with finite four-order central moment and with their third-order central moment different from null. Alternatively, the set Ω might be defined as the set of all non-normal c.d.f. (or for sake of convenience, those with finite second moments). The sets $\mathscr{F}$ and Ω are disjoint. Let ξ be a random variable with unknown c.d.f. F. We formulate the null hypothesis H_0 in such a way, that the c.d.f. F belongs to the set $\mathscr{F}$, i.e. H_0: $F \in \mathscr{F}$. This hypothesis is tested against the alternative hypothesis, which states that the c.d.f. F belongs to the set Ω, in other words against H_1: $F \in \Omega$.

The procedure of testing the hypothesis H_0 is carried out by a test of normality (e.g. W test, Shapiro—Wilk, [13]). In the statistical literature the application of these tests is treated mostly for simple samples. Before applying one of the well-known tests of normality for the earlier described linear model, it is necessary to transform the vector $\mathbf{y}$ in such a way that the new random vector should have elements satisfying the conditions for simple sample. To find such transformations is our basic task, they enable the assessment of the validity of the assumption on normality in linear models.

As it is known, in the analysis of the linear model the least square method (LS method) is applied. It allows to find an estimator $\boldsymbol{\beta}^0$ of the vector parameter $\boldsymbol{\beta}$ of model (1.1), which minimizes the quadratic form $(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}^0)'(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}^0)$. Then the components of the vector $\mathbf{X}\boldsymbol{\beta}^0$ are best linear unbiased estimators (BLU estimators) of the components of the vector $\mathbf{X}\boldsymbol{\beta}$. We introduce a vector $\boldsymbol{\varepsilon}$ as follows:

Definition 1.1. The vector $\boldsymbol{\varepsilon}=\mathbf{y}-\mathbf{X}\boldsymbol{\beta}^0$, where $\boldsymbol{\beta}^0$ is an arbitrary solution of $\mathbf{X}'\mathbf{X}\boldsymbol{\beta}^0=\mathbf{X}'\mathbf{y}$, is called the residual vector of the least square method (r.v.).

The r.v. is to be found. The general solution of the equation $\mathbf{X}'\mathbf{X}\boldsymbol{\beta}^0=\mathbf{X}'\mathbf{y}$ is of the form $\boldsymbol{\beta}^0=\mathbf{G}\mathbf{X}'\mathbf{y}+(\mathbf{I}-\mathbf{G}\mathbf{X}'\mathbf{X})\mathbf{z}$, where $\mathbf{z}$ is an arbitrary vector of suitable dimension, $\mathbf{G}=(\mathbf{X}'\mathbf{X})^-$ is a g-inverse of the matrix $\mathbf{X}'\mathbf{X}$. Hence we have $\mathbf{X}\boldsymbol{\beta}^0=\mathbf{X}\mathbf{G}\mathbf{X}'\mathbf{y}+\mathbf{X}(\mathbf{I}-\mathbf{G}\mathbf{X}'\mathbf{X})\mathbf{z}=\mathbf{X}\mathbf{G}\mathbf{X}'\mathbf{y}$ according to the equation $\mathbf{X}\mathbf{G}\mathbf{X}'\mathbf{X}=\mathbf{X}$ which is satisfied by any matrix $\mathbf{X}$. Therefore the vector $\boldsymbol{\varepsilon}$ in the model $(\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ gets a form

$$\boldsymbol{\varepsilon} = \mathbf{y}-\mathbf{X}\mathbf{G}\mathbf{X}'\mathbf{y} = (\mathbf{I}-\mathbf{X}\mathbf{G}\mathbf{X}')\mathbf{y} = \boldsymbol{\Psi}\mathbf{y}, \tag{1.2}$$

where $\boldsymbol{\Psi}=\mathbf{I}-\mathbf{X}(\mathbf{X}'\mathbf{X})^-\mathbf{X}'$ is an $n\times n$ matrix with constant elements.

The properties of the vector $\boldsymbol{\varepsilon}$ will be discussed in Section 2.1. Because the components of the vector $\boldsymbol{\varepsilon}$ form no simple sample, we propose one of following linear transformations:

(A) $\boldsymbol{\omega}=\boldsymbol{\varepsilon}+\sigma\boldsymbol{\eta}$, where $\boldsymbol{\eta}$ is an n-dimensional normal random variable with $E(\boldsymbol{\eta})=\mathbf{0}$ and $D(\boldsymbol{\eta})=\mathbf{I}-\boldsymbol{\Psi}$, the variables $\mathbf{y}$ and $\boldsymbol{\eta}$ are independent.

(B) $\boldsymbol{\omega}=\boldsymbol{\Phi}\boldsymbol{\varepsilon}$, where $\boldsymbol{\Phi}\in\mathscr{M}_{n-r,n}$ $(r=r(\mathbf{X}))$ is a matrix satisfying the equation $\boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{\Phi}'=\mathbf{I}$, while $r(\mathbf{X})$ denote the rank of the matrix $\mathbf{X}$.

(C) $\omega = \mathbf{C}\mathbf{y}$, where $\mathbf{C} \in \mathcal{M}_{n-r,n}$ is a matrix satisfying the conditions $\mathbf{CX} = \mathbf{0}$ and $\mathbf{CC}' = \mathbf{I}$.

The transformations (B) and (C) are equivalent since we may put $\mathbf{C} = \boldsymbol{\Phi}\boldsymbol{\Psi}$. Conversely, if $\mathbf{C}$ is defined by the assumption of (C) then $\mathbf{Ce} = \mathbf{Cy}$.

It will be shown in Section 2.2 that the components of the vector ω satisfy the conditions for simple sample under the null hypothesis. We introduce the following definition.

DEFINITION 1.2. The vector ω which is such a transform of the random vector $\mathbf{y}$, that its components are uncorrelated random variables with expectations null and identical variance will be called an adjusted residual vector from a least square method (a.r.v.).

In this paper some methods for assessing the validity of the assumptions of normality of random error vectors in linear models are presented. The vector $\mathbf{y}$ is transformed into r.v. and then the r.v. are transformed into a.r.v. by linear transformation. Also Theil's method of calculation of the a.r.v. directly from the vector $\mathbf{y}$ is described in a generalized formulation. The suitable consideration for the model with restrictions are also included. Because the components of the a.r.v. satisfy the conditions for simple sample, therefore the appropriately chosen a.r.v. can be used for testing normality. For the practical application of the methods described suitable numerical algorithms have been constructed.

2. The methods of calculation of a.r.v.

2.1. *The properties of r.v.* The basic properties of the vector ε are connected with that of $\boldsymbol{\Psi}$.

LEMMA 2.1. *The matrix $\boldsymbol{\Psi} \in \mathcal{M}_{n,n}$ is symmetrical, idempotent, is of rank $n - r(\mathbf{X})$ and is invariant in respect of the choice of $\boldsymbol{\beta}^0$ (or in respect of choice of $\mathbf{G}$).*

The next properties of the matrix $\boldsymbol{\Psi}$ are as follows:

(a) The matrix $\boldsymbol{\Psi}$ can be presented in form $\boldsymbol{\Psi} = \mathbf{I} - \mathbf{XX}^+$, where $\mathbf{X}^+$ is a Moore—Penrose inverse of the matrix $\mathbf{X}$;

(b) The matrix $\boldsymbol{\Psi}$ satisfies the equations: $\boldsymbol{\Psi}\mathbf{X} = \mathbf{0}$ and $\mathbf{X}'\boldsymbol{\Psi} = \mathbf{0}$.

The properties of the vector ε are the following:

(a) The covariance matrix of the vector ε can be presented in the form $D(\varepsilon) = D(\mathbf{y}) - D(\mathbf{X}\boldsymbol{\beta}^0)$;

(b) The vector ε is uncorrelated with the BLU estimator $\mathbf{B}\boldsymbol{\beta}^0$ $(\mathbf{B} \in \mathcal{M}_{n,q})$, of the set of parametric functions $\mathbf{B}\boldsymbol{\beta}$;

(c) The vector ε is orthogonal to every column of the matrix $\mathbf{X}$;

(d) The vector ε may be expressed in the form $\varepsilon = (\mathbf{I} - \mathbf{XX}^+)\mathbf{y}$;

(e) The scalar product of a vector $\mathbf{X}\boldsymbol{\beta}^0$ and ε is zero;

(f) The sum of squares $\mathbf{y}'\mathbf{y}$ may be expressed in the form $\mathbf{y}'\mathbf{y} = \varepsilon'\varepsilon + \mathbf{y}'\mathbf{X}\boldsymbol{\beta}^0$.

The assumptions, taken for the model (1.1) imply $E(\varepsilon) = \mathbf{0}$. The theorem below determines the covariance matrix of the vector ε.

THEOREM 2.1. *The covariance matrix of the vector ε is $D(\varepsilon) = \sigma^2\boldsymbol{\Psi}$ and is a null matrix if and only if $n = q$ and $r(\mathbf{X}) = q$.*

The components of the r.v. are linear functions of the components of vector $\mathbf{y}$. It is interesting to see, that the vector $\boldsymbol{\varepsilon}$ has minimal norm in a class of all vectors $\tilde{\boldsymbol{\varepsilon}}$ of the form $\tilde{\boldsymbol{\varepsilon}}=\mathbf{y}-\mathbf{X}\tilde{\boldsymbol{\beta}}$, where $\tilde{\boldsymbol{\beta}}$ is a linear function of vector $\mathbf{y}$ such that $E(\tilde{\boldsymbol{\varepsilon}})=\mathbf{0}$. It is a result of the LS method (Scheffé, [12]).

2.2. *The calculation of the a.r.v.* There exist a number of methods for calculating of the a.r.v. (see Definition 1.2) for the model $(\mathbf{y}, \mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ with desirable properties. We present five methods for calculation of the vector $\boldsymbol{\omega}$, four of them are connected with the transformation of vector $\boldsymbol{\varepsilon}$, two of type (A) and two of type (B), and one with the transformation of vector $\mathbf{y}$ of type (C).

Let $\boldsymbol{\tau}$ be a normal random vector of suitable dimension with parameters $E(\boldsymbol{\tau})=\mathbf{0}$ and $D(\boldsymbol{\tau})=\mathbf{I}$.

METHOD I. Let $\mathbf{P}\in\mathcal{M}_{q,q}$ be a matrix satisfying the equation $\mathbf{P}\mathbf{P}'=(\mathbf{X}'\mathbf{X})^-$ and let $\boldsymbol{\tau}$ be a q-dimensional normally distributed random variable. We construct the variable $\boldsymbol{\eta}=\mathbf{X}\mathbf{P}\boldsymbol{\tau}$ and define the a.r.v. (Golub, et al., [6], p. 64) in the form

$$\boldsymbol{\omega}_I = \boldsymbol{\varepsilon}+\sigma\mathbf{X}\mathbf{P}\boldsymbol{\tau},$$

with distribution parameters $E(\boldsymbol{\omega}_I)=E(\boldsymbol{\varepsilon})+\sigma\mathbf{X}\mathbf{P}E(\boldsymbol{\tau})=\mathbf{0}$ and $D(\boldsymbol{\omega}_I)=D(\boldsymbol{\varepsilon})+ +\sigma^2\mathbf{X}\mathbf{P}\mathbf{P}'\mathbf{X}'=\sigma^2\boldsymbol{\Psi}+\sigma^2(\mathbf{I}-\boldsymbol{\Psi})=\sigma^2\mathbf{I}$. The components of the vector $\boldsymbol{\tau}$ can be selected from a table of random numbers or can be generated by Monte Carlo method.

Hence, if $\mathbf{y}\sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$ and $\boldsymbol{\tau}\sim N(\mathbf{0}, \mathbf{I})$, then $\boldsymbol{\omega}_I\sim N(\mathbf{0}, \sigma^2\mathbf{I})$, i.e. the components of $\boldsymbol{\omega}_I$ are independent random variables each with the distribution $N(0, \sigma^2)$. The above method has a disadvantage caused by the fact that many matrices $\mathbf{P}$ satisfy the equation $\mathbf{P}\mathbf{P}'=\mathbf{G}$, and also on many ways one may select a g-inverse of the matrix $\mathbf{G}$. Moreover this method depends on the selection of the components of the vector $\boldsymbol{\tau}$. This disadvantage in the choice of the matrix $\mathbf{P}$ and of the g-inverse does not influence the probability of the acceptance of the null hypothesis.

METHOD II. Let $\boldsymbol{\tau}$ be an n-dimensional random variable. We define $\boldsymbol{\eta}=(\mathbf{I}-\boldsymbol{\Psi})\boldsymbol{\tau}$ and introduce the a.r.v. in the form (Golub, et al., [6] , p. 65)

$$\boldsymbol{\omega}_{II} = \boldsymbol{\varepsilon}+\sigma(\mathbf{I}-\boldsymbol{\Psi})\boldsymbol{\tau}$$

with distribution parameters $E(\boldsymbol{\omega}_{II})=E(\boldsymbol{\varepsilon})+\sigma(\mathbf{I}-\boldsymbol{\Psi})E(\boldsymbol{\tau})=\mathbf{0}$ and $D(\boldsymbol{\omega}_{II})=D(\boldsymbol{\varepsilon})+ +\sigma^2(\mathbf{I}-\boldsymbol{\Psi})D(\boldsymbol{\tau})(\mathbf{I}-\boldsymbol{\Psi})'=\sigma^2\boldsymbol{\Psi}+\sigma^2(\mathbf{I}-\boldsymbol{\Psi})(\mathbf{I}-\boldsymbol{\Psi})=\sigma^2\boldsymbol{\Psi}+\sigma^2(\mathbf{I}-\boldsymbol{\Psi})=\sigma^2\mathbf{I}$ where the matrix $\mathbf{I}-\boldsymbol{\Psi}$ is idempotent.

With the definition of the vector $\boldsymbol{\omega}_{II}$ we can see, that values of its components depend only on the component values of vector $\boldsymbol{\tau}$, but are independent from the selection of g-inverse of the matrix $\mathbf{G}$. Certainly, if $\boldsymbol{\tau}\sim N(\mathbf{0}, \mathbf{I})$, then $\boldsymbol{\eta}=\mathbf{X}\mathbf{G}\mathbf{X}'\boldsymbol{\tau}\sim \sim N(\mathbf{0}, \mathbf{X}\mathbf{G}\mathbf{X}')$ and moreover if $\mathbf{y}\sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I})$, then $\boldsymbol{\omega}_{II}\sim N(\mathbf{0}, \sigma^2\mathbf{I})$. The selection $\boldsymbol{\tau}$ is done analogically as in the Method I. For numerical calculation the form

$$\boldsymbol{\omega}_{II} = \mathbf{y}-\mathbf{X}\mathbf{X}^+(\mathbf{y}-\sigma\boldsymbol{\tau})$$

is more convenient, because the calculation of the matrix $\mathbf{X}\mathbf{X}^+$ is equivalent with calculating orthogonal projection operators for the space of columns of the matrix $\mathbf{X}$.

METHOD III. We select the matrix $\mathbf{B}\in\mathcal{M}_{n,n}$ (Rao, [19], p. 19) in such a way that $\mathbf{B}\boldsymbol{\Psi}\mathbf{B}'=\mathbf{D}$, where $\mathbf{D}\in\mathcal{M}_{n,n}$ is a diagonal matrix with $n-r$ positive diagonal elements and with the last diagonal elements equal to zero. Moreover, we select the matrix $\mathbf{M} \in \mathcal{M}_{n,n-r}$ the columns of which are unit vectors with 1 on the position indicating a row of the matrix $\mathbf{D}$ with positive element. We see, that $\mathbf{M}'\mathbf{M} = \mathbf{I}$. We construct, in addition, the matrix $\mathbf{D}^{-1/2}$ according to the following: if in the matrix $\mathbf{D}$ the element $d_{ii} \neq 0$, then the diagonal element of the matrix $\mathbf{D}^{-1/2}$ is $1/\sqrt{d_{ii}}$, in the opposite case it is equal to zero. Assuming that $\boldsymbol{\Phi} = \mathbf{M}'\mathbf{D}^{-1/2}\mathbf{B}$ the a.r.v. is built up of $n-r$ elements and has the form

$$\omega_{\text{III}} = \mathbf{M}'\mathbf{D}^{-1/2}\mathbf{B}\boldsymbol{\varepsilon},$$

for which $\boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{\Phi}' = \mathbf{M}'\mathbf{D}^{-1/2}\mathbf{B}\mathbf{B}'\mathbf{D}^{-1/2}\mathbf{M} = \mathbf{M}'\mathbf{D}^{-1/2}\mathbf{D}\mathbf{D}^{-1/2}\mathbf{M} = \mathbf{M}'\mathbf{M} = \mathbf{I}$, $E(\omega_{\text{III}}) = \mathbf{0}$ and $D(\omega_{\text{III}}) = \sigma^2\mathbf{I}$.

METHOD IV. The latent roots λ_j, $j = 1, \ldots, n$, of the matrix $\boldsymbol{\Psi}$ are 1 and 0, with multiplicities $n-r$ and r, respectively. We denote by $\mathbf{c}_j$, $j = 1, \ldots, n-r$ the n-dimensional latent vectors for the latent root $\lambda_j = 1$, then $\boldsymbol{\Psi}\mathbf{c}_j = \mathbf{c}_j$. We assume, that the vectors $\mathbf{c}_j$ are orthonormal. This means we may accomplish a decomposition

$$\boldsymbol{\Psi} = \sum_{j=1}^{n-r} \mathbf{c}_j\mathbf{c}_j' = \mathbf{Q}'\mathbf{Q},$$

where $\mathbf{Q}' = (\mathbf{c}_1, \ldots, \mathbf{c}_{n-r})$ satisfies $\mathbf{Q}\mathbf{Q}' = \mathbf{I}$. We construct the $(n-r)$-dimensional a.r.v. in the form

$$\omega_{\text{IV}} = \mathbf{Q}\boldsymbol{\varepsilon},$$

where $\boldsymbol{\Phi}\boldsymbol{\Psi}\boldsymbol{\Phi}' = \mathbf{Q}\mathbf{Q}'\mathbf{Q}\mathbf{Q}' = \mathbf{I}$ and also $E(\omega_{\text{IV}}) = \mathbf{0}$ and $D(\omega_{\text{IV}}) = \sigma^2\mathbf{I}$.

The selection of vectors $\mathbf{c}_1, \ldots, \mathbf{c}_{n-r}$ may be accomplished by the Gram—Schmidt orthogonalization method (see, e.g. Birkhoff and MacLane, [2]). This selection can be carried out on many ways. Therefore the selection of the matrix $\mathbf{Q}$ is not unique.

The null distribution of ω_{III} and ω_{IV} is normal, but their distributions under the alternative hypothesis is not investigated. It is clear that this distribution differs from that of $\mathbf{e}$ under the alternative hypothesis. Unsuitable choice of the matrix $\boldsymbol{\Phi}$ may cause that the distribution of ω will be nearly normal even if that of $\mathbf{e}$ differs from the normal law markedly. It is known, e.g. that the distribution of the sample mean usually tends to the normal law very rapidly. Accordingly, it should be avoided that the absolute values of the elements of the row-vectors of $\boldsymbol{\Phi}\boldsymbol{\Psi}$ should be equal or nearly equal. On the other hand it seems advantageous that one element in each row should have high and all the others low absolute values (cf. Method V).

From numerical point of view Method II is the best of the presented methods, beyond the generation of random vector τ, there is no need to execute any additional numerical operation connected with the decomposition of the matrix $\mathbf{G}$ (Method I) or of the matrix $\boldsymbol{\Phi}$ (Methods III and IV).

METHOD V (Theil's method). Here a generalized formulation of the method of Theil [16] is given. He assumed that $r(\mathbf{X}) = q$, we, however, admit that $r(\mathbf{X}) \leq q$. The matrix $\mathbf{C}$ satisfies the conditions $\mathbf{C}\mathbf{X} = \mathbf{0}$ and $\mathbf{C}\mathbf{C}' = \mathbf{I}$ if and only if its row vectors are orthonormal and belong to an orthogonal complement of space spanned by the columns of the matrix $\mathbf{X}$. To achieve this aim we propose a partition of $\mathbf{y}$ on two subvectors with r and $n-r$ components. Because of the equality $\mathbf{X}\boldsymbol{\beta} + \mathbf{e} = \mathbf{X}\boldsymbol{\beta}^0 + \boldsymbol{\varepsilon}$

the vector $\mathbf{y}$ can be written as

$$\begin{bmatrix}\mathbf{y}_0\\ \mathbf{y}_1\end{bmatrix} = \begin{bmatrix}\mathbf{X}_0\\ \mathbf{X}_1\end{bmatrix}\beta + \begin{bmatrix}\mathbf{e}_0\\ \mathbf{e}_1\end{bmatrix} = \begin{bmatrix}\mathbf{X}_0\\ \mathbf{X}_1\end{bmatrix}\beta^0 + \begin{bmatrix}\varepsilon_0\\ \varepsilon_1\end{bmatrix} \tag{2.1}$$

where the matrices with subscript 0 have r rows and those with subscript 1 have $n-r$ rows. The matrix $\mathbf{X}_0 \in \mathcal{M}_{r,q}$ is of full row rank. Our objective is to find a linear unbiased a.r.v. with a covariance matrix $\sigma^2\mathbf{I}$, that corresponds to the $(n-r)$-dimensional disturbance subvector $\mathbf{e}_1$. By the partition (2.1) we calculate the $(n-r)$-dimensional a.r.v. as following. As we assume, that $r(\mathbf{X})=r$, then also $r(\mathbf{X}\mathbf{X}_0^+)=r$ and the matrix $(\mathbf{X}\mathbf{X}_0^+)'\mathbf{X}\mathbf{X}_0^+ \in \mathcal{M}_{r,r}$ is symmetrical and positive definite. Its inverse matrix $\mathbf{X}_0(\mathbf{X}'\mathbf{X})^+\mathbf{X}_0'$ is positive definite and invariant in respect of the selection of the Moore—Penrose inverse of the matrix $\mathbf{X}'\mathbf{X}$. All latent roots d_h^2, $h=1, \ldots, r$ of the matrix $\mathbf{X}_0(\mathbf{X}'\mathbf{X})^+ \times \times \mathbf{X}'$ are positive and $\leqq 1$. Denote by k $(\leqq r)$ the number of latent roots $d_h^2 < 1$, and the corresponding r-dimensional orthonormal characteristic vectors by $\mathbf{v}_1, \ldots, \mathbf{v}_k$. Let d_h, $h=1, \ldots, k$ be the positive square roots of the numbers d_h^2. The a.r.v. are defined as follows

$$\omega_{\text{V}} = \varepsilon_1 - \mathbf{X}_1\mathbf{X}_0^+ \sum_{h=1}^{k} \frac{d_h}{1+d_h} \mathbf{v}_h\mathbf{v}_h'\varepsilon_0$$

and we have $\omega_{\text{V}}'\omega_{\text{V}} = \varepsilon'\varepsilon$.

The disadvantage of this method is a great number of possible partitions of the matrix $\mathbf{X}$ in (2.1). There exist $\binom{n}{r}$ partitions and for every partition the a.r.v. can be calculated. We propose to select such a vector ω_{V}, for which (cf. Theil, [16] p. 206) the expected sum of squares of the error vector

$$E[(\omega_{\text{V}} - \mathbf{e}_1)'(\omega_{\text{V}} - \mathbf{e}_1)] = 2\sigma^2 \sum_{h=1}^{k} (1 - d_h),$$

is minimalized.

Let be remarked, too, that the distributions of the components of ω_{III}, are all different under the alternative hypothesis even in such cases when method V yields identically distributed variables (see, e.g., Sarkadi [11]). The latter situation seems more advantageous from the point of view of testing for normality.

Methods I—II are based on the assumption that the parameter σ^2 is known. If σ^2 in unknown, it may be replaced by the unbiased estimator s^2 of the parameter σ^2 in the form $s^2 = \varepsilon'\varepsilon/(n-r)$. The results are the more accurate the greater the number of degrees of freedom $n-r$ is. Alternatively, we may apply the idea of Durbin ([4], p. 52) if σ^2 is unknown. Accordingly ω_{II} is to be defined as

$$\omega_{\text{II}} = s_0\varepsilon/s + (\mathbf{I} - \boldsymbol{\Psi})\boldsymbol{\tau}$$

where s_0^2 is a random number which is χ^2-distributed with $n-r$ degrees of freedom. In this case $\boldsymbol{\omega}_{\text{II}}$ is normal, otherwise this is not true.

2.3. *Calculation of the a.r.v. in model with restriction.* The calculation of the a.r.v. is the model $(\mathbf{y}, \mathbf{X}\beta, \sigma^2\mathbf{I})$ can be carried out also under the assumption, that the parameter vector β fulfils the consistent restrictions $\mathbf{R}\beta = \mathbf{c}$, where $\mathbf{R} \in \mathcal{M}_{k,q}$ and $\mathbf{c} \in \mathcal{M}_{k,1}$ are known. The model with restrictions has the notation $(\mathbf{y}, \mathbf{X}\beta | \mathbf{R}\beta = \mathbf{c}, \sigma^2\mathbf{I})$. This model can be reduced to a model without restrictions (Rao and Mitra,

[10], p. 144)

$$(\mathbf{y}-\mathbf{XR}^{-}\mathbf{c},\ \mathbf{X}(\mathbf{I}-\mathbf{R}^{-}\mathbf{R})\theta,\ \sigma^2\mathbf{I}), \tag{2.2}$$

where the matrix $\mathbf{X}(\mathbf{I}-\mathbf{R}^{-}\mathbf{R})$ is a design matrix, $\mathbf{y}-\mathbf{XR}^{-}\mathbf{c}$ is the observable vector and $\theta\in\mathcal{M}_{q,1}$ the parameter vector. The imposed restrictions for the model do not change the distribution parameters of the a.r.v. Therefore the calculation of the a.r.v. for model (2.2) is carried out analogically as described above.

3. The numerical aspect of the calculation of a.r.v. and testing the hypothesis on the normality of errors

3.1. *The algorithms of calculation of the a.r.v.*

METHOD I. 1. Calculate the matrices $\mathbf{XX}^+$ and $\Psi=\mathbf{I}-\mathbf{XX}^+$.

2. Compute the r.v. If σ^2 is unknown, compute $s^2=\varepsilon'\varepsilon/(n-r)$.

3. Select the matrix $\mathbf{P}\in\mathcal{M}_{q,q}$ to satisfy the equation $(\mathbf{X}'\mathbf{X})^{-}=\mathbf{PP}'$, by the generalized Cholesky method.

4. Generate by computer, by Monte Carlo method, the random vector τ and s_0.

5. Compute the vector $\eta=s\mathbf{XP}\tau/s_0$ (or $\sigma\mathbf{XP}\tau$), where $s=\sqrt{s^2}$, and $\omega_{\mathrm{I}}=\varepsilon+\eta$.

METHOD II. The points 1, 2, 4 are the same as in Method I. Compute the vector $\omega_{\mathrm{II}}=\mathbf{y}-\mathbf{XX}^+(\mathbf{y}-s\tau/s_0)$.

METHOD III. 1. Same as point 1 in the Method I.

2. Compute the nonsingular matrix $\mathbf{B}$ by $\mathbf{B}\Psi\mathbf{B}'=\mathbf{D}$, where $\mathbf{D}$ is a diagonal matrix with $n-r$ positive elements, in the following way. Decompose the matrix Ψ for $\Psi=\mathbf{SS}'$, where $\mathbf{S}$ is triangular matrix of the form $\mathbf{S}=\mathbf{LD}^{1/2}$ while $\mathbf{D}^{1/2}$ is a diagonal matrix with diagonal elements equal to the corresponding elements of the matrix $\mathbf{S}$. $\mathbf{L}$ is a triangular matrix with all diagonal elements equal to one and the other elements are calculated by the equation

$$l_{ij}=\begin{cases} s_{ij}/s_{jj}, & \text{when} \quad s_{jj}\neq 0,\\ 0, & \text{when} \quad s_{jj}=0,\end{cases}$$

for $j=1,\dots,n-1$, $i=j+1,\dots,n$. We have the decomposition of the matrix $\Psi=\mathbf{SS}'=\mathbf{LD}^{1/2}(\mathbf{LD}^{1/2})'=\mathbf{LDL}'$ and therefore $\mathbf{B}=\mathbf{L}^{-1}$.

3. Compute the matrix $\mathbf{M}\in\mathcal{M}_{n,n-r}$ and $\mathbf{D}^{-1/2}$ and compute the a.r.v. in form ω_{III}.

METHOD IV. 1. Same as point 1 in the Method I.

2. Split the matrix $\mathbf{X}$ in two submatrices $\mathbf{X}=[\mathbf{X}_0'\vdots\mathbf{X}_1']'$, where $\mathbf{X}_0$ is a matrix of full row rank and $\varepsilon=[\varepsilon_0'|\varepsilon_1']'$.

3. Compute the matrices $\mathbf{X}_0^+=\mathbf{X}_0'(\mathbf{X}_0\mathbf{X}_0')^{-1}$ and $\mathbf{X}_1\mathbf{X}_0^+$.

4. Compute the matrix $(\mathbf{X}'\mathbf{X})^+$ (see Wilkinson and Reinsch, [17], p. 144).

5. Compute the matrix $\mathbf{X}_0(\mathbf{X}'\mathbf{X})^+\mathbf{X}_0'$ and its latent roots d_h^2, $h=1,\dots,k$ for which $d_h^2<1$.

6. Compute the orthonormal latent vectors $\mathbf{v}_1,\dots,\mathbf{v}_k$ of the matrix $\mathbf{X}_0(\mathbf{X}'\mathbf{X})^+\mathbf{X}_0'$ corresponding to the latent roots $d_1^2,\dots,d_k^2$ and compute ω_{V}.

3.2. *Assessing the validity of the assumption on normality of the random error vector by the Shapiro—Wilk test.*

A. Model $(\mathbf{y}, \mathbf{X}\beta, \sigma^2\mathbf{I})$. 1. Calculate the a.r.v. by one of the methods presented in Section 3.1. From now on, the components of the a.r.v. will be treated as simple sample with elements $x_1, \ldots, x_m$.

2. Compute the sum of squares of deviations from the sample

$$S^2 = \mathbf{x}'(\mathbf{I}-\mathbf{1}\mathbf{1}'/m)\mathbf{x}.$$

3. Order the sample $\{x_i\}$ into an ordered sample $y_1 \leqq y_2 \leqq \ldots \leqq y_m$.

4. Read the coefficients $a_{i,m}$, $i=1, \ldots, [m/2]$ (where $[m/2]$ is entier of $m/2$) from the table of Shapiro and Wilk ([13], p. 596).

5. (a) if m is even $(m=2l)$, compute

$$b = a_{2l,m}(y_{2l}-y_1)+a_{2l-1,m}(y_{2l-1}-y_2)+\ldots+a_{l+1,m}(y_{l+1}-y_l),$$

(b) if m is odd $(m=2l+1)$, then compute

$$b = a_{2l+1,m}(y_{2l+1}-y_1)+a_{2l,m}(y_{2l}-y_2)+\ldots+a_{l+2,m}(y_{l+2}-y_l).$$

6. Compute $W=b^2/S^2$.

7. The hypothesis on the normality of the random error vectors in model (1.1) will be rejected when $W \leqq W_{\alpha,m}$, where $W_{\alpha,m}$ is the $100\alpha\%$ point of the distribution of W.

B. Model $(\mathbf{y}, \mathbf{X}\beta | \mathbf{R}\beta=\mathbf{c}, \sigma^2\mathbf{I})$. 1. Compute the matrices $\mathbf{R}^-$, $\mathbf{R}^-\mathbf{R}$ and $\mathbf{X}(\mathbf{I}-\mathbf{R}^-\mathbf{R})$ and the vector $\mathbf{y}=\mathbf{X}\mathbf{R}^-\mathbf{c}$.

2. Compute for the model $(\mathbf{y}-\mathbf{X}\mathbf{R}^-\mathbf{c}, \mathbf{X}(\mathbf{I}-\mathbf{R}^-\mathbf{R})\theta, \sigma^2\mathbf{I})$ the a.r.v. with one of the methods presented in Section 3.1.

3. Next as in points 2—7 in the model $(\mathbf{y}, \mathbf{X}\beta, \sigma^2\mathbf{I})$.

Acknowledgement. I am indebted to three anonymous referees of this paper for their careful reading of this paper, and for valuable comments which helped me to improve its organization and broaden its content.

REFERENCES

[1] Benhelli, M., Contribution à l'étude de l'autocorrelation des erreurs dans le modèle linéaire, Thèse de 3ème cycle, Université Paul Sabatier, Toulouse, 1980.

[2] Birkhoff, G. and MacLane, S., *A survey of modern algebra*, Macmillan Co., New York, 1941. *MR* **3**—99; Rev. ed. Macmillan Co., New York, 1953. *MR* **14**—939.

[3] Caussinus, H., Sur l'analyse des résidus dans le modèle linéaire, *Statist. Anal. Données* **3** (1980), 29—39.

[4] Durbin, J., Some methods of constructing exact tests, *Biometrika* **48** (1961), 41—55. *MR* **23**# A3609.

[5] Farebrother, R. W., BLUS Residuals, *Applied Statistics* **25** (1976), 317—322.

[6] Golub, G. H., Guttman, I. and Dutter, R., Examination of pseudo-residuals of outliers for detecting spurosity in the general univariate linear model, *Multivariate Statistical Inference* (Proc. Res. Sem., Dalhousie Univ., Halifax, N. S., 1972), North-Holland Publ. Co., Amsterdam; American Elsevier, New York, 1973, 63—108. *MR* **51** #9380.

[7] Graybill, F. A., *An introduction to linear statistical models*, Vol. 1, McGraw-Hill Series in Probability and Statistics, McGraw-Hill Book Co., Inc., New York—Toronto—London, 1961. *MR* **23**# A3612.

[8] Koerts J. and Abrahamse, A. P. J., On the power of the BLUS procedure, *J. Amer. Statist. Assoc.* **63** (1968), 1227—1236. *MR* **38** # 3967.
[9] Rao, C. R., *Linear statistical inference and its applications*, John Wiley and Sons, Inc., New York—London—Sydney, 1965. *MR* **36** # 4668 and *Zbl.* **137**, 362.
[10] Rao, C. R. and Mitra, S. K., *Generalized inverse of matrices and its applications*, John Wiley and Sons Inc., New York—London—Sydney, 1971. *MR* **49** # 2780.
[11] Sarkadi, K., On testing for normality, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5** (1960), 269—275. *MR* **25** # 4603.
[12] Scheffé, H., *The analysis of variance*, John Wiley and Sons Inc., New York; Chapman and Hall, Ltd., London, 1959. *MR* **22** # 7217.
[13] Shapiro, S. S. and Wilk, M. B., An analysis of variance test for normality: complete samples, *Biometrika* **52** (1965), 591—611. *MR* **34** # 5213.
[14] Theil, H., The analysis of disturbance in regression analysis, *J. Amer. Statist. Assoc.* **60** (1965), 1067—1079. *MR* **32** # 8455.
[15] Theil, H., A simplification of the BLUS procedure for analysing regression disturbances, *J. Amer. Statist. Assoc.* **63** (1968), 242—251. *MR* **37** # 1031.
[16] Theil, H., *Principles of econometrics*, North-Holland Publishing Co., Amsterdam, 1971.
[17] Wilkinson, J. H. and Reinsch, C., *Linear Algebra (Handbook for Automatic Computation*, vol. 2), Springer-Verlag, Berlin—New York, 1971.

(Received October 23, 1981)

AKADEMIA ROLNICZA W POZNANIU
ZAKŁAD METOD MATEMATYCZNYCH I STATYSTYCZNYCH
UL. WOJSKA POLSKIEGO 28
PL—60-637 POZNAŃ
POLAND

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 403--408

UPPER ESTIMATE FOR THE EIGENVALUES OF AN ORTHONORMAL SYSTEM CONSISTING OF EIGENFUNCTIONS OF A LINEAR DIFFERENTIAL OPERATOR

V. KOMORNIK

Let $G=[a, b]\subset\mathbf{R}$ be a compact interval, $p_2, \ldots, p_n$: $G\to\mathbf{C}$ Lebesgue integrable functions and consider the formal differential operator

$$Lu = u^{(n)}+p_2u^{(n-2)}+\ldots+p_nu.$$

As it is usual, a function not a.e. vanishing u: $G\to\mathbf{C}$ is called the eigenfunction of the operator L with the eigenvalue $\lambda\in\mathbf{C}$ provided that u, together with its first $n-1$ derivatives is absolute continuous on G and for almost all $x\in G$,

$$(Lu)(x) = \lambda u(x).$$

Let now $(u_m)\subset L^2(G)$ be an arbitrary orthonormal system, consisting of eigenfunctions of the operator L and denote λ_m the eigenvalue of u_m. Then a natural question is whether the sequence (λ_m) may have a (finite) cluster point. It is well-known that for the case $n=2$ the answer is negative if $p_2\in L^q(G)$ for some $q>1$: it was shown by V. A. Il'in and I. Joó in [2] for the case when all the eigenvalues are real and nonnegative and after it by I. Joó in [4] for the general case.

Developing the ideas of the papers [2], [3], [4] and applying some results of [6] we shall prove that the answer is always negative (even if the functions $p_2, \ldots, p_n$ are only integrable):

THEOREM. *Let $G=[a, b]\subset\mathbf{R}$ be a compact interval, $p_2, \ldots, p_n$: $G\to\mathbf{C}$ Lebesgue integrable functions and $(u_m)\subset L^2(G)$ an arbitrary orthonormal system, consisting of eigenfunctions of the operator*

$$Lu = u^{(n)}+p_2u^{(n-2)}+\ldots+p_nu.$$

Then, denoting the eigenvalue of u_m by λ_m, we have

$$|\lambda_m|\to+\infty \qquad (m\to\infty).$$

Moreover, there exists a constant $\mathcal{K}$, depending only on the numbers $|b-a|$, $\|p_2\|_1, \ldots$ $\ldots, \|p_n\|_1$ such that — denoting by μ_m an arbitrary nth root of λ_m — for any $\mu\in\mathbf{C}$,

$$\sum_{|\mu_m-\mu|\leq 1} 1 < \mathcal{K}(1+|\mu|).$$

1980 *Mathematics Subject Classification.* Primary 47E05; Secondary 45C05.
Key words and phrases. Eigenfunction.

Let us introduce for brevity the following notations: for any $\lambda\in\mathbf{C}$ and $t\in\mathbf{R}$ define

$$D(\lambda, t) \equiv \begin{cases} \begin{vmatrix} e^{\varrho_1 t} & \dots & e^{\varrho_n t} \\ \vdots & \dots & \vdots \\ e^{n\varrho_1 t} & \dots & e^{n\varrho_n t} \end{vmatrix} & \text{if } \lambda \neq 0, \\ \begin{vmatrix} 1 & t & \dots & t^{n-1} \\ \vdots & & \dots & \vdots \\ 1 & nt & \dots & (nt)^{n-1} \end{vmatrix} & \text{if } \lambda = 0, \end{cases}$$

where $\varrho_1, \dots, \varrho_n$ are the different nth roots of λ for $\lambda\neq 0$, in the increasing order of their arguments, taking in the interval $[0, 2\pi)$. Obviously, $D(\lambda, t)\neq 0$ if $t\neq 0$ and $|t\sqrt[n]{\lambda}|<\pi$. Let $D_{kq}(\lambda, t)$ denote the minor of $D(\lambda, t)$, corresponding to the q-th element of the k-th row, and define for any $1\leq k\leq n$, $\lambda\in\mathbf{C}$, $0\neq t\in\mathbf{R}$, $|t\sqrt[n]{\lambda}|<\pi$,

$$f_k(\lambda, t) \equiv \begin{cases} \sum_{q=1}^{n} \dfrac{D_{kq}(\lambda, t)}{D(\lambda, t)} & \text{if } \lambda \neq 0, \\ \dfrac{D_{k1}(\lambda, t)}{D(\lambda, t)} & \text{if } \lambda = 0. \end{cases}$$

Let us also introduce the function $g\colon \mathbf{C}\times\mathbf{R}\to\mathbf{C}$,

$$g(\lambda, t) \equiv \begin{cases} \sum_{q=1}^{n} \dfrac{\varrho_q}{n\lambda} e^{\varrho_q t} & \text{if } \lambda \neq 0, \\ \dfrac{t^{n-1}}{(n-1)!} & \text{if } \lambda = 0. \end{cases}$$

Recall the following proposition, proved in [6]:

PROPOSITION. *Let u be an arbitrary eigenfunction of the operator L with the eigenvalue λ. Then for any $x\in G$, $x+nt\in G$, $t\neq 0$, $|t\sqrt[n]{\lambda}|<\pi$, the following formula is valid:*

$$u(x) = \sum_{k=1}^{n} f_k(\lambda, t)\, u(x+kt) +$$

(1)

$$+ \sum_{k=1}^{n}\sum_{j=2}^{n} f_k(\lambda, t) \int_{x}^{x+kt} g(\lambda, x+kt-\tau)\, p_j(\tau)\, u^{(n-j)}(\tau)\, d\tau.$$

We shall also use the following properties:

LEMMA. (i) *For any fixed $\lambda\in\mathbf{C}$, the functions $f_k(\lambda, \cdot)$ can be continuously extended to the whole real line.*

(ii) *There exists a constant* $C>0$, *depending only on n such that for any* $1\leqq k\leqq n$, $t\in\mathbf{R}$, $\lambda=\mu^n\in\mathbf{C}$ *and* $\lambda_0=\mu_0^n\in\mathbf{C}$,

(2) $$|f_k(\lambda, t)| < C \qquad \text{if} \quad |\mu t| \leqq 1;$$

(3) $$|f_k(\lambda, t)-f_k(\lambda_0, t)| \leqq C|\mu-\mu_0||t| \quad \text{if} \quad |\mu t| \leqq 1 \quad \text{and} \quad |\mu_0 t| \leqq 1;$$

(4) $$|g(\lambda, t)| \leqq C|t|^{n-1} \qquad \text{if} \quad |\mu t| \leqq 2n.$$

Proof. (i) follows at once from [6], Lemma 5. To prove (2) and (3), we remark that by Lemmas 4, 5 in [6] we have

$$f_k(\lambda, t) = h_k(\mu t) \quad \text{for all} \quad \lambda = \mu^n\in\mathbf{C} \quad \text{and} \quad t\in\mathbf{R},$$

where $h_k\colon \mathbf{C}\to\mathbf{C}$ is a holomorphic function $(1\leqq k\leqq n)$. Finally, (4) follows from [6], Lemma 6.

We shall also use the following estimates, proved in [6]:

Proposition. *There exists a positive number D, depending only on the numbers* $|b-a|, \|p_2\|_1, \ldots, \|p_n\|_1$, *such that all the eigenfunctions u of the operator L having the eigenvalue* $\lambda=\mu^n\in\mathbf{C}$, *satisfy the inequality*

(5) $$\|u^{(j)}\|_\infty \leqq D(1+|\mu|)^{j+\frac{1}{p}}\|u\|_p, \quad 0\leqq j<n, \quad 1\leqq p\leqq\infty.$$

Let us now turn to the proof of the theorem. Let $(u_m)\subset L^2(G)$ be an orthonormal system where u_m is an eigenfunction of L with the eigenvalue $\lambda_m=\mu_m^n\in\mathbf{C}$ and let μ be an arbitrary complex number, $\lambda=\mu^n$. Define

$$R \equiv \min\left(\frac{b-a}{2n}, \frac{1}{2+|\mu|}, \frac{\varepsilon}{\|p_2\|_1}, \sqrt{\frac{\varepsilon}{\|p_3\|_1}}, \ldots, \sqrt[n-1]{\frac{\varepsilon}{\|p_n\|_1}}, \frac{1}{4nCD\sqrt{1+|\mu|}\sqrt{b-a}}\right)$$

(6) $$\text{where} \quad \varepsilon \equiv \left(4n^{n+1}C^2D^2\sqrt{1+|\mu|}\sqrt{b-a}\right)^{-1}.$$

Fix an arbitrary number $a\leqq x\leqq\dfrac{a+b}{2}$ and choose an arbitrary u_m such that $|\mu_m-\mu|\leqq 1$.

For any $0\leqq t\leqq R$ we have by (1)

$$u_m(x) = \sum_{k=1}^{n} f_k(\lambda_m, t)\,u_m(x+kt) +$$

$$+\sum_{k=1}^{n}\sum_{j=2}^{n} f_k(\lambda_m, t)\int_x^{x+kt} g(\lambda_m, x+kt-\tau)\,p_j(\tau)\,u_m^{(n-j)}(\tau)\,d\tau.$$

Integrating both sides by t from 0 to R, we obtain

$$Ru_m(x) = \sum_{k=1}^{n} \int_0^R f_k(\lambda, t)\, u_m(x+kt)\, dt +$$

$$+\sum_{k=1}^{n} \int_0^R [f_k(\lambda_m, t) - f_k(\lambda, t)]\, u_m(x+kt)\, dt + \tag{7}$$

$$+\sum_{k=1}^{n} \sum_{j=2}^{n} \int_0^R f_k(\lambda_m, t) \int_x^{x+kt} g(\lambda_m, x+kt-\tau)\, p_j(\tau)\, u_m^{(n-j)}(\tau)\, d\tau\, dt.$$

Introduce now the function $w\colon G \to \mathbf{C}$,

$$w(t) = \sum_{k=1}^{n} \frac{1}{k} \overline{f_k\left(\lambda, \frac{t-x}{k}\right)} \cdot \chi_{\{x \le t \le x+kR\}}(t);$$

it follows from (2) and (6) that

$$w \in L^2(G), \quad \|w\|_2 \le nC\sqrt{R}. \tag{8}$$

Moreover, one can easily see the equality

$$\sum_{k=1}^{n} \int_0^R f_k(\lambda, t) u_m(x+kt)\, dt = \langle u_m, w \rangle,$$

whence

$$\left|\sum_{k=1}^{n} \int_0^R f_k(\lambda, t)\, u_m(x+kt)\, dt\right| = |\langle u_m, w \rangle|. \tag{9}$$

It follows from (3), (5) and (6) that

$$\left|\sum_{k=1}^{n} \int_0^R [f_k(\lambda_m, t) - f_k(\lambda, t)] u_m(x+kt)\, dt\right| \le$$

$$\le \sum_{k=1}^{n} \int_0^R Ct\, dt \|u_m\|_\infty \le nCR^2 \|u_m\|_\infty \le \frac{1}{4} R \|u_m\|_2 (b-a)^{-\frac{1}{2}}. \tag{10}$$

Finally, by (2), (4), (5) and (6) we have

$$\left|\sum_{k=1}^{n} \sum_{j=2}^{n} \int_0^R f_k(\lambda_m, t) \int_x^{x+kt} g(\lambda_m, x+kt-\tau)\, p_j(\tau)\, u_m^{(n-j)}(\tau)\, d\tau\, dt\right| \le$$

$$\le (nR)^n C^2 \sum_{j=2}^{n} \|p_j\|_1 \|u_m^{(n-j)}\|_\infty \le$$

$$\le (nR)^n C^2 \sum_{j=2}^{n} \|p_j\|_1 D(1+|\mu_m|)^{n-j} \|u_m\|_\infty \le \tag{11}$$

$$\le n^n C^2 DR \sum_{j=2}^{n} [R^{j-1} \|p_j\|_1][R(2+|\mu|)^{n-j} \|u_m\|_\infty \le$$

$$\le n^{n+1} C^2 DR\varepsilon \|u_m\|_\infty \le \frac{1}{4} R \|u_m\|_2 (b-a)^{-\frac{1}{2}}.$$

From (7), (9), (10) and (11) we conclude

$$R|u_m(x)| \leqq |\langle u_m, w\rangle| + \frac{1}{2} R\|u_m\|_2 (b-a)^{-\frac{1}{2}}$$

and hence

$$R^2|u_m(x)|^2 \leqq 2|\langle u_m, w\rangle|^2 + \frac{1}{2} R^2\|u_m\|_2^2 (b-a)^{-1}. \tag{12}$$

Let now M be an arbitrary finite index set such that

$$|\mu_m - \mu| \leqq 1 \quad \text{for every} \quad m \in M.$$

We have by (12)

$$R^2 \sum_{m\in M} |u_m(x)|^2 \leqq 2 \sum_{m\in M} |\langle u_m, w\rangle|^2 + \frac{1}{2} R^2 \sum_{m\in M} \|u_m\|_2^2 (b-a)^{-1}$$

applying the Bessel inequality, we obtain in view of (8)

$$R^2 \sum_{m\in M} |u_m(x)|^2 \leqq 2n^2C^2R + \frac{1}{2} R^2 \sum_{m\in M} \|u_m\|_2^2 (b-a)^{-1}.$$

This is true for all $a \leqq x \leqq \frac{a+b}{2}$, but one can prove this inequality quite similarly for all $\frac{a+b}{2} \leqq x \leqq b$, too.

Integrating by x from a to b, and taking into account that

$$\|u_m\|_2 = 1, \qquad m = 1, 2, \ldots,$$

we obtain

$$\sum_{m\in M} 1 \leqq 4n^2C^2(b-a)R^{-1}.$$

The left-hand side expression is the number of elements of M, while the right-hand side does not depend on M. Therefore

$$\sum_{|\mu_m-\mu|\leqq 1} 1 \leqq 4n^2C^2(b-a)R^{-1},$$

and the theorem follows in view of (6).

REMARK 1. In case $n=2$, one can easily see that $f_2(\lambda, t) = -1$,

$$f_1(\lambda, t) = \begin{cases} 2\operatorname{ch}(\varrho_1 t) & \text{if} \quad \lambda \neq 0, \\ 2 & \text{if} \quad \lambda = 0, \end{cases}$$

$$g(\lambda, t) = \begin{cases} \dfrac{\operatorname{sh} \varrho_1 t}{t} & \text{if} \quad \lambda \neq 0, \\ t & \text{if} \quad \lambda = 0. \end{cases}$$

Therefore in the estimates (2)—(4) the conditions

$$|\mu t| \leqq 1, \quad |\mu_0 t| \leqq 1, \quad |\mu t| \leqq 2n$$

can be replaced by the weaker conditions

$$|\mathrm{Re}\,\mu t| \leqq 1, \quad |\mathrm{Re}\,\mu_0 t| \leqq 1, \quad |\mathrm{Re}\,\mu t| \leqq 2n$$

(see [5]). Therefore our theorem can be sharpened to

$$\sum_{|\mu_m-\mu|\leqq 1} 1 < \mathscr{K}(1+|\mathrm{Re}\,\mu|).$$

The only change in the proof is that we have to write in the definition of R $(2+|\mathrm{Re}\,\mu|)$ instead of $(2+|\mu|)$. This estimate was proved by I. Joó in (4) for the case when $p_2 \in L^q(G)$ for some $q>1$.

REMARK 2. The proof of our theorem equally works for the more general case when the system (u_m) is not necessarily orthonormal but the following two conditions are satisfied:

(a) $$\sup\left\{\sum_{m=1}^{\infty} |\langle w, u_m\rangle|^2 :\ w\in L^2(G), \quad \|w\|_2 \leqq 1\right\} < \infty,$$

(b) $$\inf\{\|u_m\|_2 :\ m=1,2,\ldots\} > 0.$$

Consequently, our theorem is also true if (u_m) is a Riesz basis.

REFERENCES

[1] NEUMARK, M. A., *Lineare Differentialoperatoren*, Mathematische Monographien, Bd. XI, Akademie-Verlag, Berlin, 1960. *MR* **35** # 6884 (for the English translation see *MR* **35** # 6885 and **41** # 7485).
[2] IL'IN, V. A. and JOÓ, I., Uniform estimation of eigenfunctions and an upper bound on the number of eigenvalues of the Sturm—Liouville operator with a potential from the class L^p, *Differencial'nye Uravnenija* **15** (1979), 1164—1174, 1340 (in Russian). *MR* **80i**: 34033.
[3] JOÓ, I., Upper estimates for the eigenfunctions of the Schrödinger operator, *Acta Sci. Math. (Szeged)* **44** (1982), 87—93.
[4] JOÓ, I., Spectral investigation of the Schrödinger operator, Thesis, 1979 (in Hungarian).
[5] KOMORNIK, V., Lower estimates for the eigenfunctions of the Schrödinger operator, *Acta Sci. Math. (Szeged)* **44** (1982), 95—98.
[6] KOMORNIK, V., Upper estimates for the eigenfunctions of a linear differential operator, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* (to appear).

(Received October 29, 1981)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
ANALÍZIS II. TANSZÉK
POSTAFIÓK 323
H—1445 BUDAPEST 8
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 409—416

BIVARIATE POISSON PROCESSES DEFINED ON AN ABSTRACT SPACE

MIECZYSŁAW POLAK

Summary

In this note we consider bivariate Poisson processes defined on an abstract space. The main result is given in Theorem 2. The method is based on some ideas of A. Rényi [3] and L. Győrfi [1]. Some applications of Theorem 2 are also given.

Let $\mathfrak{X}$ be an abstract set, $\mathscr{F}$ a ring of subsets of $\mathfrak{X}$, and $\lambda_1, \lambda_2, \lambda_3$ additive set functions on $\mathscr{F}$. Let $(\Omega, \mathscr{A}, \mathsf{P})$ denote a probability space, and $X(E, \omega)$, $Y(E, \omega)$ nonnegative integer-valued functions, which are for each $E \in \mathscr{F}$, random variables defined on $(\Omega, \mathscr{A}, \mathsf{P})$ and, for each $\omega \in \Omega$, additive functions on $\mathscr{F}$, i.e., for arbitrary $E_1, E_2 \in \mathscr{F}$

$$X(E_1 \cup E_2, \omega) = X(E_1, \omega) + X(E_2, \omega), \tag{1}$$
$$Y(E_1 \cup E_2, \omega) = Y(E_1, \omega) + Y(E_2, \omega)$$

provided $E_1 \cap E_2 = \emptyset$.

Further on, we assume that

1° for each $\varepsilon > 0$ and $E \in \mathscr{F}$, there exists a disjoint decomposition $E_1, E_2, \ldots$ $\ldots, E_n$ of E with $E_i \in \mathscr{F}$ and $\max_{1 \leqq j \leqq 3} \lambda_j(E_i) < \varepsilon$, $i = 1, 2, \ldots, n$;

2° for each $E, F \in \mathscr{F}$,

$$\mathsf{P}[X(E) = 0, Y(F) = 0] = e^{-[\lambda_1(E) + \lambda_2(F) + \lambda_3(E \cup F)]}, \tag{2}$$

$$\mathsf{P}[X(E) = 0] = e^{-[\lambda_1(E) + \lambda_3(E)]}, \tag{3}$$

$$\mathsf{P}[Y(E) = 0] = e^{-[\lambda_2(E) + \lambda_3(E)]} \tag{4}$$

and

$$\begin{gathered} \mathsf{P}[X(E) \geqq 2, Y(E) = 0] + \mathsf{P}[X(E) = 0, Y(E) \geqq 2] + \\ \mathsf{P}[X(E) \geqq 2, Y(E) = 1] + \mathsf{P}[X(E) \geqq 1, V(E) \geqq 2] \leqq \\ [\lambda_1(E) + \lambda_2(E) + \lambda_3(E)] \cdot \delta(\lambda_1(E), \lambda_2(E), \lambda_3(E)) \end{gathered} \tag{5}$$

where

$$\lim_{x_1, x_2, x_3 \to 0} \delta(x_1, x_2, x_3) = 0. \tag{6}$$

Let now $X(E, \omega)$, $Y(E, \omega)$, $E \in \mathscr{F}$, $\omega \in \Omega$ be nonnegative integer-valued functions satisfying condition (1).

Define

$$X^*(E) = I[X(E) \geqq 1], \quad Y^*(E) = I[Y(E) \geqq 1]$$

where $I[A]$ denotes the indicator of an event A.

1980 *Mathematics Subject Classification.* Primary 60F05; Secondary 60G50.

Key words and phrases. Bivariate Poisson processes, additive set function, characteristic function, random vector.

We begin our considerations by the

LEMMA. *If conditions* (2)—(4) *are met then, for any disjoint sets* $E_1, E_2, \dots, E_n \in \mathcal{F}$, *the random vectors* $[X^*(E_1), Y^*(E_1)]$, $[X^*(E_2), Y^*(E_2)], \dots, [X^*(E_n), Y^*(E_n)]$ *are independent.*

PROOF. Let k, l be integers such that $2 \leqq k+l \leqq n$ and let $1 \leqq i_1 < i_2 < \dots < i_k \leqq n$; $1 \leqq j_1 < j_2 < \dots < j_l \leqq n$ be sequences of positive integers satisfying the condition $\{i_1, i_2, \dots, i_k\} \cap \{j_1, j_2, \dots, j_l\} = \emptyset$.

Put

$$A_k = \sum_{r=1}^{k} E_{i_r}, \quad B_l = \sum_{s=1}^{l} E_{j_s}.$$

It is easy to check that the sets A_k, B_l are disjoint.

Under the above assumptions we have

$$\mathsf{P}\left\{\bigcap_{r=1}^{k} [X^*(E_{i_r}) = 0], \bigcap_{s=1}^{l} [Y^*(E_{j_s}) = 0]\right\} =$$

$$\mathsf{P}\left\{\sum_{r=1}^{k} X^*(E_{i_r}) = 0, \sum_{s=1}^{l} Y^*(E_{j_s}) = 0\right\} =$$

$$\mathsf{P}[X(A_k) = 0,\ Y(B_l) = 0] = e^{-[\lambda_1(A_k) + \lambda_2(B_l) + \lambda_3(A_k \cup B_l)]} =$$

$$\prod_{r=1}^{k} \mathsf{P}[X^*(E_{i_r}) = 0] \prod_{s=1}^{l} \mathsf{P}[Y^*(E_{j_s}) = 0].$$

Thus the Lemma is proved.

THEOREM 1. *If conditions* 1° *and* 2° *are met then, for any disjoint sets* $E_1, E_2, \dots, E_n$, *the random vectors* $[X(E_1), Y(E_1)]$, $[X(E_2), Y(E_2)], \dots, [X(E_n), Y(E_n)]$ *are independent.*

PROOF. We prove Theorem 1 in the case $n=2$. For an arbitrary n the proof can be done by induction.

Let E, F be disjoint sets from $\mathcal{F}$. Now we show that the random vectors $[X(E), Y(E)]$, $[X(F), Y(F)]$ are independent.

It follows from 1° that for every $\varepsilon > 0$ there exist a natural number n and two sequences of disjoint sets $\{E_i,\ i=1, 2, \dots, n\}$; $\{F_i,\ i=1, 2, \dots, n\}$ such that $E = \sum_{i=1}^{n} E_i$, $F = \sum_{i=1}^{n} F_i$. Moreover, for $i = 1, 2, \dots, n$ we have

$$\max_{1 \leqq j \leqq 3} \lambda_j(E_i) < \varepsilon, \quad \max_{1 \leqq j \leqq 3} \lambda_j(F_i) < \varepsilon. \tag{7}$$

The Lemma implies that the random vectors $[X^*(E_1), Y^*(E_1)]$, $[X^*(E_2), Y^*(E_2)], \dots, [X^*(E_n), Y^*(E_n)], [X^*(F_1), Y^*(F_1)], [X^*(F_2), Y^*(F_2)], \dots, [X^*(F_n), Y^*(F_n)]$

are independent. Then the random vectors

$$\sum_{i=1}^{n}[X^*(E_i), Y^*(E_i)], \qquad \sum_{i=1}^{n}[X^*(F_i), Y^*(F_i)]$$

are also independent.

Using (5), we obtain now

$$\mathsf{P}\left\{[X(E), Y(E)] \neq \sum_{i=1}^{n}[X^*(E_i), Y^*(E_i)]\right\} \leqq$$

$$\leqq \sum_{i=1}^{n}[\lambda_1(E_i)+\lambda_2(E_i)+\lambda_3(E_i)]\cdot\delta\big(\lambda_1(E_i), \lambda_2(E_i), \lambda_3(E_i)\big).$$

It follows from relations (6) and (7) that for every $\varepsilon>0$,

$$\max_{1\leqq i\leqq n}\delta\big(\lambda_1(E_i), \lambda_2(E_i), \lambda_3(E_i)\big) < \varepsilon.$$

Hence

$$\mathsf{P}\left\{[X(E), Y(E)] \neq \sum_{i=1}^{n}[X^*(E_i), Y^*(E_i)]\right\} \leqq [\lambda_1(E)+\lambda_2(E)+\lambda_3(E)]\varepsilon. \tag{8}$$

In the same way we obtain

$$\mathsf{P}\left\{[X(F), Y(F)] \neq \sum_{i=1}^{n}[X^*(F_i), Y^*(F_i)]\right\} \leqq [\lambda_1(E)+\lambda_2(E)+\lambda_3(E)]\varepsilon. \tag{9}$$

From the independence of the random vectors

$$\sum_{i=1}^{n}[X^*(E_i), Y^*(E_i)], \qquad \sum_{i=1}^{n}[X^*(F_i), Y^*(F_i)]$$

and from (8) and (9), we have

$$|\mathsf{P}[X(E)=k_1, Y(E)=k_2, X(F)=t_1, Y(F)=t_2]-$$
$$-\mathsf{P}[X(E)=k_1, Y(E)=k_2]\mathsf{P}[X(F)=t_1, Y(F)=t_2]| \leqq$$
$$\leqq 3[\lambda_1(E\cup F)+\lambda_2(E\cup F)+\lambda_3(E\cup F)]\varepsilon.$$

The last relation completes the proof.

THEOREM 2. *If conditions* 1° *and* 2° *are satisfied then, for every* $E\in\mathscr{F}$ *and any nonnegative integers* n, m, *we have*

$$\mathsf{P}[X(E)=n, Y(E)=m] = \sum_{k=0}^{\min(n,m)}\frac{\lambda_1^{n-k}(E)\lambda_2^{m-k}(E)\lambda_3^{k}(E)}{(n-k)!(m-k)!k!}e^{-[\lambda_1(E)+\lambda_2(E)+\lambda_3(E)]}.$$

PROOF. Let $\varphi_E(t_1, t_2)$ denote, for each $E\in\mathscr{F}$, the characteristic function of the random vector $[X(E), Y(E)]$. It is easy to see that

$$|\varphi_E(t_1, t_2)| \geqq 2e^{-[\lambda_1(E)+\lambda_2(E)+\lambda_3(E)]}-1,$$

so we have

$$|\varphi_E(t_1, t_2)| \neq 0 \quad \text{for} \quad \lambda_1(E)+\lambda_2(E)+\lambda_3(E) < \log 2.$$

Because of condition 1°, E has a disjoint decomposition $E=\sum_{k=1}^{n} E_k$ with $\lambda_1(E_i)+$ $+\lambda_2(E_i)+\lambda_3(E_i)<\log 2$; i.e., $|\varphi_{E_i}(t_1, t_2)|\neq 0$, $i=1, 2, \ldots, n$.

But, because of Theorem 1, the random vectors $[X(E_1), Y(E_1)], [X(E_2), Y(E_2)], \ldots$ $\ldots, [X(E_n), Y(E_n)]$ are independent, hence

$$\varphi_E(t_1, t_2) = \prod_{i=1}^{n} \varphi_{E_i}(t_1, t_2) \neq 0 \tag{10}$$

and we have

$$\log \varphi_E(t_1, t_2) = \sum_{i=1}^{n} \log \varphi_{E_i}(t_1, t_2). \tag{11}$$

From condition 2° we have

$$\begin{aligned} |\varphi_{E_i}(t_1, t_2) - e^{-[\lambda_1(E_i)+\lambda_2(E_i)+\lambda_3(E_i)]} - e^{it_1-\lambda_2(E_i)-\lambda_3(E_i)} \times \\ \times (1-e^{-\lambda_1(E_i)}) - e^{it_2-\lambda_1(E_i)-\lambda_3(E_i)}(1-e^{-\lambda_2(E_i)}) - \\ - e^{i(t_1+t_2)}(1-e^{-[\lambda_1(E_i)+\lambda_3(E_i)]} - e^{-[\lambda_2(E_i)+\lambda_3(E_i)]} + e^{-[\lambda_1(E_i)+\lambda_2(E_i)+\lambda_3(E_i)]})| \leqq \\ \leqq 2\{[\lambda_1(E_i)+\lambda_2(E_i)+\lambda_3(E_i)]\delta(\lambda_1(E_i), \lambda_2(E_i), \lambda_3(E_i))\}. \end{aligned}$$

Hence there exists a complex valued function $\varrho(x_1, x_2, x_3)$ of real variables for which $|\varrho(x_1, x_2, x_3)|\leqq 1$, $\varrho(0, 0, 0)=0$ and

$$\begin{aligned} \varphi_{E_i}(t_1, t_2) = e^{-[\lambda_1(E_i)+\lambda_2(E_i)+\lambda_3(E_i)]} + \\ + e^{it_1-\lambda_2(E_i)-\lambda_3(E_i)}(1-e^{-\lambda_1(E_i)}) + e^{it_2-\lambda_1(E_i)-\lambda_3(E_i)}(1-e^{-\lambda_2(E_i)}) + \\ + e^{i(t_1+t_2)}(1-e^{-[\lambda_1(E_i)+\lambda_3(E_i)]} - e^{-[\lambda_2(E_i)+\lambda_3(E_i)]} + e^{-[\lambda_1(E_i)+\lambda_2(E_i)+\lambda_3(E_i)]}) + \\ + 2[\lambda_1(E_i)+\lambda_2(E_i)+\lambda_3(E_i)]\delta(\lambda_1(E_i), \lambda_2(E_i), \lambda_3(E_i))\varrho(\lambda_1(E_i), \lambda_2(E_i), \lambda_3(E_i)). \end{aligned}$$

Expanding $\log \varphi_{E_i}(t_1, t_2)$ in Taylor series we have

$$\begin{aligned} \log \varphi_{E_i}(t_1, t_2) = \lambda_1(E_i)(e^{it_1}-1)+\lambda_2(E_i)(e^{it_2}-1)+\lambda_3(E_i)(e^{i(t_1+t_2)}-1)+ \\ + \sum_{k=1}^{3} \lambda_k(E_i)\varepsilon_k(\lambda_1(E_i), \lambda_2(E_i), \lambda_3(E_i)) \end{aligned} \tag{12}$$

where for $k = 1, 2, 3$

$$\lim_{x_1, x_2, x_3 \to 0} \varepsilon_k(x_1, x_2, x_3) = 0. \tag{13}$$

Substituting (12) into (11), we have

$$\begin{aligned} \log \varphi_E(t_1, t_2) = \lambda_1(E)(e^{it_1}-1)+\lambda_2(E)(e^{it_2}-1)+\lambda_3(E)(e^{i(t_1+t_2)}-1)+ \\ + \sum_{i=1}^{n} \sum_{k=1}^{3} [\lambda_k(E_i)\varepsilon_k(\lambda_1(E_i), \lambda_2(E_i), \lambda_3(E_i))]. \end{aligned} \tag{14}$$

It follows from (14) that

$$T = \sum_{i=1}^{n} \sum_{k=1}^{3} [\lambda_k(E_i)\varepsilon_k(\lambda_1(E_i), \lambda_2(E_i), \lambda_3(E_i))]$$

is a constant, and from (13) we have

$$|T| \leqq \sum_{k=1}^{3} \lambda_k(E) \max_{1\leqq i\leqq n} \varepsilon_k(\lambda_1(E_i), \lambda_2(E_i), \lambda_3(E_i)) \to 0$$

as $\lambda_1(E_i), \lambda_2(E_i), \lambda_3(E_i) \to 0$. Hence $T \equiv 0$.

Theorem 2 follows from the last identity and from (14).

THEOREM 3. *Let, for each* $E, F \in \mathscr{F}$,

$$\mathsf{P}[X_i(E) = 0, Y_i(F) = 0] = e^{-[\lambda_{1i}(E)+\lambda_{2i}(F)+\lambda_{3i}(E\cup F)]},$$

$$\mathsf{P}[X_i(E) = 0] = e^{-[\lambda_{1i}(E)+\lambda_{3i}(E)]},$$

(15) $$\mathsf{P}[Y_i(E) = 0] = e^{-[\lambda_{2i}(E)+\lambda_{3i}(E)]},$$

$$\mathsf{P}[X_i(E) = n, Y_i(E) = m] = \sum_{k=0}^{\min(n,m)} \frac{\lambda_{1i}^{n-k}(E)\lambda_{2i}^{m-k}(E)\lambda_{3i}^{k}(E)}{(n-k)!\,(m-k)!\,k!} e^{-[\lambda_{1i}(E)+\lambda_{2i}(E)+\lambda_{3i}(E)]}$$

$$i = 1, 2, \ldots, s; \; n, m = 0, 1$$

and let the set functions

$$\lambda_1(E) = \sum_{i=1}^{s} \lambda_{1i}(E), \quad \lambda_2(E) = \sum_{i=1}^{s} \lambda_{2i}(E), \quad \lambda_3(E) = \sum_{i=1}^{s} \lambda_{3i}(E)$$

be measures on $\mathscr{F}$ *as in* 1°. *If for each* $E, F \in \mathscr{F}$ *we have*

(16) $$\mathsf{P}\left\{\bigcap_{i=1}^{s} [X_i(E) = n_i, Y_i(F) = m_i]\right\} = \prod_{i=1}^{s} \mathsf{P}[X_i(E) = n_i, Y_i(F) = m_i]$$

for all nonnegative integers $n_1, n_2, \ldots, n_s; m_1, m_2, \ldots, m_s$ *with*

$$\sum_{i=1}^{s} n_i \leqq 1, \quad \sum_{i=1}^{s} m_i \leqq 1,$$

then

$$\mathsf{P}\left[\sum_{i=1}^{s} X_i(E) = n, \; \sum_{i=1}^{s} Y_i(E) = m\right] =$$

$$= \sum_{k=0}^{\min(n,m)} \frac{\lambda_1^{n-k}(E)\lambda_2^{m-k}(E)\lambda_3^{k}(E)}{(n-k)!\,(m-k)!\,k!} e^{-[\lambda_1(E)+\lambda_2(E)+\lambda_3(E)]}$$

for any $E \in \mathscr{F}$.

PROOF. Let $E, F \in \mathcal{F}$ be arbitrary sets, then from (16) and (15) we get

$$\mathsf{P}\left[\sum_{i=1}^{s} X_i(E) = 0,\ \sum_{i=1}^{s} Y_i(F) = 0\right] = \prod_{i=1}^{s} \mathsf{P}[X_i(E) = 0, Y_i(F) = 0] =$$
$$= e^{-[\lambda_1(E)+\lambda_2(F)+\lambda_3(E\cup F)]},$$

$$\mathsf{P}\left[\sum_{i=1}^{s} X_i(E) = 0\right] = e^{-[\lambda_1(E)+\lambda_3(E)]}, \tag{17}$$

$$\mathsf{P}\left[\sum_{i=1}^{s} Y_i(E) = 0\right] = e^{-[\lambda_2(E)+\lambda_3(E)]}.$$

Now it is easy to see that

$$\mathsf{P}\left[\sum_{i=1}^{s} X_i(E) = 1,\ \sum_{i=1}^{s} Y_i(E) = 1\right] = \sum_{l,j=1}^{s} \prod_{i=1}^{s} \mathsf{P}[X_i(E) = \delta_{il}, Y_i(E) = \delta_{ij}]$$

$$= \sum_{lj=1}^{s} \prod_{i=1}^{s} \sum_{k=0}^{\min(\delta_{il},\delta_{ij})} \frac{\lambda_{1i}^{\delta_{il}-k}(E)\lambda_{2i}^{\delta_{ij}-k}(E)\lambda_{3i}^{k}(E)}{(\delta_{il}-k)!\,(\delta_{ij}-k)!\,k!} e^{-[\lambda_{1i}(E)+\lambda_{2i}(E)+\lambda_{3i}(E)]} = \tag{18}$$
$$= [\lambda_1(E)+\lambda_2(E)+\lambda_3(E)]e^{-[\lambda_1(E)+\lambda_2(E)+\lambda_3(E)]}$$

and also

$$\mathsf{P}\left[\sum_{i=1}^{s} X_i(E) = 1,\ \sum_{i=1}^{s} Y_i(E) = 0\right] = \sum_{j=1}^{s} \prod_{i=1}^{s} \lambda_{1i}^{\delta_{ij}}(E) e^{-[\lambda_{1i}(E)+\lambda_{2i}(E)+\lambda_{3i}(E)]} = \tag{19}$$
$$= \lambda_1(E)e^{-[\lambda_1(E)+\lambda_2(E)+\lambda_3(E)]}.$$

In the same way it can be proved that

$$\mathsf{P}\left[\sum_{i=1}^{s} X_i(E) = 0,\ \sum_{i=1}^{s} Y_i(E) = 1\right] = \lambda_2(E)e^{-[\lambda_1(E)+\lambda_2(E)+\lambda_3(E)]}. \tag{20}$$

Now let us put

$$I(E) = 1 - \mathsf{P}\left[\sum_{i=1}^{s} X_i(E) = 1,\ \sum_{i=1}^{s} Y_i(E) = 1\right] -$$
$$- \mathsf{P}\left[\sum_{i=1}^{s} X_i(E) = 1,\ \sum_{i=1}^{s} Y_i(E) = 0\right] -$$
$$- \mathsf{P}\left[\sum_{i=1}^{s} X_i(E) = 0,\ \sum_{i=1}^{s} Y_i(E) = 1\right] -$$
$$- \mathsf{P}\left[\sum_{i=1}^{s} X_i(E) = 0,\ \sum_{i=1}^{s} Y_i(E) = 0\right].$$

Taking into account (17), (18), (19) and (20), we have

$$I(E) = 1 - [1-\lambda_1(E)-\lambda_2(E)-\lambda_3(E)-\lambda_1(E)\lambda_2(E)]e^{-[\lambda_1(E)+\lambda_2(E)+\lambda_3(E)]}. \tag{21}$$

Because of (17) and (21), condition 2° is met and $\lambda_1, \lambda_2, \lambda_3$ meet condition 1°. Thus we apply Theorem 2 and the proof of Theorem 3 is complete.

THEOREM 4. *Let, for any* $E, F \in \mathcal{F}$,

$$\mathsf{P}\left[\sum_{i=1}^{s} X_i(E)=0,\ \sum_{i=1}^{s} Y_i(F)=0\right] = e^{-\lambda_1(E)+\lambda_2(E)+\lambda_3(E\cup F)}, \tag{22}$$

$$\mathsf{P}\left[\sum_{i=1}^{s} X_i(E)=n,\ \sum_{i=1}^{s} Y_i(E)=m\right] = \sum_{k=0}^{\min(n,m)} \frac{\lambda_1^{n-k}(E)\lambda_2^{m-s}(E)\lambda_3^{k}(E)}{(n-k)!\,(m-k)!\,k!} e^{-[\lambda_1(E)+\lambda_2(E)+\lambda_3(E)]} \tag{23}$$

where $\lambda_1(E)$, $\lambda_2(E)$ *and* $\lambda_3(E)$ *are measures on* $\mathcal{F}$ *satisfying condition* 1°.

If, for all $E, F \in \mathcal{F}$

$$\mathsf{P}\left\{\prod_{i=1}^{s} [X_i(E)=0,\ Y_i(F)=0]\right\} = \prod_{i=1}^{s} \{\mathsf{P}[X_i(E)=0,\ Y_i(F)=0]\}, \tag{24}$$

$$\mathsf{P}\left\{\prod_{i=1}^{s} [X_i(E)=0]\right\} = \prod_{i=1}^{s} \mathsf{P}[X_i(E)=0], \tag{25}$$

$$\mathsf{P}\left\{\prod_{i=1}^{s} [Y_i(E)=0]\right\} = \prod_{i=1}^{s} \mathsf{P}[Y_i(E)=0], \tag{26}$$

and

$$\mathsf{P}[X_i(E)=0,\ Y_i(F)=0] = a_i^{F_1(\lambda_1(E),\lambda_2(E),\lambda_3(E\cup F))}, \tag{27}$$

$$\mathsf{P}[X_i(E)=0] = a_i^{F_2(\lambda_1(E),\lambda_3(E))}, \tag{28}$$

$$\mathsf{P}[Y_i(E)=0] = a_i^{F_3(\lambda_2(E),\lambda_3(E))}, \quad i=1,2,\dots,s, \tag{29}$$

the $a_i > 1$ *being arbitrary numbers, and* F_1, F_2, F_3 *arbitrary real-valued functions, then*

$$\mathsf{P}[X_i(E)=n,\ Y_i(E)=m] = \sum_{k=0}^{\min(n,m)} \frac{\mu_{1i}^{n-k}(E)\mu_{2i}^{m-k}(E)\mu_{3i}^{k}(E)}{(n-k)!\,(m-k)!\,k!} e^{-[\mu_{1i}(E)+\mu_{2i}(E)+\mu_{3i}(E)]}$$

$$i=1,2,\dots,s;\ n, m \geqq 0$$

for

$$\mu_{ki}(E) = \lambda_{ki}(E)\frac{\log a_i}{\sum_{i=1}^{s}\log a_i}, \qquad k=1,2,3.$$

PROOF. Because of (22) and (24)

$$e^{-[\lambda_1(E)+\lambda_2(F)+\lambda_3(E\cup F)]} = \mathsf{P}\left[\sum_{i=1}^{s} X_i(E)=0,\ \sum_{i=1}^{s} Y_i(F)=0\right] =$$

$$= \mathsf{P}\left\{\prod_{i=1}^{s} [X_i(E)=0,\ Y_i(F)=0]\right\} = \prod_{i=1}^{s} \mathsf{P}[X_i(E)=0,\ Y_i(F)=0].$$

From (27) we get

$$e^{-[\lambda_1(E)+\lambda_2(F)+\lambda_3(E\cup F)]} = \prod_{i=1}^{n} a_i^{F_1(\lambda_1(E),\lambda_2(F),\lambda_3(E\cup F))}.$$

Hence

$$F_1(\lambda_1(E), \lambda_2(F), \lambda_3(E\cup F)) = -[\lambda_1(E)+\lambda_2(F)+\lambda_3(E\cup F)] \Big/ \sum_{i=1}^{s} \log a_i$$

and

$$\mathrm{P}[X_i(E)=0, Y_i(F)=0] = e^{-[\lambda_1(E)+\lambda_2(F)+\lambda_3(E\cup F)]\log a_i \big/ \sum_{i=1}^{s} \log a_i}.$$

In the same way it can be proved that

$$\mathrm{P}[X_i(E)=0] = e^{-[\lambda_1(E)+\lambda_3(E)]\log a_i \big/ \sum_{i=1}^{s} \log a_i}$$

and

$$\mathrm{P}[Y_i(E)=0] = e^{-[\lambda_2(E)+\lambda_3(E)]\log a_i \big/ \sum_{i=1}^{s} \log a_i}, \qquad i=1,2,\dots,s.$$

Observe that (2)—(4) in condition 2°, and also (5) are satisfied. Indeed, let us define

$$A = \{(n,m)\colon n, m\in \mathbf{N}\times\mathbf{N},\ n\geqq 2,\ m=0\cup n=0,\ m\geqq 2\cup n\geqq 2,$$
$$m=1\cup n=1,\ m\geqq 2\cup n\geqq 2,\ m\geqq 2\}.$$

Then we have, for $i=1,2,\dots,s$,

$$\mathrm{P}\{[X_i(E), Y_i(E)]\in A\} \leqq \mathrm{P}\left\{\left[\sum_{i=1}^{s} X_i(E),\ \sum_{i=1}^{s} Y_i E\right]\in A\right\}.$$

Condition 1° is also met for all $i=1,2,\dots,s$ because if $\lambda_1, \lambda_2, \lambda_3$ satisfy condition 1°, then for any arbitrary $c>0$, $c\lambda_i(E)$, $i=1,2,3$ also satisfy condition 1°.

Since both conditions of Theorem 2 hold, the proof of Theorem 4 is complete.

REFERENCES

[1] Győrfi, L., Poisson processes defined on an abstract space, *Studia Sci. Math. Hungar.* **7** (1972), 243—248. *MR* **48** # 7365.

[2] Kawamura, K., The structure of bivariate Poisson distribution, *Kōdai Math. Sem. Rep.* **25** (1973), 246—256. *MR* **48** # 5143.

[3] Rényi, A., Remarks on the Poisson process, *Studia Sci. Math. Hungar.* **2** (1967), 119—123. *MR* **35** # 3726.

(Received November 2, 1981)

INSTYTUT MATEMATYKI
UNIWERSYTET MARII CURIE-SKŁODOWSKIEJ
PL. MARII CURIE-SKŁODOWSKIEJ 1
PL—20-031 LUBLIN
POLAND

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 417—424

ON AN IMBALANCE PROBLEM OF G. WAGNER CONCERNING POLYNOMIALS

JÓZSEF BECK

Answering a question of P. Erdős [2], G. Wagner [3] proved the following beautiful theorem: Given an arbitrary infinite sequence $\omega=(\xi_1, \xi_2, \ldots)$ of points on the unit circle $\{z\colon |z|=1\}$, the sequence of polynomials $p_n(z)=p_n(z;\omega)=\prod_{j=1}^{n}(z-\xi_j)$ *does not* remain uniformly bounded on the unit circle. This result is in the same spirit as the classical Van Aardenne—Ehrenfest theorem (see [1]) on irregularities of distribution of sequences.

In connection with the extension of the problem above for arbitrary compact subsets of the complex plane, Wagner raised an elementary problem as follows (see [4]).

On the unit circle $\{z\colon |z|=1\}$ consider polynomials of degree $(n-1)$ having as their (simple or multiple) zeros n-th roots of unity only. That is, consider the polynomials $p(z)$ of the form

$$(1)\qquad \prod_{j=0}^{n-1}(z-\zeta_j)^{a_j}, \quad \text{where} \quad \zeta_j=e^{2\pi i\, j/n}$$

and the sum of multiplicities a_j equals $(n-1)$. Does there exist a function $f(n)$ with $f(n)\to\infty$ as $n\to\infty$ such that, for any $p(z)$ of type (1), $\max_{|z|=1}|p(z)|>f(n)$?

Observe that for polynomials $p(z)$ of type (1) having *simple* zeros only, $\max_{|z|=1}|p(z)|=n$. Wagner suspected that there is a universal $\delta>0$ such that for any polynomial $p(z)$ of type (1), $\max_{|z|=1}|p(z)|>n^{\delta}$.

Erdős disagreed, he believed that there exists a polynomial $p(z)$ of type (1) such that $\max_{|z|=1}|p(z)|<C$, where C is *independent* of n. Our objective is to prove the validity of Erdős' belief.

THEOREM. *There exists a polynomial $p(z)$ of type* (1) *such that* $\max_{|z|=1}|p(z)|\ll 1$.

The well-known Vinogradov notation $f(n)\ll g(n)$ means that $f(n)=O(g(n))$, i.e., $|f(n)/g(n)|$ remains bounded as n tends to infinity.

PROOF. We start with the definition of the desired polynomial $p(z)$, that is, we shall associate multiplicities a_j, $0\leqq a_j\leqq 2$ with the n-th roots of unity $\zeta_j=e^{2\pi i\cdot j/n}$,

---

1980 *Mathematics Subject Classifications.* Primary 10K30; Secondary 10F40.
Key words and phrases. Irregularities of point distributions, polynomials with multiple zeros, maximum norm.

$0 \leqq j \leqq n-1$ such that

$$\sum_{j=0}^{n-1} a_j = n-1 \quad \text{and} \quad \max_{|z|=1} \prod_{j=0}^{n-1} |z-\zeta_j|^{a_j} \ll 1.$$

We shall denote by l_j, $0 \leqq j \leqq s$ and k_j, $0 \leqq j \leqq s-1$ the indices for which ζ_{l_j} and ζ_{k_j} have multiplicities 0 and 2, respectively. The sequences $\{l_j\}$ and $\{k_j\}$ will have the following basic properties:

$$0 = l_0 < k_0 < l_1 < k_1 < \ldots < l_{s-1} < k_{s-1} < l_s = n-1$$

and

$$\left| \int_0^x \frac{y}{n} dy - \sum_{i: k_i \leqq x} (k_i - l_i) - (x - l_{i_0}) \right| \ll 1 \tag{2}$$

uniformly for all $0 \leqq x \leqq n-1$. Here i_0 denotes the index for which $l_{i_0} \leqq x < k_{i_0}$.

Let $l_0 = 0$. Let $l_1 > 0$ be the first integer such that $\int_0^{l_1} (y/n)\, dy \geqq 1$. Assume now that $l_1 < \ldots < l_{i-1} < n/4$ are already defined, then let $l_i > l_{i-1}$ be the first integer such that

$$\int_0^{l_i} (y/n)\, dy \geqq i.$$

Let l_q denote the first index $\geqq n/4$. Then let

$$l_{q+i} = l_q + 4i, \quad i = 1, 2, \ldots \quad \text{if only} \quad l_{q+i} \leqq n/2.$$

Let l_{r-1} denote the last index $\leqq n/2$, then let $l_r = l_{r-1} + 4$. Let us extend l_{r+i} (almost) *symmetrically:*

$$l_{r+i} = l_r + (l_r - l_{r-i}) \quad \text{if only} \quad l_{r+i} < n-1.$$

Let l_{s-1} denote the last index $< n-1$. Finally, let $l_s = n-1$.

Now we define the sequence k_j, $j = 0, 1, \ldots, s-1$. Let $k_j = l_j + 1$ if only $j \leqq q$. For $q < j \leqq r-1$ let $k_j - l_j$ be the integral part of the expression

$$\int_0^{l_{j+1}} \frac{y}{n} dy - \sum_{i=0}^{j-1} (k_i - l_i).$$

For $r \leqq j \leqq s-1$ we extend k_j as follows: let

$$l_{j+1} - k_j = k_{2r-j-1} - l_{2r-j-1}.$$

We mention some further consequences of the definitions above:

$$l_1 \gg\ll n^{1/2} \quad \text{and} \quad l_{i+1} - l_i \gg\ll n/l_i \quad \text{for} \quad 1 \leqq i \leqq r; \tag{3a}$$

(the notation $f(n) \gg\ll g(n)$ means that both $f(n) \gg g(n)$ and $f(n) \ll g(n)$)

$$n - l_{s-1} \gg\ll n^{1/2} \quad \text{and} \quad l_{s-i} - l_{s-i-1} = l_{i+1} - l_i \quad \text{for} \quad 1 \leqq i \leqq r; \tag{3b}$$

$$1 \leqq k_i - l_i \leqq 2 \quad \text{if} \quad 0 \leqq i \leqq r \quad \text{and} \quad 1 \leqq l_{i+1} - k_i \leqq 2 \quad \text{if} \quad r \leqq i \leqq s-1. \tag{4}$$

Now we are ready to define the desired $p(z)$: let $p(z)=\prod_{j=0}^{n-1}(z-\zeta_j)^{a_j}$, where

$$a_j=\begin{cases}0 & \text{if} \quad j\in\{l_0, l_1, \dots, l_s\}\\ 2 & \text{if} \quad j\in\{k_0, k_1, \dots, k_{s-1}\}\\ 1 & \quad\text{otherwise.}\end{cases}$$

In what follows we shall show that the polynomial $p(z)$ of degree $n-1$ is bounded in $|z|=1$. We shall verify this only in the arc $\{e^{2\pi it/n}: -1/2\leqq t\leqq l_r\}$ (we recall that $l_r\sim n/2$). The case of the complementing arc goes similarly using the fact the rootsystem of $p(z)$ is almost symmetrical.

Let $z_0=e^{2\pi i t/n}$, where

$$l_u-\frac{1}{2}\leqq t<l_{u+1}-\frac{1}{2}, \quad 0\leqq u\leqq r.$$

We have

$$p(z_0)=P_1P_2P_3,$$

where

$$P_1=(z_0-\zeta_{k_u})\prod_{\substack{h=0\\ h\neq l_u, l_{u+1}}}^{n-1}(z_0-\zeta_h), \quad P_2=\prod_{i=u+1}^{s-1}\left(\frac{z_0-\zeta_{k_i}}{z_0-\zeta_{l_{i+1}}}\right)$$

and

$$P_3=\prod_{i=0}^{u-1}\left(\frac{z_0-\zeta_{k_i}}{z_0-\zeta_{l_i}}\right).$$

First we show

$$|P_1|\ll\frac{n}{l_{u+1}-t}. \tag{5}$$

If $t\leqq k_u+1$ then $|z_0-\zeta_{k_u}|\ll 1/n$. Therefore, using the facts $|z_0-\zeta_{l_{u+1}}|\gg(l_{u+1}-t)/n$ and $\left|\frac{z_0^n-1}{z_0-\zeta_{l_u}}\right|\leqq n$, we obtain

$$|P_1|=\left|\frac{z_0^n-1}{z_0-\zeta_{l_u}}\right|\left|\frac{z_0-\zeta_{k_u}}{z_0-\zeta_{l_{u+1}}}\right|\ll\frac{n}{l_{u+1}-t}.$$

If $t>k_u+1$ then $|(z_0-\zeta_{k_u})/(z_0-\zeta_{l_u})|\ll 1$, therefore using again the inequality $|z_0-\zeta_{l_{u+1}}|\gg(l_{u+1}-t)/n$ we get

$$|P_1|=|z_0^n-1|\frac{1}{|z_0-\zeta_{l_{u+1}}|}\left|\frac{z_0-\zeta_{k_u}}{z_0-\zeta_{l_u}}\right|\ll\frac{n}{l_{u+1}-t},$$

which completes (5).

Let

$$P_2'=\prod_i{}^{*}\left(\frac{z_0-\zeta_{k_i}}{z_0-\zeta_{l_{i+1}}}\right), \quad P_2''=\prod_i{}^{**}\left(\frac{z_0-\zeta_{k_i}}{z_0-\zeta_{l_{i+1}}}\right)$$

where the products $\prod^{*}$ and $\prod^{**}$ are extended over all $i\geqq u+1$ for which

$$\left|\frac{z_0-\zeta_{k_i}}{z_0-\zeta_{l_{i+1}}}\right| \quad \text{is} \quad <1 \quad \text{or} \quad \geqq 1, \quad \text{respectively.}$$

Clearly, $P_2=P_2'P_2''$. We claim

(6) $$|P_2''|\ll 1.$$

Simple calculation shows that for the factors of the product P_2'' we have (see (4)):

(7) $$\left|\frac{z_0-\zeta_{k_i}}{z_0-\zeta_{l_{i+1}}}\right|\leqq\left|\frac{1-\zeta_{l_{(i+1)}-2}}{1-\zeta_{l_{i+1}}}\right|\leqq 1+\frac{c}{n-l_{i+1}}$$

with a suitable absolute constant $c>0$. Let $m_i=n-l_{u-i}$, $i=0,1,2,\ldots$. By (3a) and (3b), $m_0=1$, $m_1\gg n^{1/2}$, $m_{i+1}-m_i\gg n/m_i$, $i=1,2,\ldots$. Now we need a general lemma.

LEMMA 1. *Let a sequence* $1=m_0<m_1<\ldots<m_v$ *of integers be given such that* $m_1\gg n^{1/2}$, $m_{i+1}-m_i\gg n/m_i$, $i=1,\ldots,v-1$. *Then* $\sum_{i=0}^{v}\frac{1}{m_i}\ll 1$.

PROOF. The hypothesis of the lemma yields that the interval $[x,2x)$ contains $\ll x^2/n$ elements of the sequence $m_0,m_1,\ldots,m_v$. Let $\lfloor\log_2 n\rfloor$ denote the integral part of the binary logarithm of n. Then

$$\sum_{i=0}^{v}\frac{1}{m_i}=\sum_{j=0}^{\lfloor\log_2 n\rfloor}\sum_{2^j\leqq m_i<2^{j+1}}\frac{1}{m_i}\ll\sum_{j=0}^{\lfloor\log_2 n\rfloor}\frac{4^j}{n}\frac{1}{2^j}=$$

$$=\sum_{j=0}^{\lfloor\log_2 n\rfloor}\frac{2^j}{n}\leqq 1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\ldots=2.\quad\square$$

By (7) and Lemma 1

$$|P_2''|\leqq\prod_i{}^{**}\left(1+\frac{c}{n-l_{i+1}}\right)\leqq\prod_{j=0}^{r}\left(1+\frac{c}{m_j}\right)\leqq$$

$$\leqq\exp\left\{c\sum_{j=0}^{r}\frac{1}{m_j}\right\}=e^{O(1)}\ll 1,$$

which proves (6).

In order to finish the proof it suffices to verify the following estimates:

(8) $$|P_2'|\ll\left(\frac{l_{u+1}-t}{n}\right)^{1-\frac{k_{u+1}}{n}};$$

(9) $$|P_3|\ll\min\left\{1,\left(\frac{t-k_{u-1}}{t}\right)^{-\frac{k_{u-1}}{n}}\right\}.$$

Indeed, if $t\in\left[l_0-\frac{1}{2}=-\frac{1}{2},\ l_1-\frac{1}{2}\right)$ then $P_3=1$ (empty product). Thus, by (3), (5),

(6) and (8)

$$|p(z_0)| = |P_1||P_2'||P_2''||P_3| \ll \left(\frac{n}{l_{u+1}-t}\right)\left(\frac{l_{u+1}-t}{n}\right)^{1-\frac{k_{u+1}}{n}} =$$

$$= \left(\frac{n}{l_{u+1}-t}\right)^{\frac{k_{u+1}}{n}} < n^{O(1)/\sqrt{n}} \ll 1.$$

If $l_r \geqq t \geqq l_1 - \frac{1}{2}$, then again by (3), (5), (6), (8) and (9) we have

$$|p(z_0)| = |P_1||P_2'||P_2''||P_3| \ll$$

$$\ll \left(\frac{n}{l_{u+1}-t}\right)\left(\frac{l_{u+1}-t}{n}\right)^{1-\frac{k_{u+1}}{n}}\left(\frac{t-k_{u-1}}{t}\right)^{\frac{k_{u-1}}{n}} =$$

$$= \left(\frac{n}{l_{u+1}-t}\right)^{\frac{k_{u+1}}{n}}\left(\frac{t-k_{u-1}}{t}\right)^{\frac{k_{u-1}}{n}} \ll \left\{\frac{n(t-k_{u-1})}{(l_{u+1}-t)t}\right\}^{\frac{k_{u-1}}{n}} n^{\frac{k_{u+1}-k_{u-1}}{n}} \ll$$

$$\ll \left(\frac{n(n/t)}{t}\right)^{t/n} n^{O(1)/\sqrt{n}} \ll \left(\frac{n}{t}\right)^{\frac{2t}{n}} \ll 1,$$

which completes the proof of the theorem.

Both (8) and (9) are corollaries of the following general lemma.

LEMMA 2. *Let be given the positive real numbers* n, $0<\alpha<1$, $b_1>b_2>\ldots>b_m$ *and coefficients* $\varepsilon_j=1$ *or* 0, $1\leqq j\leqq m$ *such that* $m\leqq\alpha\cdot n$,

(10) $$b_j \ll \frac{1}{j}$$

uniformly for all $1\leqq j\leqq m$; *and*

(11) $$\left|\sum_{j=0}^{x}\varepsilon_j - \int_0^x\left(\alpha-\frac{y}{n}\right)dy\right| \ll 1$$

uniformly for all $0\leqq x\leqq m$. *Then*

$$\left|\sum_{j=1}^{m}\varepsilon_j b_j - \alpha\sum_{j=1}^{m} b_j\right| \ll 1.$$

PROOF. Clearly,

$$\left|\sum_{j=1}^{m}\varepsilon_j b_j - \alpha\sum_{j=1}^{m} b_j\right| \leqq \left|\sum_{j=1}^{m}\varepsilon_j b_j - \sum_{j=1}^{m}\left(\int_{j-1}^{j}\left(\alpha-\frac{y}{n}\right)dy\right)b_j\right| +$$

$$+\left|\sum_{j=1}^{m}\left(\int_{j-1}^{j}\left(\alpha-\frac{y}{n}\right)dy\right)b_j - \sum_{j=1}^{m}\alpha b_j\right| = S_1+S_2.$$

By (10), (11) and Abel's summation method, $S_1 \ll 1$. On the other hand, by (10)

$$S_2 = \sum_{j=1}^{m} \left(\int_{j-1}^{j} \frac{y}{n} dy \right) b_j \ll \sum_{j=1}^{m} \frac{j}{n} \frac{1}{j} \leqq 1. \quad \square$$

Now we deduce (8) from Lemma 2. For notational convenience let

$$D_j = \left| \frac{z_0 - \zeta_j}{z_0 - \zeta_{j+1}} \right|, \quad J = \left[k_{u+1}, \frac{n}{3} + k_{u+1} \right],$$

$$A = \bigcup_{i=0}^{s-1} [k_i, l_{i+1}), \quad \delta_i = \begin{cases} 1 & \text{if } i \in A \\ 0 & \text{if } i \in [0, n-1] \setminus A. \end{cases}$$

From the definition of P_2' it follows that

$$|P_2'| \leqq \prod_{i: k_i \in J} \left| \frac{z_0 - \zeta_{k_i}}{z_0 - \zeta_{l_{i+1}}} \right|. \tag{12}$$

The right-hand side of (12) can be written in the form $\prod_{j \in J} D_j^{\delta_j}$, therefore

$$|P_2'| \leqq \prod_{j \in J} D_j^{\delta_j} = \exp \left\{ - \sum_{j \in J} \delta_j \log \frac{1}{D_j} \right\}. \tag{13}$$

We recall that $z_0 = e^{2\pi i t/n}$. An easy calculation shows that for all $j \in J$

$$\frac{1}{D_j} = \left| \frac{z_0 - \zeta_{j+1}}{z_0 - \zeta_j} \right| \leqq 1 + \frac{c}{j-t}$$

with a suitable universal constant $C > 0$. Thus

$$\log \frac{1}{D_j} \ll \frac{1}{j-t} \tag{14}$$

uniformly for all $j \in J$. Moreover, observe that for $j \in J$

$$\frac{1}{D_j} > \frac{1}{D_{j+1}}. \tag{15}$$

Let us denote $K = k_{u+1}$. We claim

$$\left| \sum_{j=K}^{K+x} \delta_j - \int_0^x \frac{n-K-y}{n} dy \right| \ll 1 \tag{16}$$

uniformly for $0 \leqq x \leqq n/3$.

In order to prove (16) let $\bar{\delta}_j = 1 - \delta_j$. By definition

$$\sum_{j=0}^{W} \bar{\delta}_j = \sum_{j: k_j \leqq W} (k_j - l_j) + (W - l_{j_0}),$$

where j_0 denotes the index for which $l_{j_0} \leqq W < k_{j_0}$. Thus, by (2)

$$\left|\sum_{j=0}^{W} \delta_j - \int_0^W (y/n)\,dy\right| \ll 1 \tag{17}$$

uniformly for $0 \leqq W \leqq n-1$. Using the identity $\sum_{j=K}^{K+x} \delta_j = (x+1) - \sum_{j=K}^{K+x} \bar{\delta}_j$ and (17) we have

$$\left|\sum_{j=K}^{K+x} \delta_j - \int_0^x \frac{n-K-y}{n}\,dy\right| = \left|(x+1) - \sum_{j=K}^{K+x} \bar{\delta}_j - \int_0^x \frac{n-K-y}{n}\,dy\right| =$$

$$= O(1) + \left|(x+1) - \int_K^{K+x} (y/n)\,dy - \int_0^x \frac{n-K-y}{n}\,dy\right| = O(1) + 0 \ll 1,$$

completing the verification of (16).

By (14), (15), (16) and Lemma 2 we obtain

$$\left|\sum_{j \in J} \delta_j \log \frac{1}{D_j} - \frac{n-K}{n} \sum_{j \in J} \log \frac{1}{D_j}\right| \ll 1. \tag{18}$$

By (13) and (18)

$$|P_2'| \leqq \exp\left\{-\sum_{j \in J} \delta_j \log \frac{1}{D_j}\right\} = \exp\left\{-\frac{n-K}{n} \sum_{j \in J} \log \frac{1}{D_j} + O(1)\right\} \ll$$

$$\ll \exp\left\{-\frac{n-K}{n} \sum_{j \in J} \log \frac{1}{D_j}\right\} = \Big(\prod_{j \in J} D_j\Big)^{\frac{n-K}{n}}.$$

Clearly,

$$\prod_{j \in J} D_j = \prod_{j \in \left[K, K+\frac{n}{3}\right)} \left|\frac{z_0 - \zeta_j}{z_0 - \zeta_{j+1}}\right| = \left|\frac{z_0 - \zeta_K}{z_0 - \zeta_{K+\frac{n}{3}}}\right| \ll \frac{K-t}{n} \ll \frac{l_{u+1}-t}{n},$$

hence

$$|P_2'| \ll \left(\frac{l_{u+1}-t}{n}\right)^{\frac{n-K}{n}}$$

which proves (8).

The proof of (9) goes similarly. If $t \in \left[l_0 - \frac{1}{2} = -\frac{1}{2},\ l_1 - \frac{1}{2}\right)$ then $P_3 = 1$ (empty product). Now assume that $t \geqq l_1 - \frac{1}{2}$. Let $K^* = k_{u-1}$ and $J^* = [0, K^* - 1]$. We have

$$|P_3| = \prod_{i=0}^{u-1} \left|\frac{z_0 - \zeta_{k_i}}{z_0 - \zeta_{l_i}}\right| = \prod_{j \in J^*} D_j^{\delta_j - 1} = \prod_{j \in J^*} D_j^{-\bar{\delta}_j} =$$

$$= \exp\left\{-\sum_{j \in J^*} \bar{\delta}_j \log D_j\right\}. \tag{19}$$

For all $j\in J^*$, $1<D_j<D_{j+1}$ and $\log D_j \ll \frac{1}{t-j}$. Furthermore, by (2)

$$\left|\sum_{j=0}^{x} \bar{\delta}_j - \int_0^x (y/n)\,dy\right| \ll 1 \quad \text{uniformly for all } x\in J^*.$$

Applying Lemma 2 we obtain

$$\left|\sum_{j\in J^*} \bar{\delta}_j \log D_j - \frac{K^*}{n} \sum_{j\in J^*} \log D_j\right| \ll 1.$$

Returning to (19)

$$|P_3| = \exp\Big\{-\sum_{j\in J^*} \bar{\delta}_j \log D_j\Big\} \ll \exp\Big\{-\frac{K^*}{n} \sum_{j\in J^*} \log D_j\Big\} =$$

$$= \Big(\prod_{j\in J^*} D_j\Big)^{-\frac{K^*}{n}} = \left|\frac{z_0-\zeta_{K^*}}{z_0-\zeta_{l_0}}\right|^{\frac{K^*}{n}} = \left|\frac{z_0-\zeta_{K^*}}{z_0-1}\right|^{\frac{K^*}{n}} \ll \left(\frac{t-K^*}{t}\right)^{\frac{K^*}{n}}.$$

This completes the proof of (9) and completes the proof of the theorem. □

REFERENCES

[1] VAN AARDENNE—EHRENFEST, T., Proof of the impossibility of a just distribution of an infinite sequence of points over an interval, *Nederl. Akad. Wetensch., Proc.* **48** (1945), 266—271 = *Indagationes Math.* 7 (1945), 71—76. *MR* 7—376.

[2] ERDŐS, P., Problems and results on diophantine approximations, *Compositio Math.* **16** (1964), 52—65. *MR* **31** #3382.

[3] WAGNER, G., On a problem of Erdős in Diophantine approximation, *Bull. London Math. Soc.* **12** (1980), 81—88. *MR* **81f**: 10048.

[4] WAGNER, G., Irregularities of charge distributions, Colloq. Math. Soc. J. Bolyai, North-Holland Publishing Co., Amsterdam.

(Received November 11, 1981)

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

It is natural to conjecture, that in order to ensure the validity of (1), θ_n should be chosen as $\theta_n \leqq 1/\log n$. Surprisingly enough however, the condition $\theta_n \to 0$ suffices.

THEOREM 2. *Let $\xi_1, \xi_2, \ldots$ be as in Theorem 1, and $\zeta_{1,n}, \ldots, \zeta_{n,n}$ an independent 0—1 sequence with $\mathsf{P}(\zeta_{i,n}=1)=\theta_n$. Let $k=\sum_{i=1}^{n} \zeta_{i,n}$ and i_t the index of the t-th 1 in $\zeta_{1,n}, \ldots, \zeta_{n,n}$ $t=1, \ldots, k$, $i_0=0$. Let $v_t(n)$ be the length of the longest run of 1's within $\xi_{i_{t-1}+1}, \ldots, \xi_{i_t}$ and $v(n)=\max v_t(n)$. Suppose that $\theta_n \to 0$, then*

$$\mathsf{P}\left(\lim \frac{v(n)}{\log n} = 1\right) = 1.$$

PROOF of Theorem 1. We start with quoting the result of Erdős—Rény:i

Let $\tilde{v}(n)$ be the length of the longest run of 1's in $\xi_1, \ldots, \xi_n$. Then

$$\mathsf{P}\left(\lim \frac{\tilde{v}(n)}{\log n} = 1\right) = 1.$$

Because of this result, it suffices to show

$$\mathsf{P}\left(\liminf \frac{v(n)}{\log n} \geqq 1\right) = 1.$$

For arbitrary $\varepsilon > 0$ we introduce the following notations:

$m := [(1-\varepsilon)\log n]$, $[x]$ denotes the largest integer not greater than x

$$v_t(n) := \max_l \{\exists_j\colon\ \xi_{j+1} = \ldots = \xi_{j+l} = 1;\ (t-1)k_n \leqq j < tk_n\},$$

$$A_{t,j} := \{\xi_{(t-1)k_n+(j-1)m+1} = \ldots = \xi_{(t-1)k_n+jm} = 1\} \qquad t = 1, \ldots, \left[\frac{n}{k_n}\right].$$

With this notations we get immediately an upper bound for $\mathsf{P}(v(n) < m)$, especially

$$\mathsf{P}(v(n) < m) \leqq \mathsf{P}\left(\bigcap_{t=1}^{\left[\frac{n}{k_n}\right]} (v_t(n) < m)\right) \leqq \mathsf{P}\left(\bigcap_{t=1}^{\left[\frac{n}{k_n}\right]} \bigcap_{j=1}^{\left[\frac{k_n}{m}\right]} A_{t,j}^c\right) \leqq$$

$$\leqq (1-2^{-m})^{\left[\frac{n}{k_n}\right]\left[\frac{k_n}{m}\right]} \leqq (1-n^{-(1-\varepsilon)})^{\max\left(\left[\frac{n}{k_n}\right], \left[\frac{k_n}{m}\right]\right)}.$$

Therefore we get:

$$\mathsf{P}(v(n) < m) \leqq (1-n^{-(1-\varepsilon)})^{n^{1-\frac{\varepsilon}{2}}} \quad \text{for} \quad k_n < n^{\frac{\varepsilon}{2}}$$

$$\mathsf{P}(v(n) < m) \leqq (1-n^{-(1-\varepsilon)})^{\frac{n^{1-\frac{\varepsilon}{2}}}{m}} \quad \text{for} \quad k_n > n^{1-\frac{\varepsilon}{2}}$$

$$\mathsf{P}(v(n) < m) \leqq (1-n^{-(1-\varepsilon)})^{\left(\frac{n}{k_n}-1\right)\left(\frac{k_n}{m}-1\right)} \quad \text{for} \quad n^{\frac{\varepsilon}{2}} \leqq k_n \leqq n^{1-\frac{\varepsilon}{2}} \quad \text{and}$$

n large enough, So we get

$$\sum_{n\in N} \mathsf{P}(\nu(n) < m) \leqq C \sum_{N} e^{-n^{\frac{\varepsilon}{2}}} < \infty$$

with some positive constant C. An application of the Borel—Cantelli-lemma completes the proof. ∎

Proof of Theorem 2. We will independently modify the values of the orginal variables ξ_i such that we could look for the longest run in the whole new sequence. This modification consists of defining the last variable of all blocks to be zero:

$$\eta_i := \begin{cases} 0 & \text{if } \ i = i_t \ \text{ for some } \ t \\ \xi_i & \text{otherwise.} \end{cases}$$

Then the sequence $\eta_1, \ldots, \eta_n$ is i.i.d. with probability $\mathsf{P}(\eta_i=1)=\frac{1-\theta_n}{2}$, and $\nu(n)$ is the longest run of 1's in $\eta_1, \ldots, \eta_n$. Since $\mathsf{P}(\eta_i=1)<\frac{1}{2}$, we can a fortiori use Erdős—Rényi's theorem in the same way as before. Therefore we need to prove

$$\mathsf{P}\left(\liminf \frac{\nu(n)}{\log n} \geqq 1\right) = 1.$$

Let $m=[(1-\varepsilon)\log n]$ and denote the event, that at least one of the random variables $\eta_i, \ldots, \eta_{i+m-1}$ is zero by B_i. Then

$$\mathsf{P}(\nu(n) < m) \leqq \mathsf{P}\left(B_1 B_{1+m} \ldots B_{1+m\left[\frac{n-m}{m}\right]}\right).$$

Since B_i, B_j are independent for $|i-j|\geqq m$, we get

$$\mathsf{P}(\nu(n) < m) \leqq \mathsf{P}(B_1)^{\left[\frac{n-m}{m}\right]} \leqq \left(1-\frac{1-\theta_n}{2}\right)^{-1} \exp\left(-\frac{n(1-\theta_n)}{m2^m}\right).$$

Therefore $\lim \theta_n=0$ yields

$$\mathsf{P}(\nu(n) < m) \leqq C\exp\left(-n^{\frac{\varepsilon}{2}}\right)$$

for some positive constant C and n large enough. As before this inequality and the Borel—Cantelli lemma yields

$$\mathsf{P}\left(\liminf \frac{\nu(n)}{\log n} \geqq 1\right) = 1. \quad ∎$$

Remark 3. It follows from the results of Komlós—Tusnády [2], that $\lim \theta_n=0$ is a necessary condition. Indeed, for the case $\limsup \theta_n \geqq h>0$ one can easily deduce from their results, that

$$\mathsf{P}\left(\liminf \frac{\nu(n)}{\log n} \leqq \frac{\log 2}{\log 2-\log(1-h)} < 1\right) = 1.$$

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 429—444

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

U. BOLLE

1

In dieser Arbeit sollen die in [1] angegebenen Dichteabschätzungen für die noch fehlenden Dimensionen bewiesen werden. Ich benutze dabei die Bezeichnungsweise der genannten Arbeit. Insbesondere bezeichne $d_k^{(m)}$ bzw. $D_k^{(m)}$ die optimale Dichte einer gitterförmigen k-Packung bzw. k-Überdeckung mit Einheitskugeln im $\mathbf{R}^m$. Insgesamt wird sich damit ergeben:

SATZ. *Sei* $m\in\mathbf{N}$, $m\geqq 2$. *Dann gelten*
für $m\not\equiv 1(4)$

$$\frac{d_k^{(m)}}{k}\leqq 1-c_m k^{-\frac{m+1}{2m}},\qquad \frac{D_k^{(m)}}{k}\geqq 1+C_m k^{-\frac{m+1}{2m}}; \tag{1}$$

für $m\equiv 1(4)$

$$\frac{d_k^{(m)}}{k}\leqq 1-c_m k^{-\frac{m+3}{2m}},\qquad \frac{D_k^{(m)}}{k}\geqq 1+C_m k^{-\frac{m+3}{2m}} \tag{2}$$

($c_m, C_m>0$ *nur von* m *abhängig*).

[1] enthält den Beweis von (1) für $m\equiv 3(4)$, von (2) für Packungen und $m\equiv 1(8)$ bzw. Überdeckungen und $m\equiv 5(8)$.

2

Um die Ergebnisse aus [1] benutzen zu können, bleiben einige Aussagen nachzutragen. Zunächst ist zu zeigen, daß die Voraussetzung von Lemma 3 [1] für gerade m erfüllt ist.

LEMMA 1. *Seien* $m, M, n\in\mathbf{N}$; $m=2M$; $\frac{2}{n+1}\leqq h\leqq\frac{2}{n}$. *Dann ist* $S=S(y,h)$ *(s. die Formel unten) als Funktion von* y *für kein* h *konstant.*

BEWEIS. Ich beweise die Behauptung nur für gerade $n=2N$, der andere Fall verläuft ganz entsprechend. Da

$$\frac{2}{n+1}\leqq h\leqq\frac{2}{n},\quad \text{ist}\quad h=\frac{2}{n+2x}=\frac{1}{N+x}$$

1980 *Mathematics Subject Classification.* Primary 10E30; Secondary 05B40.
Key words and phrases. Multiple packing, multiple covering, lattice packing of spheres, lattice covering with spheres.

mit einem geeigneten $x \in \left[0; \frac{1}{2}\right]$.

$$S(y,h) = V_{m-1} \sum_{|ih-y|\leq 1} \left(1-(ih-y)^2\right)^{M-\frac{1}{2}} =$$

$$= V_{m-1}\left(\sum_{i=-N+1}^{N} \left(1-(ih-y)^2\right)^{M-\frac{1}{2}} + r(y,h)\right),$$

wobei

$$r(y,h) = \begin{cases} \left(1-(Nh+y)^2\right)^{M-\frac{1}{2}} & \text{für} \quad y \leq xh \\ 0 & \text{für} \quad y > xh. \end{cases}$$

Sei zunächst $x>0$. Dann gilt $|ih-y|<1$ für $-N+1 \leq i \leq N$, so daß die Terme unter dem Summenzeichen im angegebenen Bereich beliebig oft nach y differenzierbar sind, $r^{(M)}$ hat dagegen bei $y=xh$ eine Singularität.

BEHAUPTUNG. *Sei* $f(y)=\left(1-(Nh+y)^2\right)^{M-\frac{1}{2}}$ *für* $0 \leq y \leq xh$. *Dann gilt*

$$f^{(t)}(y) = P_t(y)\left(1-(Nh+y)^2\right)^{M-\frac{1}{2}-t},$$

wobei P_t *ein Polynom in* y *ist und* $P_t(xh) \neq 0$.

BEWEIS *(durch vollständige Induktion).*

$t = 1$.

$$f'(y) = -2\left(M-\frac{1}{2}\right)(Nh+y)\left(1-(Nh+y)^2\right)^{M-\frac{1}{2}-1}$$

also

$$P_1(y) = -2\left(M-\frac{1}{2}\right)(Nh+y) \quad \text{und} \quad P_1(xh) = -2\left(M-\frac{1}{2}\right) \neq 0.$$

$t \geq 1$.

$$f^{(t+1)}(y) = P_{t+1}(y)\left(1-(Nh+y)^2\right)^{M-\frac{1}{2}-t-1}$$

mit

$$P_{t+1}(y) = \left(1-(Nh+y)^2\right)P_t'(y) - 2\left(M-\frac{1}{2}-t\right)(Nh+y)P_t(y);$$

$$P_{t+1}(xh) = -2\left(M-\frac{1}{2}-t\right)P_t(xh) \neq 0.$$

Insgesamt ist $S^{(M-1)}$ bei $y=xh$ nicht nach y differenzierbar, also S nicht konstant.

Für $x=0$ gilt:

$$S(y,h) = V_{m-1} \sum_{i=-N+1}^{N} \left(1-\left(\frac{i}{N}-y\right)^2\right)^{M-\frac{1}{2}}.$$

Hier ist die M-te Ableitung nach y für den Summanden mit $i=N$ bei $y=0$ singulär, während die übrigen Summanden im betrachteten Bereich beliebig oft differenzierbar sind. Der Beweis verläuft völlig analog zum Fall $x>0$. Aus [1] ergibt sich mit

$$a_n^{(m)} = \min_h \max_y hS(y, h),$$

$$A_n^{(m)} = \max_h \min_y hS(y, h) \qquad \left(\frac{2}{n+1} \leqq h \leqq \frac{2}{n};\ 0 \leqq y \leqq \frac{1}{2} h\right):$$

Lemma 2. *Seien* $n, m \in \mathbf{N}$, $m \geqq 2$, *so gilt*

$$a_n^{(m)} > V_m,\ A_n^{(m)} < V_m.$$

Weiter benötigen wir das Analogon der Aussage 4.2 aus [1] für gerade $m=2M$.

Lemma 3.

$$\int_0^1 (1-z^2)^{M-\frac{1}{2}} \cos\left(\frac{2\pi r z}{h}\right) dz = \frac{1}{2} \frac{\Gamma\left(M+\frac{1}{2}\right)\sqrt{\pi}}{(\pi r)^M} h^M J_M\left(\frac{2\pi r}{h}\right),$$

wobei

$$J_M(w) = \sqrt{\frac{2}{\pi w}} \left(\cos(w-\varphi_M) - \frac{4M^2-1}{8w} \sin(w-\varphi_M)\right) + O(w^{-\frac{5}{2}})$$

$\left(w>0,\ \varphi_M = \frac{2M+1}{4}\pi\right)$ *die M-te Bessel-Funktion bezeichnet.*

Beweis. [3], S. 366, 368.

Mit diesem Lemma ergibt die zu 5 [1] analoge Rechnung:

Lemma 4. *Seien* $m, n \in \mathbf{N}$, $m=2M \geqq 2$, $\frac{2}{n+1} \leqq h \leqq \frac{2}{n}$, $0 \leqq y \leqq \frac{1}{2} h$. *Dann gilt:*

$$hS(y, h) = V_m + L_m h^{M+\frac{1}{2}} \left(f_M\left(\frac{1+y}{h}\right) + f_M\left(\frac{1-y}{h}\right)\right) + O(h^{M+\frac{3}{2}}),$$

wobei

$$L_m = V_{m-1} \Gamma\left(M+\frac{1}{2}\right) \pi^{-M-\frac{1}{2}}$$

und

$$f_M(x) = \sum_{r=1}^{\infty} r^{-M-\frac{1}{2}} \sin\left(2\pi r x - \frac{2M-1}{4}\pi\right).$$

Wir werden im folgenden für den öfter vorkommenden Term

$$f_i\left(\frac{1+y}{h}\right) + f_i\left(\frac{1-y}{h}\right) \quad \text{kurz } F_i(y, h) \text{ schreiben.}$$

3

In diesem Kapitel sollen die f_M näher untersucht werden. Zunächst ist f_M für $M \geqq 1$ und jedes x konvergent. Die beiden folgenden Aussagen bekommt man durch einfaches Nachrechnen.

3.1. *Für* $M \geqq 2$ *gilt:*

$$f'_M(x) = -2\pi f_{M-1}(x).$$

3.2. *Für* $M \geqq 1$

$$f_M(0) = f_M(1) = -\sin\left(\frac{2M-1}{4}\pi\right)\zeta\left(M+\frac{1}{2}\right)$$

(ζ bezeichne die Riemannsche Zeta-Funktion)

$$f_M\left(\frac{1}{2}\right) = -(1-2^{-M+\frac{1}{2}})f_M(0).$$

Aus 3.2 folgt, daß f_M für $M \geqq 1$ in $\left(0; \frac{1}{2}\right)$ und $\left(\frac{1}{2}; 1\right)$ wenigstens je eine Nullstelle besitzt.

3.3. f_M *besitzt für* $M \geqq 1$ *genau je eine Nullstelle in* $\left(0; \frac{1}{2}\right)$ *und* $\left(\frac{1}{2}; 1\right)$. *Diese Nullstellen sind einfach.*

Beweis. Nach [3], S. 268 gilt:

$$f_M(x) = \frac{(2\pi)^{M+\frac{1}{2}}}{2\Gamma\left(M+\frac{1}{2}\right)}\zeta\left(-M+\frac{1}{2}, x\right) \quad \text{für} \quad M \geqq 0$$

und $0 < x \leqq 1$.

Dabei bezeichnet $\zeta(s, a)$ die für $\operatorname{Re}(s) > 1$ und $0 < a \leqq 1$ durch $\zeta(s, a) = \sum_{r=0}^{\infty} (r+a)^{-s}$ definierte verallgemeinerte ζ-Funktion. Diese Funktion läßt sich wie die Riemannsche ζ-Funktion zu einer in der ganzen komplexen Ebene meromorphen Funktion fortsetzen und besitzt eine einfache Polstelle bei $s = 1$ (s. [3], S. 265 ff). Wie man leicht nachrechnet, gilt für diese Funktion

$$\frac{\partial}{\partial a}\zeta(s, a) = -s\zeta(s+1, a).$$

Nun ist

$$\frac{d}{dx}\zeta\left(\frac{1}{2}, x\right) = -\frac{1}{2}\zeta\left(\frac{3}{2}, x\right) = -\frac{1}{2}\sum_{r=0}^{\infty}\frac{1}{(r+x)^{3/2}} < 0$$

für $0 < x \leqq 1$. Daher ist $f_0(x) := \frac{1}{\sqrt{2}}\zeta\left(\frac{1}{2}, x\right)$ streng monoton fallend für $0 < x \leqq 1$.

Weiter bekommt man wie bei der analytischen Fortsetzung der Riemannschen ζ-Funktion

$$\zeta\left(\frac{1}{2}, x\right) = -2\sqrt{x} + \frac{1}{\sqrt{x}} - \frac{1}{2}\sum_{r=0}^{\infty}\int_0^1 \frac{t}{(r+x+t)^{\frac{3}{2}}}\,dt.$$

Daraus folgt:

$$\lim_{x\to 0_+} f_0(x) = +\infty, \qquad f_0(1) < 0.$$

Daher besitzt f_0 genau eine einfache Nullstelle in (0; 1). Wegen $f_1'(x) = -2\pi f_0(x)$ besitzt f_1 dann genau ein relatives Extremum in (0, 1), es gibt genau zwei Nullstellen, von denen nach 3.2 eine in $\left(0, \frac{1}{2}\right)$, die andere in $\left(\frac{1}{2}, 1\right)$ liegt, und diese Nullstellen sind einfach.

Der Rest folgt durch Induktion mit Hilfe von 3.1. Sei also schon bewiesen, daß f_M $(M \geqq 1)$ genau zwei Nullstellen in den richtigen Intervallen hat.

Annahme 1. f_{M+1} besitzt zwei verschiedene Nullstellen z.B. in $\left(0, \frac{1}{2}\right)$. Wegen $f_{M+1}(0)\, f_{M+1}\left(\frac{1}{2}\right) < 0$ gibt es dann entweder sogar drei Nullstellen oder zwei, von denen eine mehrfach ist. In beiden Fällen hat f'_{M+1}, also auch f_M zwei Nullstellen in $\left(0, \frac{1}{2}\right)$. Widerspruch!

Annahme 2. f_{M+1} besitzt genau eine Nullstelle in $\left(0, \frac{1}{2}\right)$, aber diese ist nicht einfach. Dann ist sie von ungerader Vielfachheit $\geqq 3$, weil $f_{M+1}(0) f_{M+1}\left(\frac{1}{2}\right) < 0$. Dann hat aber f_M dort eine Nullstelle gerader Vielfachheit $\geqq 2$. Widerspruch!

Wir bezeichnen im folgenden mit x_M, x'_M die beiden eindeutig bestimmten Nullstellen von f_M in (0, 1), und zwar so, daß $x_M \in \left(0, \frac{1}{2}\right)$, $x'_M \in \left(\frac{1}{2}, 1\right)$.

3.4. *Für die Nullstellen* x_M, x'_M *gilt*

1. *falls* $M \equiv 0(2)$:

$$x_M = \frac{3}{8} - \frac{\vartheta_1}{2\pi}, \quad \vartheta_1 = \frac{1}{2^{M+1}} - \frac{1}{3^c} + R_1$$

$$x'_M = \frac{7}{8} + \frac{\vartheta_2}{2\pi}, \quad \vartheta_2 = \frac{1}{2^{M+1}} - \frac{1}{3^c} + R_2$$

$$c = M + \frac{1}{2}, \quad |R_1|, |R_2| < \frac{9{,}3}{4^c}$$

$$x_M < x_{M+2} < \frac{3}{8}, \quad x'_M > x'_{M+2} > \frac{7}{8}.$$

2. *falls* $M \equiv 1(2)$:

$$x_M = \frac{1}{8} - \frac{\vartheta_3}{2\pi}, \quad \vartheta_3 = \frac{1}{2^{M+1}} + \frac{1}{3^c} + R_3$$

$$x'_M = \frac{5}{8} - \frac{\vartheta_4}{2\pi}, \quad \vartheta_4 = \frac{1}{2^{M+1}} + \frac{1}{3^c} + R_4$$

$$|R_3|, |R_4| < \frac{5{,}3}{4^c}$$

$$x_M < x_{M+2} < \frac{1}{8}, \quad x'_M > x'_{M+2} > \frac{5}{8}.$$

Der Beweis dieser Aussagen verläuft jeweils analog, ich gebe ihn daher nur für $M \equiv 0(2)$ und x_M an. Sei also $M = 2\mu$, $\mu \in \mathbf{N}$. Dann gilt

$$f_M(x) = \sum_{r=1}^{\infty} r^{-c} \sin\left(2\pi r x - \mu\pi + \frac{\pi}{4}\right) =$$

$$= (-1)^{\mu} \sum_{r=1}^{\infty} r^{-c} \sin\left(2\pi r x + \frac{\pi}{4}\right).$$

Im folgenden sei $\sigma(x) := (-1)^{\mu} f_M(x)$. Offenbar ist $\sigma(0) = \frac{1}{\sqrt{2}} \zeta(c) > 0$. Weiter gilt:

$$\sigma\left(\frac{3}{8}\right) = \frac{1}{2^c} \sin\left(\frac{7}{4}\pi\right) + \frac{1}{3^c} \sin\left(\frac{5}{2}\pi\right) + \frac{1}{4^c} \sin\left(\frac{13}{4}\pi\right) + R_5; \quad |R_5| \leqq \frac{5}{c-1} 5^{-c}$$

$$= -\frac{1}{2^{M+1}} + \frac{1}{3^c} - \frac{1}{4^{M+1}} + R_5 =$$

$$= -\frac{1}{2^{M+1}} \left(1 - \sqrt{3}\left(\frac{2}{3}\right)^{M+1} + \frac{1}{2^{M+1}}\right) + R_5 < 0 \quad \text{für} \quad M \geqq 2.$$

$$\sigma\left(\frac{1}{4}\right) = \sin\left(\frac{3}{4}\pi\right) + \frac{1}{2^c} \sin\left(\frac{5}{4}\pi\right) + \frac{1}{3^c} \sin\left(\frac{7}{4}\pi\right) - \frac{1}{4^c} \sin\left(\frac{9}{4}\pi\right) +$$

$$+ \frac{1}{5^c} \sin\left(\frac{11}{4}\pi\right) + R_6, \quad |R_6| \leqq \frac{5}{c-1} 5^{-c}$$

$$= \frac{1}{\sqrt{2}} - \frac{1}{2^{M+1}} - \frac{1}{\sqrt{2}} \frac{1}{3^c} + \frac{1}{4^{M+1}} + \frac{1}{5^c} \frac{1}{\sqrt{2}} + R_6 > 0 \quad \text{für} \quad M \geqq 2.$$

Insgesamt ergibt sich $\frac{1}{4} < x_M < \frac{3}{8}$, und wenn man $x_M = \frac{3}{8} - \frac{\vartheta}{2\pi}$ setzt, so ist

$0<\vartheta<\frac{\pi}{4}$. Daraus folgt

$$\vartheta<\frac{\pi}{4}\sqrt{2}\sin(\vartheta)<1{,}12\sin(\vartheta).$$

Als grobe Abschätzung für ϑ erhält man $\vartheta<\frac{1{,}03}{2^{c-1}}$ für $M\geqq 2$.

BEWEIS.

$$0=\sigma(x_M)=\sin(\pi-\vartheta)+\frac{1}{2^c}\sin\left(\frac{7}{4}\pi-2\vartheta\right)+\ldots-\frac{1}{5^c}\sin(5\vartheta)+R_7;\quad |R_7|\leqq\frac{5}{c-1}5^{-c}\Rightarrow$$

$$\Rightarrow\sin(\vartheta)<\frac{1}{2^c}+\ldots+\frac{1}{5^c}+\frac{5}{c-1}\frac{1}{5^c}<\frac{1}{2^{c-1}}\quad\text{für}\quad M\geqq 2$$

$$\vartheta\leqq\sin(\vartheta)+\frac{\vartheta^3}{6}<\sin(\vartheta)+\frac{1}{6}(1{,}12)^3\sin^3(\vartheta)<\frac{1{,}03}{2^{c-1}}\quad\text{für}\quad M\geqq 2.$$

Schließlich gilt

$$\sin(\vartheta)=\frac{1}{2^c}\sin\left(\frac{\pi}{4}+2\vartheta\right)+\ldots+\frac{1}{5^c}\sin(5\vartheta)+R_7=\frac{1}{2^{M+1}}-\frac{1}{3^c}+R_8,$$

$$|R_8|<\frac{1{,}03}{\sqrt{2}}\frac{1}{2^{2c-1}}+\frac{9{,}62}{2^{2c-2}3^c}+\frac{1}{4^c}+\frac{1}{5^c}+\frac{1}{5^c}\frac{5}{c-1}<\frac{9}{4^c}.$$

Aus

$$|\sin(\vartheta)-\vartheta|\leqq\frac{\vartheta^3}{3!}+\frac{\vartheta^5}{5!}+\ldots<\frac{\vartheta^3}{6}\frac{1}{1-\vartheta^2}<\frac{1{,}22}{8^c}\quad\text{für}\quad M\geqq 2$$

folgt dann

$$\vartheta=\frac{1}{2^{M+1}}-\frac{1}{3^c}+R_1,$$

wie behauptet. Diese Abschätzung ergibt $x_M<x_{M+2}$, wenn

$$\frac{1}{2^{M+1}}-\frac{1}{3^c}-\frac{9{,}3}{4^c}>\frac{1}{2^{M+3}}-\frac{1}{3^{c+1}}+\frac{9{,}3}{4^{c+2}}$$

ist. Das ist für $M\geqq 5$ erfüllt. Die restlichen Fälle lassen sich durch direktes Nachrechnen erledigen. Die Ergebnisse sind in der folgenden Tabelle zusammengestellt:

M	x_M	x'_M
1	$0{,}06<x_1<0{,}07$	$0{,}65<x'_1<0{,}66$
2	$0{,}355<x_2<0{,}360$	$0{,}901<x'_2<0{,}902$
3	$0{,}110<x_3<0{,}113$	$0{,}633<x'_3<0{,}634$
4	$0{,}370<x_4<0{,}371$	$0{,}880<x'_4<0{,}881$
5	$0{,}122<x_5<0{,}123$	$0{,}627<x'_5<0{,}628$
6	$0{,}373<x_6<0{,}374$	$0{,}876<x'_6<0{,}877$

3.5. *Es gilt*

$$|f_M(x_{M-1})| > |f_M(x'_{M-1})| \quad \text{für} \quad M \equiv 0(2)$$

$$|f_M(x_{M-1})| < |f_M(x'_{M-1})| \quad \text{für} \quad M \equiv 1(2),\ M > 1.$$

Ich gebe wieder nur den Beweis für $M \equiv 0(2)$ an, um Wiederholungen zu vermeiden.

BEWEIS.

Für $M \equiv 0(2)$ gilt:

$$x_{M-1} < \frac{1}{8}, \quad \text{also} \quad |f_M(x_{M-1})| > \left|f_M\left(\frac{1}{8}\right)\right|.$$

Nun ist

$$\left|f_M\left(\frac{1}{8}\right)\right| > 1 + \frac{1}{2^{M+1}} - \frac{1}{4^c} - \frac{1}{c-1}\,\frac{2}{4^M}$$

und

$$|f_M(x'_{M-1})| < 1 - \frac{1}{2^{M+1}} + \frac{1}{4^c} + \frac{1}{c-1}\,\frac{2}{4^M} +$$

$$+\left(\frac{1}{2^M} - \frac{1}{3^{c-1}} + \frac{10{,}6}{4^M}\right)\cdot\left(\frac{1}{2^M} + \frac{1}{c-1}\,\frac{\sqrt{8}}{2^{M-1}}\right).$$

Man rechnet nach, daß sicher $\left|f_M\left(\frac{1}{8}\right)\right| > |f_M(x'_{M-1})|$ für $M \geqq 6$.

Die restlichen Fälle lassen sich wieder numerisch erledigen. Die Ergebnisse sind in den folgenden Graphen enthalten.

Um

$$a_n^{(m)} = \min_h \max_y hS(y, h), \quad \text{bzw.} \quad A_n^{(m)} = \max_h \min_y hS(y, h)$$

abzuschätzen, bleiben noch die dem Lemma 4 aus [1] entsprechenden Aussagen zu beweisen.

LEMMA 5. a) *Sei* $m = 2M+1$, $M \in \mathbf{N}$, $n \in \mathbf{N}$. *Dann gilt*

1. *für* $M \equiv 0(4)$ *und alle hinreichend kleinen* h

$$\min_h \max_y \left(T_{M+1}(y, h) - \frac{1}{2}\,\frac{M(M+1)}{M+2}\,h\,T_{M+2}(y, h)\right) \geqq c_1 h;$$

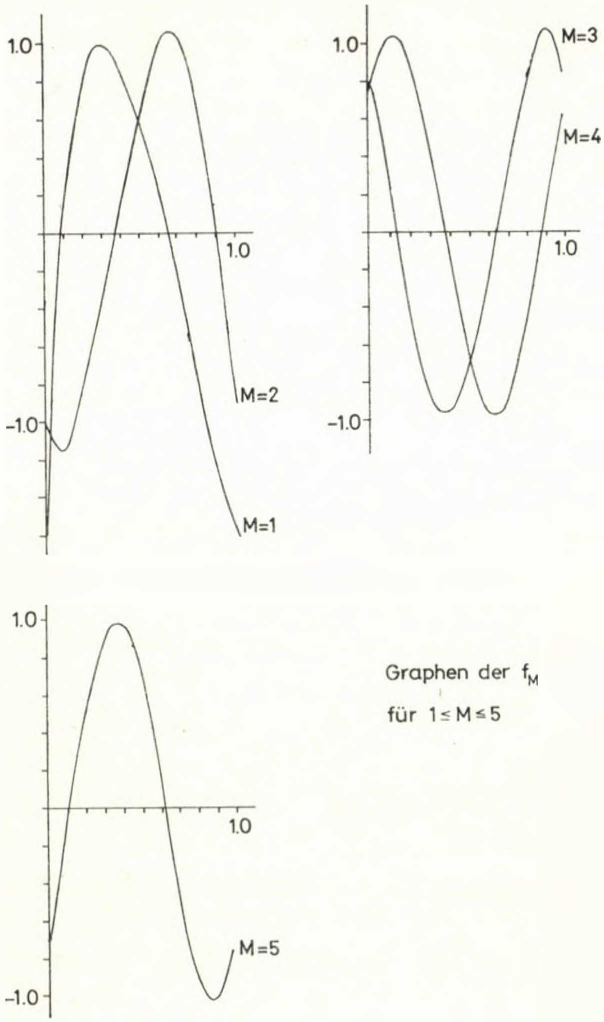
2. *für* $M \equiv 2(4)$ *und alle hinreichend kleinen* h

$$\max_h \min_y \left(T_{M+1}(y, h) - \frac{1}{2}\,\frac{M(M+1)}{M+2}\,h\,T_{M+2}(y, h)\right) \leqq -c_2 h.$$

b) *Sei* $m = 2M$, $M \in \mathbf{N}$ *und* $M > 1$. *Dann gilt*

$$\min_h \max_y F_M(y, h) \geqq c_3 > 0$$

$$\max_h \min_y F_M(y, h) \leqq -c_4 < 0.$$



Graphen der f_M für $1 \leq M \leq 5$

Dabei sind die c_i nur von m abhängig und positiv, $T_i(y, h) = B_i(a_1) + B_i(a_2)$ *die in* [1] *eingeführten Funktionen, d.h.* $a_{1/2} \in]0, 1]$ *mit* $a_{1/2} \equiv \frac{1 \pm y}{h} \bmod 1$ *und B_i das i-te Bernoulli-Polynom. Ferner gelte*

$$\frac{2}{n+1} \leqq h \leqq \frac{2}{n}, \quad 0 \leqq y \leqq \frac{1}{2} h.$$

Beweis. a) 1. Im Beweis zu a) benutze ich die Bezeichnungsweise aus [1]; hier bedeute also x_i ausnahmsweise die Nullstelle von $B_i(x)$ in $\left(0, \frac{1}{2}\right)$. Ferner sei zur

Abkürzung

$$\tau := T_{M+1}-\frac{1}{2}\frac{M(M+1)}{M+2}hT_{M+2} \qquad (\text{und} \quad m=2M+1 \quad \text{mit} \quad M\equiv 0(4)).$$

$h\in\left[\frac{2}{n+1};\frac{2}{n}\right]$ läßt sich in der Form

$$h=\frac{2}{n+2x}, \quad 0\leqq x\leqq \frac{1}{2}$$

schreiben. Sei zunächst $n=2N$ gerade, also $h=\frac{1}{N+x}$. Dann geben wir für jedes $x\in\left[0,\frac{1}{2}\right]$ ein $y_1\in\left[0,\frac{h}{2}\right]$ an, so daß $\tau(y_1,h)>0$ ist.

1.1. $0\leqq x\leqq\frac{1}{2}x_{M+2}$: $y_1=\frac{1}{2}h$

Dann gilt

$$\left.\begin{array}{l} T_{M+1}(y_1,h)\geqq 0 \\ T_{M+2}(y_1,h)\leqq 2B_{M+2}\left(\frac{5}{8}\right)<0 \end{array}\right\} \Rightarrow \tau(y_1,h)>ch,\ c>0.$$

1.2. $\frac{1}{2}x_{M+2}<x\leqq\frac{1}{2}-\varepsilon$, wobei $\varepsilon=\frac{(M+\delta)(M+1)}{8\pi^2}h$ und $\delta=\frac{1}{48}\frac{M}{2^M}$.

Mit $y_1=\frac{1}{2}h$ gilt:

$$T_{M+1}(y_1,h)\geqq 2B_{M+1}(1-\varepsilon)=-2B_{M+1}(\varepsilon)=-2\varepsilon(M+1)B_M(0)+O(h^2)$$

$$T_{M+2}(y_1,h)<2B_{M+2}(0),$$

also

$$\tau(y_1,h)\geqq 2\frac{(M+\delta)(M+1)^2}{8\pi^2}h|B_M(0)|-\frac{M(M+1)}{M+2}h|B_{M+2}(0)|+O(h^2)=$$

$$=\frac{M(M+1)}{M+2}h|B_M(0)|\left(\frac{(M+\delta)(M+1)(M+2)}{4\pi^2 M}-\left|\frac{B_{M+2}(0)}{B_m(0)}\right|\right)+O(h^2)\geqq$$

$$\geqq\frac{M(M+1)}{M+2}h|B_M(0)|\left(\frac{(M+\delta)(M+1)(M+2)}{4\pi^2 M}-\frac{(M+1)(M+2)}{4\pi^2}\right)+O(h^2)=$$

$$=\frac{1}{4\pi^2}(M+1)^2\delta h|B_M(0)|+O(h^2)\geqq c'h, \quad c'>0.$$

1.3. $\frac{1}{2}-\varepsilon<x\leqq\frac{1}{2}$: $y_1=0$

$$T_{M+1}(y_1, h) \geqq 2B_{M+1}\left(\frac{1}{2}-\varepsilon\right) \geqq -2\varepsilon(M+1)B_M\left(\frac{1}{2}\right)$$

$$T_{M+2}(y_1, h) \leqq 2B_{M+2}\left(\frac{1}{2}-\varepsilon\right) = 2B_{M+2}\left(\frac{1}{2}\right)+O(h^2),$$

also

$$\tau(y_1, h) \geqq -\frac{1}{4\pi^2}(M+\delta)(M+1)^2 hB_M\left(\frac{1}{2}\right)+\frac{M(M+1)}{M+2}h\left|B_{M+2}\left(\frac{1}{2}\right)\right|+O(h^2) =$$

$$=\frac{M(M+1)}{M+2}h\left|B_M\left(\frac{1}{2}\right)\right|\left(\left|\frac{B_{M+2}\left(\frac{1}{2}\right)}{B_M\left(\frac{1}{2}\right)}\right|-\frac{1}{4\pi^2}\frac{(M+\delta)(M+1)(M+2)}{M}\right)+O(h^2) =$$

$$=\frac{M(M+1)^2}{4\pi^2}h\left|B_M\left(\frac{1}{2}\right)\right|K+O(h^2)$$

mit

$$K=\frac{1-2^{-M-1}}{1-2^{-M+1}}\frac{\zeta(M+2)}{\zeta(M)}-1-\frac{\delta}{M}.$$

Sei für

$$s\geqq 2,\ s\in\mathbf{N}:\ e_s=(1-2^{-s+1})\zeta(s) = \sum_{j=1}^{\infty} j^{-s}-2\sum_{j=1}^{\infty}(2j)^{-s} = \sum_{j=1}^{\infty}(-1)^{j+1}j^{-s}.$$

Dann gilt:

$$\left.\begin{matrix} e_{M+2} > 1-2^{-M-2} \\ e_M < 1-2^{-M}+3^{-M}\end{matrix}\right\} \Rightarrow \frac{e_{M+2}}{e_M} > \frac{1-2^{-M-2}}{1-2^{-M}+3^{-M}} \geqq \frac{2^{M+2}-1}{2^{M+2}-\frac{4}{3}} =$$

$$=\frac{1-2^{-M-2}}{1-\frac{4}{3}2^{-M-2}} > (1-2^{-M-2})\left(1+\frac{4}{3}2^{-M-2}\right) \geqq 1+\frac{1}{24}2^{-M}.$$

Damit haben wir:

$$K=\frac{e_{M+2}}{e_M}-1-\frac{\delta}{M} \geqq \frac{1}{48}2^{-M} > 0 \Rightarrow \tau(y_1, h) > c''h > 0.$$

Insgesamt:

$$\min_h \max_y \tau(y, h) \geqq c_1 h > 0.$$

Für $n=2N+1$ läßt sich h in der Form $h=\frac{1}{N+z}$, $z\in\left[\frac{1}{2}, 1\right]$ schreiben. Ich gebe jeweils die zu wählenden y_1 an; die Rechnung ist völlig analog.

$$\frac{1}{2} \leqq z \leqq \frac{5}{8}: y_1 = 0$$

$$\frac{5}{8} < z \leqq 1-\varepsilon: y_1 = 0$$

$$1-\varepsilon < z \leqq 1: y_1 = \frac{1}{2} h.$$

2. Sei jetzt $m=2M+1$ und $M\equiv 2(4)$. Wir betrachten zunächst den Fall $n\equiv 0(2)$, also $n=2N$ und $h=\dfrac{1}{N+x}$ wie oben. ε sei jetzt $=\dfrac{M(M+1)}{8\pi^2}h$.

2.1. $$0 \leqq x \leqq \frac{1}{8}: y_1 = \frac{1}{2} h$$

$$\left.\begin{array}{l} T_{M+1}(y_1, h) \leqq 0 \\ T_{M+2}(y_1, h) \geqq 2B_{M+1}\left(\frac{5}{8}\right) \end{array}\right\} \to \tau(y_1, h) \leqq -ch$$

2.2. $$\frac{1}{8} < x \leqq \frac{1}{2}-\varepsilon: y_1 = \frac{1}{2} h$$

$$\left.\begin{array}{l} T_{M+1}(y_1, h) \leqq 2B_{M+1}(1-\varepsilon) = -2\varepsilon(M+1)|B_M(0)|+O(h^2) \\ T_{M+2}(y_1, h) \geqq -2|B_{M+2}(0)| \end{array}\right\} \Rightarrow$$

$$\tau(y_1, h) \leqq -\frac{M(M+1)^2}{4\pi^2} h|B_M(0)| + \frac{M(M+1)}{M+2} h|B_{M+2}(0)| + O(h^2) =$$

$$= \frac{M(M+1)}{M+2} h|B_M(0)| \left(\left|\frac{B_{M+2}(0)}{B_M(0)}\right| - \frac{(M+1)(M+2)}{4\pi^2}\right) + O(h^2) < -c' h.$$

2.3. $$\frac{1}{2}-\varepsilon < x \leqq \frac{1}{2}: y_1 = 0$$

$$T_{M+1}(y_1, h) \leqq 2B_{M+1}\left(\frac{1}{2}-\varepsilon\right) = 2(M+1)B_M\left(\frac{1}{2}\right) + O(h^2)$$

$$T_{M+2}(y_1, h) \geqq 2B_{M+2}\left(\frac{1}{2}-\varepsilon\right) = 2B_{M+2}\left(\frac{1}{2}\right) + O(h^2)$$

$$\Rightarrow \tau(y_1, h) \leqq -\frac{M(M+1)}{M+2} h \left|B_M\left(\frac{1}{2}\right)\right| \left(\left|\frac{B_{M+2}\left(\frac{1}{2}\right)}{B_M\left(\frac{1}{2}\right)}\right| - \frac{(M+1)(M+2)}{4\pi^2}\right) + O(h^2) =$$

$$= -\frac{M(M+1)^2}{4\pi^2} h \left|B_M\left(\frac{1}{2}\right)\right| \left(\frac{e_{M+2}}{e_M} - 1\right) + O(h^2) < -c'' h.$$

Insgesamt ergibt sich:

$$\max_h \min_y \tau(y_1, h) \leqq -c_2 h < 0.$$

Der Fall $n=2N+1$ läßt sich wieder völlig analog behandeln mit

$$\frac{1}{2} \leqq z \leqq \frac{5}{8}: \; y_1 = 0$$

$$\frac{5}{8} < z \leqq 1-\varepsilon: \; y_1 = 0$$

$$1-\varepsilon < z \leqq 1: \; y_1 = \frac{1}{2} h.$$

Damit ist der Beweis zu a) beendet.

b) Sei von nun an $m=2M$ gerade, x_M wieder die eindeutig bestimmte Nullstelle von f_M in $\left(0, \frac{1}{2}\right)$ und

$$F_M(y, h) = f_M\left(\frac{1+y}{h}\right) + f_M\left(\frac{1-y}{h}\right).$$

1. Wir betrachten zunächst den Fall $M \equiv 0(4)$.

1.1. $n=2N$ gerade, also $h=\frac{1}{N+x}$ mit $0 \leqq x \leqq \frac{1}{2}$. Dann gilt nach 3.5:

$$|f_M(x_{M-1})| > |f_M(x'_{M-1})|.$$

Wenn man $\frac{y_1}{h} = |x - x_{M-1}|$ wählt, so gilt

$$f_M\left(\frac{1 \pm y_1}{h}\right) = f_M\left(N + x \pm \frac{y_1}{h}\right) = f_M\left(x \pm \frac{y_1}{h}\right),$$

da f_M periodisch ist mit der Periode 1, also

$$F_M(y_1, h) \geqq f_M(x_{M-1}) + f_M(x'_{M-1}) \geqq |f_M(x_{M-1})| - |f_M(x'_{M-1})| > 0.$$

Daher ist auch $\min_h \max_y F_M(y, h) > 0$. Wählt man

$$\frac{y_2}{h} = \begin{cases} \frac{1}{2} & \text{für} \quad 0 \leqq x \leqq \frac{3}{8} \\ 0 & \text{für} \quad \frac{3}{8} < x \leqq \frac{1}{2}, \end{cases}$$

so erhält man

$$F_M(y_2, h) \leqq 2 \min\left\{f_M\left(\frac{1}{2}\right), f_M\left(\frac{7}{8}\right)\right\} < 0 \quad \text{für} \quad 0 \leqq x \leqq \frac{3}{8},$$

bzw.

$$F_M(y_2, h) \leqq 2 f_M\left(\frac{3}{8}\right) < 0 \quad \text{für} \quad \frac{3}{8} < x \leqq \frac{1}{2}$$

nach 3.4, also

$$\max_h \min_y F_M(y, h) < 0.$$

1.2. Für $n=2N+1$, also $h=\frac{1}{N+z}$ mit $\frac{1}{2}\leqq z\leqq 1$ verläuft alles ganz ähnlich. Ich gebe daher nur die zu wählenden y_i an:

$$\frac{y_1}{h} := \begin{cases} z-x_{M-1} & \text{für } \frac{1}{2}\leqq z\leqq \frac{1}{2}+x_{M-1} \\ 1-(z-x_{M-1}) & \text{für } \frac{1}{2}+x_{M-1}< z\leqq 1 \end{cases} \Rightarrow$$

$$\Rightarrow F_M(y_1, h) \geqq |f_M(x_{M-1})|-|f_M(x'_{M-1})| > 0$$

$$\frac{y_2}{h} := \begin{cases} 0 & \text{für } \frac{1}{2}\leqq z\leqq \frac{7}{8} \Rightarrow F_M(y_2, h) \leqq 2\max\left\{f_M\left(\frac{1}{2}\right), f_M\left(\frac{7}{8}\right)\right\} < 0 \\ \frac{1}{2} & \text{für } \frac{7}{8}< z\leqq 1 \Rightarrow F_M(y_2, h)\leqq 2f_M\left(\frac{3}{8}\right) < 0. \end{cases}$$

2. $M\equiv 1(4)$, $M\geqq 2$. Hier gilt nach 3.5

$$|f_M(x_{M-1})| < |f_M(x'_{M-1})|.$$

2.1. $n=2N$

$$\frac{y_1}{h} = \begin{cases} \frac{1}{2} & \text{für } 0\leqq x\leqq \frac{1}{8} \Rightarrow F_M(y_1, h)\geqq 2f_M\left(\frac{5}{8}\right) > 0 \\ 0 & \text{für } \frac{1}{8}< x\leqq \frac{1}{2} \Rightarrow F_M(y_1, h)\geqq 2\min\left\{f_M\left(\frac{1}{8}\right), f_M\left(\frac{1}{2}\right)\right\} > 0. \end{cases}$$

$$\frac{y_2}{h} = \begin{cases} 1-(x'_{M-1}-x) & \text{für } 0\leqq x\leqq x'_{M-1}-\frac{1}{2} \\ x'_{M-1}-x & \text{für } x'_{M-1}-\frac{1}{2}< x\leqq \frac{1}{2} \end{cases} \Rightarrow$$

$$\Rightarrow F_M(y_2, h) \leqq -|f_M(x'_{M-1})|+|f_M(x_{M-1})| < 0.$$

2.2. $n=2N+1$.

$$\frac{y_1}{h} = \begin{cases} 0 & \text{für } \frac{1}{2}\leqq z\leqq \frac{5}{8} \Rightarrow F_M(y_1, h)\geqq 2f_M\left(\frac{5}{8}\right) > 0 \\ \frac{1}{2} & \text{für } \frac{5}{8}< z\leqq 1 \Rightarrow F_M(y_1, h)\geqq 2\min\left\{f_M\left(\frac{1}{8}\right), f_M\left(\frac{1}{2}\right)\right\} > 0 \end{cases}$$

$$\frac{y_2}{h} = |x'_{M-1}-z| \Rightarrow F_M(y_2, h) \leqq -|f_M(x'_{M-1})|+|f_M(x_{M-1})| < 0.$$

3. $M \equiv 2(4)$.
3.1. $n=2N$.

$$\frac{y_1}{h} = \begin{cases} \frac{1}{2} & \text{für} \quad 0 \leqq x \leqq \frac{3}{8} \Rightarrow F_M(y_1, h) \geqq 2 \min\left\{f_M\left(\frac{1}{2}\right), f_M\left(\frac{7}{8}\right)\right\} > 0 \\ 0 & \text{für} \quad \frac{3}{8} < x \leqq \frac{1}{2} \Rightarrow F_M(y_1, h) \geqq 2 f_M\left(\frac{3}{8}\right) > 0, \end{cases}$$

$$\frac{y_2}{h} = |x - x_{M-1}| \Rightarrow F_M(y_2, h) \leqq -|f_M(x_{M-1})| + |f_M(x'_{M-1})| < 0.$$

3.2. $n=2N+1$

$$\frac{y_1}{h} = \begin{cases} 0 & \text{für} \quad \frac{1}{2} \leqq z \leqq \frac{7}{8} \Rightarrow F_M(y_1, h) \geqq 2 \min\left\{f_M\left(\frac{1}{2}\right), f_M\left(\frac{7}{8}\right)\right\} > 0 \\ \frac{1}{2} & \text{für} \quad \frac{7}{8} < z \leqq 1 \Rightarrow F_M(y_1, h) \geqq 2 f_M\left(\frac{3}{8}\right) > 0, \end{cases}$$

$$\frac{y_2}{h} = \left\{\begin{matrix} z - x_{M-1} & \text{für} \quad \frac{1}{2} \leqq z \leqq \frac{1}{2} + x_{M-1} \\ 1 - (z - x_{M-1}) & \text{für} \quad \frac{1}{2} + x_{M-1} < z \leqq 1 \end{matrix}\right\} \Rightarrow$$

$$\Rightarrow F_M(y_2, h) \leqq -|f_M(x_{M-1})| + |f_M(x'_{M-1})| < 0.$$

4. $M \equiv 3(4)$.
4.1. $n=2N$.

$$\frac{y_1}{h} = \left\{\begin{matrix} 1 - (x'_{M-1} - x) & \text{für} \quad 0 \leqq x \leqq x'_{M-1} - \frac{1}{2} \\ x'_{M-1} - x & \text{für} \quad x'_{M-1} - \frac{1}{2} < x \leqq \frac{1}{2} \end{matrix}\right\} \Rightarrow$$

$$\Rightarrow F_M(y_1, h) \geqq |f_M(x'_{M-1})| - |f_M(x_{M-1})| > 0,$$

$$\frac{y_2}{h} = \begin{cases} \frac{1}{2} & \text{für} \quad 0 \leqq x \leqq \frac{1}{8} \Rightarrow F_M(y_2, h) \leqq 2 f_M\left(\frac{5}{8}\right) < 0 \\ 0 & \text{für} \quad \frac{1}{8} < x \leqq \frac{1}{2} \Rightarrow F_M(y_2, h) \leqq 2 f_M\left(\frac{1}{8}\right) < 0. \end{cases}$$

4.2. $n=2N+1$.

$$\frac{y_1}{h} = |x'_{M-1} - z| \Rightarrow F_M(y_1, h) \geqq |f_M(x'_{M-1})| - |f_M(x_{M-1})| > 0,$$

$$\frac{y_2}{h} = \begin{cases} 0 & \text{für} \quad \frac{1}{2} \leqq z \leqq \frac{5}{8} \Rightarrow F_M(y_2, h) \leqq 2 f_M\left(\frac{5}{8}\right) < 0 \\ \frac{1}{2} & \text{für} \quad \frac{5}{8} < z \leqq 1 \Rightarrow F_M(y_2, h) \leqq 2 f_M\left(\frac{1}{8}\right) < 0. \end{cases}$$

Der Fall $m=2$ ist in [2] ausführlich behandelt. Er läßt sich ebenso wie die übrigen m erledigen, wenn man noch zeigt, daß $\max f_1(x) < |f_1(0)|$ $(x \in [0, 1])$. Man hat dann folgende y_i zu wählen:

$n=2N$:

$$\frac{y_1}{h} = \begin{cases} \frac{1}{2} & \text{für} \quad 0 \leqq x \leqq 0{,}1 \\ 0 & \text{für} \quad 0{,}1 < x \leqq 0{,}5 \end{cases} \Rightarrow F_1(y_1, h) \geqq 2\min\{f_1(0{,}1), f_1(0{,}6)\} > 0,$$

$$\frac{y_2}{h} = x \Rightarrow F_1(y_2, h) \leqq f_1(0) + \max f_1(x) < 0.$$

$n=2N+1$.

$$\frac{y_1}{h} = \begin{cases} 0 & \text{für} \quad 0{,}5 \leqq z \leqq 0{,}6 \\ \frac{1}{2} & \text{für} \quad 0{,}6 < z \leqq 1 \end{cases} \Rightarrow F_1(y_1, h) \geqq 2\min\{f_1(0{,}1), f_1(0{,}6)\} > 0,$$

$$\frac{y_2}{h} = 1 - z \Rightarrow F_1(y_2, h) \leqq f_1(1) + \max f_1(z) < 0.$$

Der Beweis des Satzes folgt nun fast wörtlich wie in [1].

LITERATURVERZEICHNIS

[1] Bolle, U., Dichteabschätzungen für mehrfache gitterförmige Kugelanordnungen in $\mathbf{R}^m$, *Studia Sci. Math. Hungar.* **14** (1979), 51—68.

[2] Bolle, U., Mehrfache Kreisanordnungen in der euklidischen Ebene, Dissertation, Dortmund, 1976.

[3] Whittaker, E. T. und Watson, G. N., *A course of modern analysis*. An introduction to the general theory of infinite processes and of analytic functions; with an account of the principal transcendental functions, Fourth edition, Reprinted, Cambridge University Press, New York, 1962. *MR* **31** #2375.

(Eingegangen am 25. November 1981)

DANGENSDORF 71
D—3131 LÜBBOW 1
FEDERAL REPUBLIC OF GERMANY

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 445—448

AN INVARIANCE THEOREM FOR BIRTH AND DEATH PROCESSES WITH AN ABSORBING BARRIER

W. EBERL JR.

Summary

In this paper it is shown for time homogeneous birth and death processes with an absorbing barrier that the distribution of absorption into this state is invariant against reversion of the order of the birth and death rates.

For some $n \in \mathbf{N}$ let be given positive (real) numbers $\lambda_0, \ldots, \lambda_{n-1}$ and $\mu_1, \ldots, \mu_{n-1}$. Then, according to known theorems on time homogeneous Markov processes (cf. e.g. [1]) there always exist a unique transition function and correspondingly a birth and death process $\mathscr{X}=(X_t)_{t\in[0,\infty)}$ with the (finite) state space $\{0, 1, \ldots, n\}$, with reflecting barrier 0 and absorbing state n such that the birth and death rates take on the preassigned values $\lambda_0, \ldots, \lambda_{n-1}$ resp. $\mu_1, \ldots, \mu_{n-1}$. For simplicity we additionally assume the process to start from the state 0 at time 0 with probability 1.

In various papers the distribution of absorption into state n, i.e. the distribution of the random variable $T_n = \inf\{t \mid X_t = n\}$, resp. its limiting distribution for $n \to \infty$ was investigated, mainly for special classes of transition rates (cf. e.g. [2], [3], [4], [5], [6], [8], [9]).

In the sequel we shall prove by elementary means that for fixed $n \in \mathbf{N}$ the distribution of absorption does not change if the order of the birth and of the death rates is reversed. First, we provide some prerequisites. For $j=1, \ldots, n$ let $T_j := \inf\{t > 0 \mid X_t = j\}$ be the first entrance time into state j, which for $j=n$ is just the above introduced time of absorption. Denoting the reciprocal $1/\Phi_j$ of the Laplace transform Φ_j of the distribution of T_j with Q_j, in [2], p. 82 (5), the following recursion formula was given:

$$\lambda_j Q_{j+1}(s) = (\lambda_j + \mu_j + s) Q_j(s) - \mu_j Q_{j-1}(s) \qquad (1 \leqq j \leqq n-1)$$

$$Q_0 \equiv 1, \quad Q_1(s) = 1 + s/\lambda_0. \tag{1}$$

With the aid of this relationship we can prove the following

THEOREM. *Let $n \in \mathbf{N}$ and let be given positive (real) numbers $\lambda_0, \ldots, \lambda_{n-1}$ resp. $\mu_1, \ldots, \mu_{n-1}$. Further, let P denote the distribution of absorption into state n for a homogeneous birth and death process $\mathscr{X}$ with the above enumerated properties and with the birth and death rates $\lambda_0, \ldots, \lambda_{n-1}$ resp. $\mu_1, \ldots, \mu_{n-1}$. Finally, let P^* designate the distribution of absorption for a corresponding birth and death process $\mathscr{Y}$ with the birth*

1980 *Mathematics Subject Classification*. Primary 60J80.
Key words and phrases. Birth-and-death processes.

and death rates $\lambda_{n-1}, \ldots, \lambda_0$ *resp.* $\mu_{n-1}, \ldots, \mu_1$. *Then*

$$P = P^*.$$

PROOF. We consider the characteristic polynomial D_n of the infinitesimal generator of $\mathscr{X}$ and its principal minors D_j $(1 \leqq j \leqq n-1)$, i.e. we put for $j=1, \ldots, n$

$$D_j(s) = \begin{vmatrix} s+\lambda_0 & -\lambda_0 & & & & & \\ -\mu_1 & s+\lambda_1+\mu_1 & -\lambda_1 & & & 0 & \\ 0 & -\mu_2 & s+\lambda_2+\mu_2 & -\lambda_2 & & & \\ & & \cdot & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot & \cdot \\ & & & -\mu_{j-2} & s+\lambda_{j-2}+\mu_{j-2} & & -\lambda_{j-2} \\ & & & 0 & -\mu_{j-1} & & s+\lambda_{j-1}+\mu_{j-1} \end{vmatrix} \tag{2}$$

with the additional convention $\mu_0 := 0$ for $D_1(s) = s + \lambda_0$. Expanding (2) with respect to the last row we obtain immediately the relationship

$$D_j(s) = (\lambda_{j-1}+\mu_{j-1}+s)D_{j-1}(s) - \mu_{j-1}\lambda_{j-2}D_{j-2}(s), \tag{3}$$

valid for $j=2, \ldots, n$ if we define $D_0 \equiv 1$ (cf. [7]). Hence, putting

$$\tilde{D}_0 \equiv 1, \quad \tilde{D}_j = D_j/(\lambda_0 \ldots \lambda_{j-1}) \qquad (j = 1, \ldots, n)$$

and dividing both sides of (3) by $\lambda_0 \ldots \lambda_{j-1}$, we get

$$\lambda_{j-1}\tilde{D}_j(s) = (\lambda_{j-1}+\mu_{j-1}+s)\tilde{D}_{j-1}(s) - \mu_{j-1}\tilde{D}_{j-2}(s) \qquad (j = 2, \ldots, n)$$

with the initial conditions $\tilde{D}_0 \equiv 1$, $\tilde{D}_1(s) = 1 + s/\lambda_0$. Therefore the polynomials $\tilde{D}_j$ $(j=0, \ldots, n)$ satisfy the same recurrence relation with the same initial conditions as the polynomials Q_j $(j=0, \ldots, n)$, which yields

$$Q_0 = D_0 = 1, \quad Q_j = D_j/(\lambda_0 \ldots \lambda_{j-1}) \qquad (j = 1, \ldots, n). \tag{4}$$

In virtue of (4) and because of the one-to-one correspondence between the Laplace transforms and distributions it suffices to show the validity of

$$D_n = D_n^*, \tag{5}$$

where correspondingly D_n^* denotes the characteristic polynomial of the infinitesimal generator of the process $\mathscr{Y}$. On verifying (5), we introduce the following matrices:

$$\Lambda = \begin{pmatrix} \lambda_0 & & & \\ & \lambda_1 & & 0 \\ & & \ddots & \\ 0 & & & \lambda_{n-1} \end{pmatrix}, \quad M = \begin{pmatrix} 0 & & & \\ \mu_1 & 0 & & 0 \\ & \ddots & \ddots & \\ 0 & & \mu_{n-1} & 0 \end{pmatrix},$$

$$\Lambda^* = \begin{pmatrix} \lambda_{n-1} & & & \\ & \lambda_{n-2} & & 0 \\ & & \ddots & \\ 0 & & & \lambda_0 \end{pmatrix}, \quad M^* = \begin{pmatrix} 0 & & & \\ \mu_{n-1} & 0 & & 0 \\ & \ddots & \ddots & \\ 0 & & \mu_1 & 0 \end{pmatrix},$$

$$V = \begin{pmatrix} 1 & -1 & & & \\ & 1 & -1 & & 0 \\ & & \ddots & \ddots & \\ & 0 & & 1 & -1 \\ & & & & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 0 & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & 0 \end{pmatrix}.$$

With these notations the following identities can readily be seen to hold:

$$\text{(6)} \qquad \begin{gathered} \Lambda^* = U\Lambda U, \quad M^* = UM^T U, \\ U^{-1} = U^T = U, \quad UV = V^T U, \end{gathered}$$

where e.g. M^T stands for the transpose of M. Besides we get for the characteristic polynomials the representations

$$\text{(7)} \qquad D_n(s) = \|(\Lambda - M)V + sI\|, \quad D_n^*(s) = \|(\Lambda^* - M^*)V + sI\|,$$

where I is the $n \times n$ unit matrix.

Using the identities (6) we conclude

$$\begin{aligned} (\Lambda^* - M^*)V + sI &= U(\Lambda - M^T)UV + sI = \\ &= U(V^{-1})^T V^T (\Lambda - M^T)UV + sI = \\ &= U(V^{-1})^T [(\Lambda - M)V]^T V^T U + sI = \\ &= \{UV[(\Lambda - M)V]V^{-1}U\}^T + sI = \\ &= \{UV[(\Lambda - M)V + sI](UV)^{-1}\}^T. \end{aligned}$$

Therefore the matrices $(\Lambda^* - M^*)V + sI$ and $(\Lambda - M)V + sI$ are similar, thus having the same determinants. Because of (7) this proves the validity of (5) and therefore the theorem.

REMARK 1. We have proved in the above theorem that the distribution of absorption remains unaltered if we reverse the order of all birth and of all death rates. Actually, this result does not remain true any more if only some of the rates are permuted. This can be seen by simple counterexamples.

REMARK 2. In the papers [2], [3], [4] the asymptotic distribution of absorption was determined for sequences $(\mathscr{X}^{(n)})_{n \in \mathbf{N}}$ of birth and death processes with the above enumerated properties for special classes of birth and death rates. Especially, constant resp. linearly increasing birth and/or death rates were considered. Our theorem can be applied in an obvious way to get these limiting distributions for rates decreasing linearly with the reached state. But there is an evident difference; whereas considering sequences $(\mathscr{X}^{(n)})$ with increasing rates the later were supposed to start with fixed (i.e. for all $n \in \mathbf{N}$ the same) values for the "first" states 0,1 a.s.o., correspondingly in the decreasing case one has to start for any $n \in \mathbf{N}$ with these values from behind (i.e. from state n). Thus, for sufficiently large n, in the former case the rates of a (fixed) state remain unaltered if we change n, whereas in the later case they vary with n.

REFERENCES

[1] Chung, K. L., *Markov chains with stationary transition probabilities,* Second edition, Die Grundlehren der mathematischen Wissenschaften, Bd. 104, Springer-Verlag New York, Inc., New York, 1967. *MR* **36** # 961.

[2] Eberl, W., Jr., Die asymptotische Verteilung der Absorptionszeit zweier Klassen von Geburts- und Todesprozessen, *Operations Research Verfahren* Bd. XII (1972), 78—97.

[3] Eberl, W., Jr., Die asymptotische Absorptionszeitverteilung bei Geburts- und Todesprozessen mit linear ansteigenden Geburts- und Sterberaten, *Sechste Oberwolfach-Tagung über Operations Research* (1973), Teil I, *Operations Research Verfahren,* Bd. XVIII, Hain, Meisenheim am Glan, 1974, 57—77. *MR* **56** # 6901.

[4] Eberl, W., Jr. Asymptotic absorption distributions in birth and death processes, Preprint Series, Nr. 4, Department of Mathematics, Fernuniversität Hagen, 1979, 1—27.

[5] Heathcote, C. R. and Moyal, J. E., The random walk (in continuous time) and its application to the theory of queues, *Biometrika* **46** (1959), 400—411. *MR* **22** # 3033.

[6] Kendall, D. G., Stochastic processes and population growth, *J. Roy. Statist. Soc. Ser. B* **11** (1949), 230—264. *MR* **11**—672.

[7] Ledermann, W. and Reuter, G. E. H., Spectral theory for the differential equations of simple birth and death processes, *Philos. Trans. Roy. Soc. London Ser. A.* **246** (1954), 321—369. *MR* **15**—625.

[8] Saaty, T. L., Some stochastic processes with absorbing barriers, *J. Roy. Statist. Soc. Ser. B* **23** (1961), 319—334. *MR* **25** # 1578.

[9] Schmidt, W., Die Verteilung der Adsorptionszeit und ihre Schätzung beim Markoffschen Prozeß mit einem reflektierenden und einem adsorbierenden Rand, Thesis, Göttingen, 1966.

(Received November 26, 1981)

FACHBEREICH MATHEMATIK
LEHRSTUHL FÜR WAHRSCHEINLICHKEITSTHEORIE
UND STATISTIK
FERNUNIVERSITÄT HAGEN
POSTFACH 940
LÜTZOWSTRASSE 125
D—5800 HAGEN 1
FEDERAL REPUBLIC OF GERMANY

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 449—452

PACKING OF r-CONVEX DISCS

L. FEJES TÓTH

We define an r-convex domain as the intersection of circular discs of radius r. We shall denote a domain and its area with the same symbol, and prove the following

THEOREM. *If the r-convex discs $c_1, \ldots, c_n$ of perimeter $p_1, \ldots, p_n$ are packed into a convex polygon h with at most six sides then the density $d=(c_1+\ldots+c_n)/h$ of the discs in h satisfies the inequality*

$$\frac{1}{d} \geqq 1+\frac{p^3}{864rc}$$

where $p=(p_1+\ldots+p_n)/n$ is the average perimeter and $c=(c_1+\ldots+c_n)/n$ the average area of the discs.

As an example we consider the case when the discs are equal lenses, i.e. congruent copies of the intersection of two circles, say, of radius 1. The area of a lens of perimeter p is equal to $\frac{p}{2}-\sin\frac{p}{2}$. Thus

$$\frac{1}{d} \geqq 1+\frac{p^3}{864\left(\frac{p}{2}-\sin\frac{p}{2}\right)} > 1+\frac{p^3}{864\frac{1}{3!}\left(\frac{p}{2}\right)^3} = \frac{19}{18},$$

i.e. $d<18/19=0.9473\ldots$. On the other hand, it is known [1] that d can get arbitrarily close to $\sqrt{8}/3=0.9428\ldots$.

PROOF of the Theorem. We may assume that $r=1$. We also may assume that c_i $(i=1, \ldots, n)$ is the intersection of a finite number of circles of unit radius. Let a_k be a k-gon of minimal area circumscribed about $c=c_i$. Let $S_1, \ldots, S_k$ be the side-midpoints of a_k in their cyclic order, i.e. the points at which a_k touches c. Let s_j be the region bounded by the parts S_jS_{j+1} of the boundaries of c and a_k. Let l_j be the length of the arc S_jS_{j+1} of the boundary of c. If the closed arc S_jS_{j+1} does not contain a vertex of c then

$$s_j = \tan\frac{l_j}{2}-\frac{l_j}{2},$$

1980 *Mathematics Subject Classification.* Primary 52A45.
Key words and phrases. Convexity, packing.

otherwise

$$s_j \geqq \tan\frac{l_j}{2} - \frac{l_j}{2}.$$

This is obvious if S_j or S_{j+1} is a vertex of c but no vertex of c lies on the open arc S_jS_{j+1}. If the open arc $S_j S_{j+1}$ contains a vertex of c then the above inequality follows from the following

REMARK. If in the quadrangle $ABCD$ with $\sphericalangle B > 180°$ we fix the length of the sides AB and BC and the angles $\sphericalangle A$ and $\sphericalangle C$ then the area of the quadrangle is an increasing function of $\sphericalangle B$.

First we consider the case when $\sphericalangle ADB \leqq 90°$ and $\sphericalangle BDC \leqq 90°$. Now replacing $\sphericalangle B$ by a greater angle, the new quadrangle can be decomposed into triangles congruent to the triangles ABD and BCD and a quadrangle (Fig. 1).

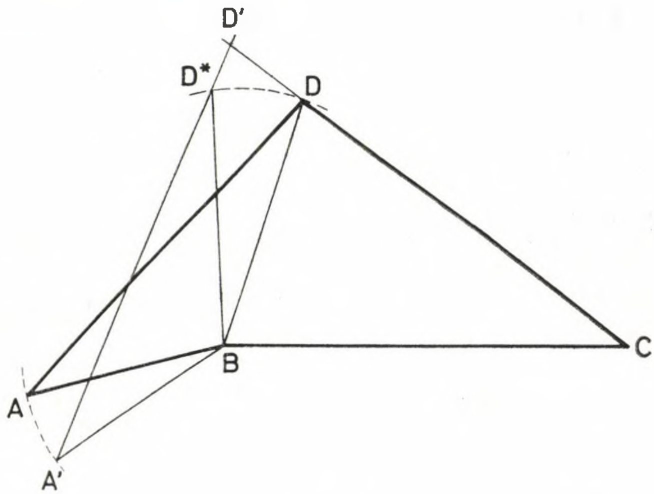


Fig. 1

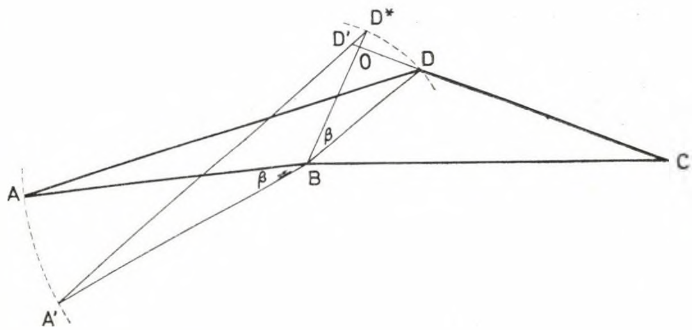


Fig. 2

Supposing now that, say, $\sphericalangle BDC > 90°$, we fix the side BC and increase $\sphericalangle B$ by an angle β obtaining the new quadrangle $A'BCD'$ (Fig. 2). Let the rotation which carries the side BA of $ABCD$ into the side BA' of $A'BCD'$ carry D into D^*. If β is sufficiently small the segments BD^* and DD' intersect one another in a point, say, O, and we have

$$A'BCD' - ABCD = A'BD^* + BCD + OBD - OD'D^* - ABD - BCD =$$

$$= OBD - OD'D^*.$$

But the order of magnitude of OBD is β while that of $OD'D^*$ is β^2. Thus for sufficiently small values of β we have $A'BCD' > ABCD$.

Now we assume that the open arc $S_j S_{j+1}$ contains one vertex V. We imagine the corresponding lines of the sides of a_k to be fixed at S_j and S_{j+1} to the arcs $S_j V$ and VS_{j+1}, and rotate these arcs about V so as to form one circular arc. Owing to the above Remark s_j decreases by this operation. If the open arc $S_j S_{j+1}$ contains several vertices of c, the above inequality can be seen by repeated application of this operation.

Now we have

$$a_k - c = \sum_{i=j}^{k} s_j \geqq \sum_{j=1}^{k} \left(\tan \frac{l_j}{2} - \frac{l_j}{2}\right) \geqq k\left(\tan \frac{p}{2k} - \frac{p}{2k}\right), \quad p = p_i.$$

Using the inequality $\tan x > x + \frac{x^3}{3}$ $(0 < x < \pi/2)$, we have

$$a_k - c > \frac{p^3}{24k^2}.$$

As the next step of the proof we use the known construction [2, 3, 4] of blowing up the discs $c_1, \ldots, c_n$ to non-overlapping convex polygons $q_1, \ldots, q_n$ of number of sides $k_1, \ldots, k_n$ such that $c_i \subset q_i \subset h$ and $k_1 + \ldots + k_n \leqq 6n$. Then we have

$$h \geqq q_1 + \ldots + q_n \geqq a_{k_1} + \ldots + a_{k_n} > c_1 + \ldots + c_n + \frac{1}{24}(p_1^3 k_1^{-2} + \ldots + p_n^3 k_n^{-2}).$$

Since the function $z = x^3 y^{-2}$ $(x > 0,\ y > 0)$ is, because of $z_{xx} > 0$ and $z_{xx} z_{yy} - z_{xy}^2 = 0$, convex, we can use Jensen's inequality obtaining

$$\frac{1}{d} = \frac{h}{c_1 + \ldots + c_n} > 1 + \frac{1}{24nc} np^3 \left(\frac{k_1 + \ldots + k_n}{n}\right)^{-2} \geqq 1 + \frac{p^3}{24c} 6^{-2} = 1 + \frac{p^3}{864c}.$$

REFERENCES

[1] FEJES TÓTH, L., Densest packing of lenses in the plane, *Mat. Lapok* **22** (1971), 209—213 (in Hungarian). *MR* **48** # 2081.

[2] FEJES TÓTH, L., Some packing and covering theorems, *Acta Sci. Math. Szeged* **12** (1950), Pars A, 62—67. *MR* **12**—352.

[3] FEJES TÓTH, L., *Lagerungen in der Ebene, auf der Kugel und im Raum*, Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung

der Anwendungsgebiete, Bd. 65, Springer-Verlag, Berlin—Göttingen—Heidelberg, 1953. *MR* **15**—248; Second edition: Springer-Verlag, Berlin—Heidelberg—New York, 1972. *MR* **50** # 5603

[4] Cassels, J. W. S., *An introduction to the geometry of numbers*, Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete, Bd. 99, Springer-Verlag, Berlin—Göttingen—Heidelberg, 1959. *MR* **28** # 1175.

(Received December 14, 1981)

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

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 453—456

A HEREDITARY RADICAL FOR NEAR-RINGS

MICHAEL HOLCOMBE

The four radicals $J_0(N)$, $D(N)$, $J_1(N)$, $J_2(N)$ of a near-ring N share many of the properties possessed by the Jacobson radical of a ring. For many purposes the radical $J_2(N)$ is the most important and has received study from several authors, in particular Betsch [1] and Laxton [7].

Kaarli [6] has shown that $J_2(N)$ is a hereditary radical in the sense that, for an ideal I of a near-ring N,

$$J_2(I) = I \cap J_2(N).$$

In this paper we study a radical, which we call $J_3(N)$, for a near-ring that possesses some interesting properties and is closely related to the radical $J_2(N)$. It was introduced in [5] where it was shown to be a special radical. Here we show that it is also a hereditary radical. This result has been announced in Pilz' book [8] without proof. If the near-ring N has an identity then $J_3(N)=J_2(N)$. There are good reasons to regard $J_3(N)$ as a more natural generalization of the Jacobson radical to near-rings than any of the others mentioned above. These reasons will be discussed later. All the near-rings in this note will be zero-symmetric [4].

Definition. Let N be a near-ring, an N-module Γ is of type 3 if and only if
(i) $\Gamma N \neq (0)$
(ii) Γ has no non-trivial N-subgroups
(iii) $\gamma n = \gamma' n$ for all $n \in N \Rightarrow \gamma = \gamma'$, where $\gamma, \gamma' \in \Gamma$.

For a given near-ring N the class of N-modules of type 3 is denoted by $M_3(N)$. The radical $J_3(N)$ of a near-ring N is defined by

$$J_3(N) = \bigcap_{\Gamma \in M_3(N)} (\Gamma)_r^N \quad \text{where} \quad (\Gamma)_r^N = \{n \in N \mid \Gamma n = 0\}.$$

If $M_3(N)$ is empty we define $J_3(N)=N$.

From the definition it is clear that $J_2(N) \subseteq J_3(N)$ and elementary examples exist to show that the inequality can be strict.

Lemma 1. *Let N be a near-ring and A a non-zero ideal of N. If Γ is an A-module of type* 3 *then Γ is also an N-module.*

Proof. Let $\gamma \in \Gamma$ with $\gamma \neq 0$, then either $\gamma A = \Gamma$ or $\gamma A = (0)$. If $\gamma A = (0)$ then $\gamma a = 0a$ for all $a \in A$ and so $\gamma = 0$. Therefore $\gamma A = \Gamma$. If $\gamma' \in \Gamma$ and $n \in N$ we define

1980 *Mathematics Subject Classification.* Primary 16A76; Secondary 16A21.
Key words and phrases. Near-ring, radical, hereditary radical.

$\gamma' n$ as follows. There exists an $a' \in A$ such that $\gamma' = \gamma a'$ and then we put $\gamma' n = \gamma(a'n)$. We now check that this operation is well-defined. Suppose that $\gamma' = \gamma a''$ for some $a'' \in A$, then $\gamma(a' - a'') = 0$. If we can show that $\gamma(a'n) = \gamma(a''n)$ for any $n \in N$ we will have succeeded. First we choose any $a \in A$, and then $na \in A$ and $(\gamma a')(na) = (\gamma a'')(na)$. Therefore $(\gamma(a'n))a = (\gamma(a''n))a$ for all $a \in A$ and thus $\gamma(a'n) = \gamma(a''n)$ as required. It is now immediate that Γ is an N-module under this operation.

LEMMA 2. *If A is a non-zero ideal of N and Γ is an A-module of type 3 then Γ is an N-module of type 3.*

PROOF. (i) $\Gamma N = (\gamma A)N = \gamma(AN) \supseteq \gamma(AA) = \Gamma A \neq (0)$.

(ii) Suppose that $\Delta \subseteq \Gamma$, $\Delta \neq (0)$ and Δ is an N-subgroup of Γ. Then $\Delta A \subseteq \Delta N \subseteq \Delta$ and so $\Delta = \Gamma$.

(iii) Let $\gamma', \gamma'' \in \Gamma$ and suppose that $\gamma' n = \gamma'' n$ for all $n \in N$. Let $a', a'' \in A$ such that $\gamma' = \gamma a'$, $\gamma'' = \gamma a''$ where γ is the same as the element γ chosen in Lemma 1.

Now $\gamma(a'n) = \gamma(a''n)$ for all $n \in N$, and, in particular, for any $a \in A$ $\gamma(a'a) = \gamma(a''a)$ and thus $(\gamma a')a = (\gamma a'')a$ for all $a \in A$. Therefore $\gamma a' = \gamma a''$. This means that Γ is an N-module of type 3.

LEMMA 3. *Suppose that Γ is an N-module of type 3 and I is an ideal of N such that $I \nsubseteq (\Gamma)_r^N$ where $(\Gamma)_r^N$ is the (right) annihilator of Γ in N. Then Γ is an I-module of type 3.*

PROOF. Clearly, Γ is an I-module and $\Gamma I \neq (0)$. Suppose that Δ is an I-subgroup of Γ, so that Δ is a subgroup of Γ and $\Delta I \subseteq \Delta$. Assume that there exists $\delta \in \Delta$ with $\delta \neq 0$. Now δI is an I-subgroup of Γ and $\delta IN \subseteq \delta I$ implying that δI is an N-subgroup of Γ. Therefore either $\delta I = (0)$ or $\delta I = \Gamma$. If $\delta I = \Gamma$ then $\Delta = \Gamma$, so we assume that $\delta I = (0)$, thus $\delta NI = (0)$. Since $\delta \neq 0$, $\delta N = \Gamma$ and so $\Gamma I = (0)$, a contradiction. Therefore $\Delta = \Gamma$ and Γ has no non-trivial I-subgroups. Now suppose that $\gamma', \gamma'' \in \Gamma$ and $\gamma' i = \gamma'' i$ for all $i \in I$. There exists $e \in I$ such that $\gamma' = \gamma' e$ for otherwise $\gamma' I = (0)$ and $\gamma' NI = \Gamma I = (0)$. We first show that $\gamma'' = \gamma'' e$. Let $n \in N$ with $\gamma' n = 0$ and suppose that $\gamma'' n \neq 0$. Then $\gamma'' nI = \Gamma$ and $\gamma' nI = 0$. Suppose that $i \in I$ and $\gamma'' ni \neq 0$ then $\gamma'' ni \neq \gamma' ni$ which is a contradiction since $ni \in I$. Therefore $\gamma'' n = 0$ and so the right annihilators in N of γ' and γ'' are equal. Since $\gamma'(n - en) = 0$ for all $n \in N$, $\gamma''(n - en) = 0$, i.e. $\gamma'' n = \gamma'' en$ for all $n \in N$ and thus $\gamma'' = \gamma'' e$.

Now for any $n' \in N$, $\gamma' n' - \gamma'' n' = \gamma' en' - \gamma'' en' = 0$ since $en' \in I$. Therefore $\gamma' = \gamma''$ since Γ is of type 3 as an N-module and the result follows.

These results lead to the main theorem:

THEOREM 1.[1] *Let N be a near-ring and I an ideal of N then*

$$J_3(I) = J_3(N) \cap I.$$

PROOF. $J_3(N) = \bigcap_{\Gamma \in M_3(N)} (\Gamma)_r^N$,

$$J_3(I) = \bigcap_{\Gamma \in M_3(I)} (\Gamma)_r^I.$$

[1] Professor R. Wiegandt has pointed out that this result, and the previous three Lemmas also hold under the weaker assumption that I is an invariant subnear-ring of N.

Let $\Gamma \in M_3(I)$, by Lemma 2. $\Gamma \in M_3(N)$ since $(\Gamma)_r^I = (\Gamma)_r^N \cap I$ we obtain $J_3(N) \cap \cap I \subseteqq J_3(I)$. Denote by $M_3^*(N)$ the class of N-modules Γ from $M_3(N)$ with the property that $\Gamma I \neq (0)$. For each $\Gamma' \in M^*(N)$ we have, by Lemma 3, $\Gamma' \in M_3(I)$. Therefore

$$J_3(I) \subseteqq I \cap \bigcap_{\Gamma' \in M_3^*(N)} (\Gamma')_r^N .$$

Since

$$J_3(I) = \bigcap_{\Gamma' \in M_3^*(N)} (\Gamma')_r^N \cap \bigcap_{\substack{\Gamma'' \in M_3(N) \\ \text{s.t. } \Gamma'' I = (0)}} (\Gamma'')_r^N$$

and

$$I \cap \bigcap_{\substack{\Gamma'' \in M_3(N) \\ \text{s.t. } \Gamma'' I = (0)}} (\Gamma'')_r^N = I$$

we have $J_3(I) \subseteqq I \cap J_3(N)$ and the result follows.

COROLLARY 1. *For any near-ring N,* $J_3(J_3(N)) = J_3(N)$.

This is an important property of all ring-theoretic radicals but has not been established before for any near-ring radical.

COROLLARY 2. *If N has a multiplicative identity* $J_1(N) = J_2(N) = J_3(N)$.

PROOF. We need just remark that if 1 is the identity and Γ is an N-module of type 2 then $\gamma 1 = \gamma$ for all $\gamma \in \Gamma$. Consequently, when $\gamma n = \gamma' n$ for all $n \in N$ then in particular $\gamma 1 = \gamma' 1$ and so $\gamma = \gamma'$. Therefore Γ will be of type 3. The rest is a standard result. All the other basic properties of the radical J_2 that are mentioned in [1] are easily verified for the case J_3.

REMARK. In [3] the Jacobson-type radicals $J_2^G(A)$, $J_0^G(A)$, $D^G(A)$ of a G-near-algebra[2] were introduced and compared with the radicals of the underlying near-rings. If A had an identity then it was shown that $J_2(A) = J_2^G(A)$, however, in general we were only able to show that $J_2^G(A) \subseteq J_2(A)$. By a simple adaptation of the proof it is possible to prove that $J_3^G(A) = J_3(A)$ for any G-near-algebra A, (where $J_3^G(A)$ is defined in the obvious way). This completely generalizes a well-known theorem in ring theory.

The Jacobson radical of a ring is an example of a *special radical property* (see [2], p. 138). We show in [5] that the radical $J_3(N)$ is also special in a similar way. These results lead us to the conclusion that the radical $J_3(N)$ is a very natural generalization of the Jacobson radical for rings.

REFERENCES

[1] BETSCH, G., Ein Radikal für Fastringe, *Math. Z.* **78** (1962), 86—90. *MR* **25** # 3068.

[2] DIVINSKY, N. J., *Rings and radicals*, Mathematical Expositions, No. 14, University of Toronto Press, Toronto, Ont.; George Allen and Unwin, London, 1965. *MR* **33** # 5654.

[3] HOLCOMBE, M., Representations of 2-primitive near-rings, and the theory of near-algebras, *Proc. Roy. Irish Acad. Sect. A* **73** (1973), 169—177. *MR* **48** # 2195.

[2] G-near-algebras arise very naturally in the theory of the representations of transformation near-rings.

[4] HOLCOMBE, M., Categorical representations of endomorphism near-rings, *J. London Math. Soc.* (2) **16** (1977), 21—37. *MR* **57** #3197.
[5] HOLCOMBE, M. and WALKER, R., Radicals in categories, *Proc. Edinburgh Math. Soc.* (2) **21** (1978/79), 111—128. *MR* **80b**: 18009.
[6] KAARLI, K., Radicals in near-rings, *Tartu Riikl. Ül. Toimetised Vih.* **390** (1976), 134—171. *MR* **57** #9760.
[7] LAXTON, R. R., A radical and its theory for distributively generated near-rings, *J. London Math. Soc.* **38** (1963), 40—49. *MR* **26** #3742.
[8] PILZ, G., *Near-rings. The theory and its applications,* North-Holland Mathematics Studies, No. 23, North-Holland Publishing Co., Amsterdam—New York—Oxford, 1977. *MR* **57** #9761.

(Received January 13, 1982)

DEPARTMENT OF PURE MATHEMATICS
THE QUEEN'S UNIVERSITY OF BELFAST
BELFAST BT7 1NN
NORTHERN IRELAND

Studia Scientiarum Mathematicarum Hungarica 17 (1982), 457—458

INTEGRAL CLOSURE AND VALUATION RINGS WITH ZERO-DIVISORS

J. GRÄTER

In this note a ring means a commutative ring with identity. All subrings of a ring R have the same identity as R. If Γ is a totally ordered abelian group written multiplicatively, then a mapping v from a ring R into $\hat{\Gamma}=\Gamma\cup\{0\}$ is a B-valuation with value group Γ, if for all $x, y\in R$

(1) $v(xy)=v(x)v(y)$

(2) $v(x+y)\leqq\max\{v(x), v(y)\}$

(3) $v(1)=1$ and $v(0)=0$.

If v is surjective, then v is called an M-valuation (compare with Bourbaki [1] and Manis [7]). A subring A of the ring R is a B- resp. M-valuation ring if there is a B- resp. M-valuation v of R with $A=\{x|x\in R\wedge v(x)\leqq 1\}$. Huckaba gives in [5] a proof for the following theorem due to Samuel and Griffin when R is the total quotient ring of A. The proof given in this paper is shorter than Huckaba's and R is not assumed to be the total quotient ring of A.

THEOREM. *Let A be a subring of the ring R. The integral closure $\bar{A}$ of A in R is the intersection of all B-valuation rings containing A.*

PROOF. Denote the intersection of all B-valuation rings containing A by S. $\bar{A}\subseteqq S$ is well-known and it remains to show that $S\subseteqq\bar{A}$ or $R\setminus\bar{A}\subseteqq R\setminus S$. Let $a\in R$ but a not integral over A. It is easily seen that $I=\{x|x\in R\wedge\exists n\in\mathbf{N}: a^n x=0\}$ is an ideal of R and $a\notin I$ as a is not integral over A. For each $x\in R$, $\tilde{x}$ means the canonical image of x in $\tilde{R}=R/I$. It is clear that $\tilde{a}$ is not a zero-divisor in $\tilde{R}$. Since a is not integral over A, $\tilde{a}$ is not integral over $\tilde{A}=(A+I)/I$. $\tilde{A}[\tilde{a}^{-1}]$ is a subring of $\tilde{R}[\tilde{a}^{-1}]$ and $\tilde{a}^{-1}\tilde{A}[\tilde{a}^{-1}]$ is a proper ideal of $\tilde{A}[\tilde{a}^{-1}]$, since $\tilde{a}$ is not integral over $\tilde{A}$. Thus there is a proper prime ideal P of $\tilde{A}[\tilde{a}^{-1}]$ with $\tilde{a}^{-1}\in P$ and by Manis [7], Proposition 1 there exists an M-valuation v of $\tilde{R}[\tilde{a}^{-1}]$ with $v(\tilde{a})>1$ and $v(\tilde{A})\leqq 1$. The restriction of v to $\tilde{R}$ induces a B-valuation w of R with $w(a)>1$ and $w(A)\leqq 1$. This completes the proof.

Now we may ask, under which circumstances the theorem holds with M-valuation rings instead of B-valuation rings. By definition (see Griffin [3]) a ring R has a large Jacobson radical J if any prime ideal containing J is maximal (J is the intersection of the maximal ideals of R). This is equivalent to the condition that for each $a\in R$ there exists $b\in R$ such that $a+b$ is a unit and $ab\in J$. Examples of such rings are rings in which every prime ideal is maximal and rings with only a finite number of

---

1980 *Mathematics Subject Classification.* Primary 13A18; Secondary 13B20.
Key words and phrases. Valuation rings, integral closure.

maximal ideals. These rings are of interest concerning our question, since the following is valid:

THEOREM. *Each B-valuation ring of a ring with large Jacobson radical is an M-valuation ring.*

PROOF. Let v be a B-valuation of R and A the B-valuation ring belonging to v. Suppose that $x\in R\setminus A$ and $y\in R$ are such that $x+y$ is invertible and $xy\in J$. By Manis [7], Proposition 1 there exists $z\in R$ with $v(xz)=1$. If $v(x+y)=v(x)$, take $z=(x+y)^{-1}$. If $v(x)<v(x+y)=v(y)$, then $x+y(x+y)^{-1}$ is invertible and $v\left(x(x+y(x+y)^{-1})^{-1}\right)=1$ is valid. In the case $v(x+y)<v(x)$, $x+y^2$ is invertible and $v(x)<v(x+y^2)=v(y^2)$. Take $z=\left(x+y^2(x+y^2)^{-1}\right)^{-1}$.

Altogether we have proved the

PROPOSITION. *Let R be a ring with large Jacobson radical. For each subring S of R the integral closure of S in R is the intersection of all M-valuation rings containing S.*

Proposition 9 of [2] and Proposition 2 of [5] are corollaries of the above proposition.

REFERENCES

[1] BOURBAKI, N., *Elements of mathematics. Commutative algebra,* Hermann, Paris; Addison—Wesley Publishing Co., Reading, Mass., 1972. *MR* **50** # 12997.
[2] GRIFFIN, M., Prüfer rings with zero divisors, *J. Reine Angew. Math.* **239/240** (1969), 55—67. *MR* **41** # 188.
[3] GRIFFIN, M. P., Valuations and Prüfer rings, *Canad. J. Math.* **26** (1974), 412—429. *MR* **49** # 10685.
[4] GRIFFIN, M. P., Multiplication rings via their total quotient rings, *Canad. J. Math.* **26** (1974), 430—449. *MR* **49** # 4993.
[5] HUCKABA, J., On valuation rings that contain zero divisors, *Proc. Amer. Math. Soc.* **40** (1973), 9—15. *MR* **47** # 6683.
[6] MANIS, M. E., Extension of valuation theory, *Bull. Amer. Math. Soc.* **73** (1967), 735—736. *MR* **36** # 1436.
[7] MANIS, M. E., Valuations on a commutative ring, *Proc. Amer. Math. Soc.* **20** (1969), 193—198. *MR* **38** # 2134.
[8] KRULL, W., Allgemeine Bewertungstheorie, *J. Reine Angew. Math.* **167** (1932), 160—196. *Zbl.* **4**, 98.

(Received January 25, 1982)

INSTITUT FÜR ALGEBRA UND ZAHLENTHEORIE
TECHNISCHE UNIVERSITÄT
D—33 BRAUNSCHWEIG
FEDERAL REPUBLIC OF GERMANY

BOOK REVIEW

Béla Andrásfai, Gráfelmélet; Folyamok, mátrixok (Graph Theory; Flows, Matrices, in Hungarian), Akadémiai Kiadó, Budapest, 1983, 262 pages. ISBN 963 05 3146 1.

The present book is a continuation of an earlier one by the same author (*Ismerkedés a gráfelmélettel,* in Hungarian, Budapest, 1971; English translation: *Introductory Graph Theory,* New York—Budapest, 1977). The author's approach is didactic, and special attention is devoted to the practical applications. The PERT-method, the transportation problem, maximal flow in a network, shortest route problem, linear electrical networks are treated, among others. Algorithmic problems of the solving methods are also considered. There are about 100 exercises together with complete solutions (at the end of the book), and 81 references (including works of J. Egerváry, T. Gallai, L. Kantorovitch, D. König, H. W. Kuhn, L. Lovász, O. Ore, H. Poincaré, H. J. Ryser). The printing work of Szegedi Nyomda is excellent as usual.

S. Lajos

PRINTED IN HUNGARY
Szegedi Nyomda, Szeged

RECENTLY ACCEPTED PAPERS

DAS, A. G., Convergence in u-second variation and RS_u integrals
KOGANOV, L. M., Операция факторизации на гиперграфах
HORVÁTH, J. and TEMESVÁRI, Á., Über Dichte und Enge von doppelgitterförmigen zweifachen Kreispackungen
TERVO, J., On the properties of k-coercive linear partial differential operators
HIEN, B. H. and SAIN, I., Category theoretic notions of ultraproducts
DALE, A. I., Some results in partial exchangeability
MEIMARIDOU-KOKKOU, A., On the tangency of multifunctions
MAKAI JR, E. and PACH, J., Controlling function classes and covering euclidean spaces
COOK, R. J., Small values of indefinite quadratic forms and polynomials in many variables
VEIDINGER, L., On the order of convergence of a finite element method for mixed boundary value problems
BALOG, A. and SÁRKÖZY, A., On sums of integers having small prime factors I
FEJES TÓTH, G., Packing and covering with incongruent circles
BIHARI, I., Lower estimates of the zeros of Bôcher's pairs with respect to equation $y''+p(x)y=0$
BIHARI, I., Zeros of the Bôcher function and its derivative with respect to differential equation $y''+p(x)y=0$ II
PETZ, D., A mean ergodic theorem with a look at martingales
BOLLE, U., Über die Dichte mehrfacher gitterförmiger Kreisanordnungen in der Ebene
TEMESVÁRI, Á. H., Die dünnste gitterförmige 5-fache Kreisüberdeckung der Ebene
KALLENBERG, O., Norm conditions for the path regularity of random processes
BALOG, A. and SÁRKÖZY, A., On sums of integers having small prime factors II
JÁRAI, A., A remark to a paper of J. Aczél and J. K. Chung
PINTZ, J., On primes in short intervals II
TOTIK, V., On a problem concerning L^p moduli of smoothness
FEJES TÓTH, L., Compact packing of circles
TEMESVÁRI, Á. H., Die dünnste 2-fache doppelgitterförmige Kreisüberdeckung der euklidischen Ebene
PAŁKA, Z., On the degree of vertices in a bichromatic random graph
MATOLCSI, T., Spectral properties of vector operators
PIXLEY, A. F., Some remarks on the two discriminators
GŁAZEK, K. and MICHALSKI, J., On polyadic groups which are term-derived from groups
DAY, A. and PICKERING, D., A note on the Arguesian lattice identity
GUESSARIAN, I., Applications of universal algebra to computer science
NIKOLOVA, T. S. and BAINOV, D. D., Краевая задача с параметрами для дифференциальных уравнений сверхнейтрального типа
DETTE, W. and MEIER, J., Bemerkungen zu einer Arbeit von J. Pintz
KEANE, W. J., Intersections of imbedded subgroups in abelian p-groups
LANG, ZS., Horocycles of a dynamical system on the plane
SORJONEN, P., Lattice-theoretical characterizations of inner product spaces
GRACZYŃSKA, E., On bases for normal identities
WSCHEBOR, M., Sur les intersections des surfaces aléatoires avec des hyperplans
BIELAK, H. and SYSŁO, M., Peripheral vertices in graphs
BIHARI, I., Asymptotic result concerning equation $x''|x'|^{n-1}+a(t)x^{n*}=0$
KOVÁCS, K., The characterization of complex valued additive functions
ALPÁR, L., Tauberian theorems for power series of two variables

CONTENTS

HU ISSN 0081—6906 Index: 26 748