Studia Scientiarum Mathematicarum Hungarica

VOLUME 29
NUMBERS 1–2
1994

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. It 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-1117 Budapest, Prielle Kornélia u. 19-35

Manuscripts and editorial correspondence should be addressed to

J. Merza
Managing Editor

P.O. Box 127
H-1364 Budapest

Tel.: (36)(1) 118-2875 Fax: (36)(1) 117-7166
e-mail: h3299mer @ ella.hu

Subscription information

Subscription price for Volume 29 (1994) in 4 issues: $ 92.00 including normal postage, airmail delivery plus $ 20.00

Orders should be addressed to

AKADÉMIAI KIADÓ
P.O.Box 245
H-1519 Budapest

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 1–8

ON PROXIMALLY FINE CONTIGUITIES

J. DEÁK

Abstract

We investigate some properties of the finest contiguity compatible with a proximity (in the sense of Čech). Counterexamples show that most of the analogous statements are false for Riesz or Lodato proximities. The paper ends up with an open problem.

Given a proximity δ in the sense of Čech [1] (respectively a Riesz or Lodato proximity), let $\Gamma^1(\delta)$ (respectively $\Gamma^1_R(\delta)$ or $\Gamma^1_L(\delta)$) denote the finest contiguity (Riesz or Lodato contiguity) compatible with δ; we aim at investigating the properties of the functors Γ^1, Γ^1_R and Γ^1_L. The two statements labelled "Theorem" answer questions raised by Prof. Á. Császár.

§ 0. Preliminaries

Let us recall some definitions and simple facts (see e.g. [2] § 0, the references given in [2], and [3] § 6).

0.1 A symmetric relation δ between subsets of X is a *proximity* on X if (i) $A\,\delta\,X$ implies $A\neq\emptyset$; (ii) $A\cap B\neq\emptyset$ implies $A\,\delta\,B$; (iii) $A\,\delta\,(B\cup C)$ iff either $A\,\delta\,B$ or $A\,\delta\,C$. The proximity δ is *Riesz* if $A\,\overline{\delta}\,B$ implies $c(A)\cap c(B)=\emptyset$, where $\overline{\delta}$ means that δ does not hold, and $c(A)=\{x:\{x\}\,\delta\,A\}$; it is *Lodato* if $A\,\overline{\delta}\,B$ implies $c(A)\,\overline{\delta}\,c(B)$. If δ is Lodato then c is the closure in a topology, which is called the topology induced by δ.

A filter $\mathfrak{f}$ on X is *δ-compressed* if $A\,\delta\,B$ whenever $A,B\in\sec\mathfrak{f}$, where $\sec\mathfrak{f}=\{S\subset X: S\cap F\neq\emptyset\ (F\in\mathfrak{f})\}$. δ' is *finer* than δ if $\delta'\subset\delta$. The supremum $\sup\limits_i\delta_i$ (with respect to this relation "finer") of the proximities δ_i does exist. The supremum of Riesz/Lodato proximities has the same property. A cover

---

1980 *Mathematics Subject Classification* (1985 Revision). Primary 54E15; Secondary 54E05.

Key words and phrases. (Riesz/Lodato) proximity, (Riesz/Lodato) contiguity, proximally fine contiguity, compressed/Cauchy filter.

Research supported by the Hungarian National Foundation for Scientific Research Grant No. 1807.

0081–6906/94/$ 4.00 ©1994 *Akadémiai Kiadó, Budapest*

$\mathfrak{c}$ of X is a *δ-cover* if for any sets A and B with $A\,\delta\,B$, there is a $C \in \mathfrak{c}$ that meets both A and B.

0.2 A *contiguity* on X is a non-empty collection Γ of finite covers of X such that (i) if $\mathfrak{c} \in \Gamma$ and $\mathfrak{c}$ refines $\mathfrak{d}$ then $\mathfrak{d} \in \Gamma$; (ii) if $\mathfrak{c}, \mathfrak{d} \in \Gamma$ then there is an element of Γ that refines both $\mathfrak{c}$ and $\mathfrak{d}$. ($\mathfrak{c}$ refines $\mathfrak{d}$ if each element of $\mathfrak{c}$ is contained in some element of $\mathfrak{d}$.) $\mathfrak{B} \subset \Gamma$ is a *base* for Γ if each element of Γ is refined by some element of $\mathfrak{B}$; it is a *subbase* for Γ if

$$\{(\bigcap)\mathfrak{F}\colon \emptyset \neq F \subset \mathfrak{B},\ \mathfrak{F} \text{ is finite}\}$$

is a base for Γ; here $(\bigcap)\mathfrak{F}$ is defined as follows: $A \in (\bigcap)\mathfrak{F}$ iff there are $A(\mathfrak{c}) \in \mathfrak{c}$ such that $A = \bigcap\{A(\mathfrak{c})\colon \mathfrak{c} \in \mathfrak{F}\}$; we shall write $\mathfrak{c}(\cap)\mathfrak{d}$ for $(\bigcap)\{\mathfrak{c}, \mathfrak{d}\}$. Γ' is *finer* than Γ if $\Gamma' \supset \Gamma$. The contiguity Γ *induces* the proximity $\delta = \delta(\Gamma)$ for which $A\,\delta\,B$ iff, for each $\mathfrak{c} \in \Gamma$, there is a $C \in \mathfrak{c}$ with $A \cap C \neq \emptyset \neq B \cap C$; in other words, Γ is *compatible* with δ.

The contiguity Γ is *Riesz* if, for each $\mathfrak{c} \in \Gamma$, $\operatorname{int}\mathfrak{c} = \{\operatorname{int} C\colon C \in \mathfrak{c}\}$ is a cover, where $\operatorname{int} C = X \setminus c(X \setminus C)$, with c defined for $\delta(\Gamma)$; Γ is *Lodato* if $\operatorname{int}\mathfrak{c} \in \Gamma$ whenever $\mathfrak{c} \in \Gamma$. A Riesz/Lodato contiguity induces a Riesz/Lodato proximity. The supremum $\sup \Gamma_i$ (with respect of the relation "finer") of the contiguities Γ_i does exist, and, assuming $i \in I \neq \emptyset$, $\bigcup\limits_i \Gamma_i$ is a subbase for it. The supremum of Riesz/Lodato contiguities has the same property. A filter $\mathfrak{f}$ on X is *Γ-Cauchy* if $\mathfrak{f} \cap \mathfrak{c} \neq \emptyset$ $(\mathfrak{c} \in \Gamma)$. Any Γ-Cauchy filter is $\delta(\Gamma)$-compressed.

0.3 Any proximity δ can be induced by contiguities; $\Gamma^0(\delta)$ is the coarsest and $\Gamma^1(\delta)$ the finest one, where the δ-covers of cardinality $\leqq 2$, i.e. the covers

$$\mathfrak{c}_{A,B} = \{X \setminus A,\ X \setminus B\} \qquad (A\,\overline{\delta}\,B),$$

form a subbase for $\Gamma^0(\delta)$, while $\Gamma^1(\delta)$ consists of all the finite δ-covers. If δ is Riesz or Lodato then so is $\Gamma^0(\delta)$; the finest compatible Riesz contiguity $\Gamma^1_R(\delta)$, respectively Lodato contiguity $\Gamma^1_L(\delta)$, can be described as follows: $\mathfrak{c} \in \Gamma^1_R(\delta)$ iff $\mathfrak{c}$ is a finite δ-cover and $\operatorname{int}\mathfrak{c}$ is a cover; $\mathfrak{c} \in \Gamma^1_L(\delta)$ iff $\mathfrak{c}$ is a finite cover and $\operatorname{int}\mathfrak{c}$ is a δ-cover. Γ^0, Γ^1, Γ^1_R and Γ^1_L are functors from the category of (Riesz/Lodato) proximities into the category of (Riesz/Lodato) contiguities; this simple fact is, however, irrelevant from the point of view of the present paper.

A cover $\mathfrak{c}$ of X will be called *strong* if the collection $\{X \setminus C\colon C \in \mathfrak{c}\}$ is disjoint. If $|\mathfrak{c}| \leqq 2$ then $\mathfrak{c}$ is evidently strong. If $\mathfrak{c}$ is strong and $|\mathfrak{c}| \geqq 3$ then $\mathfrak{c}$ is a δ-cover for any proximity δ.

§ 1. Suprema of proximally fine contiguities

If δ' is finer than δ then $\Gamma^1(\delta')$ is evidently finer than $\Gamma^1(\delta)$. Similar statements hold for Γ^1_R and Γ^1_L: check that, with int' understood in the space (X,δ'), $\operatorname{int} C \subset \operatorname{int}' C$, implying that $\operatorname{int}\mathfrak{c}$ refines $\operatorname{int}'\mathfrak{c}$; thus if $\operatorname{int}\mathfrak{c}$ is a cover (or a δ-cover) then $\operatorname{int}'\mathfrak{c}$ is also a cover (respectively a δ-cover, hence a δ'-cover). It is the aim of this section to investigate whether or no the stronger statement holds that Γ^1, Γ^1_R and Γ^1_L commute with the supremum. We begin with characterizations of proximally fine contiguities.

1.1 Given a proximity δ on X, and a finite partition (= finite cover consisting of pairwise disjoint sets) $\mathfrak{p}$ of X, define

$$\mathfrak{c}(\mathfrak{p},\delta) = \{P \cup Q : P, Q \in \mathfrak{p},\ P\,\delta\,Q\}.$$

$\mathfrak{c}(\mathfrak{p},\delta)$ is a δ-cover, since if $A\,\delta\,B$ then $A \cap P\,\delta\,B \cap Q$ for some $P, Q \in \mathfrak{p}$, see condition (iii) in the definition of a proximity.

LEMMA. *If δ is a proximity then the collection*

$$\{\mathfrak{c}(\mathfrak{p},\delta) : \mathfrak{p} \text{ is a finite partition}\} \tag{1}$$

constitutes a base for $\Gamma^1(\delta)$.

PROOF. We have already seen that (1) consists of δ-covers. Conversely, given a finite δ-cover $\mathfrak{c}$, consider the partition $\mathfrak{p}$ generated by $\mathfrak{c}$, i.e. for which $A \in \mathfrak{p}$ iff $A = \bigcap_{C\in\mathfrak{c}} f(C)$ where $f(C) = C$ or $X \setminus C$. Now $\mathfrak{c}(\mathfrak{p},\delta)$ refines $\mathfrak{c}$, since if $P, Q \in \mathfrak{p}$ and $P\,\delta\,Q$ then, by the definition of a δ-cover, there is a $C \in \mathfrak{c}$ meeting both P and Q, which means that $P \cup Q \subset C$. □

1.2 LEMMA. *Given a proximity δ, the strong δ-covers of cardinality $\leqq 3$ form a subbase for $\Gamma^1(\delta)$.*

PROOF. We intend to apply Lemma 1.1. Let $\mathfrak{p}$ be a finite partition. If $|\mathfrak{p}| \leqq 2$ then $\mathfrak{c} = \mathfrak{c}(\mathfrak{p},\delta)$ is clearly a strong δ-cover of cardinality $\leqq 3$. Assume that $|\mathfrak{p}| \geqq 3$, and for $P, Q \in \mathfrak{p}$, $P \neq Q$, define

$$\mathfrak{d}_{P,Q} = \begin{cases} \mathfrak{c}_{P,Q} & \text{if } P\,\bar{\delta}\,Q, \\ \{X \setminus P, X \setminus Q, P \cup Q\} & \text{if } P\,\delta\,Q. \end{cases}$$

$\mathfrak{d}_{P,Q}$ is a strong δ-cover. Now

$$(\textstyle\bigcap)\{\mathfrak{d}_{P,Q} : P, Q \in \mathfrak{p},\ P \neq Q\}$$

refines $\mathfrak{c}$: we have to show that if $D_{P,Q} \in \mathfrak{d}_{P,Q}$ then

$$D = \bigcap\{D_{P,Q} : P, Q \in \mathfrak{p},\ P \neq Q\}$$

is contained by some element of $\mathfrak{c}$; D is the union of some elements of $\mathfrak{p}$, and if it contains distinct sets $P_0, Q_0 \in \mathfrak{p}$ then necessarily $\mathfrak{d}_{P_0,Q_0} = P_0 \cup Q_0$ and $P_0\,\delta\,Q_0$, thus $D = P_0 \cup Q_0 \in \mathfrak{c}(\mathfrak{p},\delta)$. □

It is not true that any contiguity has a subbase consisting of coverings of cardinality $\leqq 3$:

EXAMPLE. Let $|X| = 4$, $\mathfrak{c} = \{C \subset X : |C| = 3\}$, and let $\{\mathfrak{c}\}$ be a base for Γ. Now $\mathfrak{d} \in \Gamma$ iff either $X \in \mathfrak{d}$ or $\mathfrak{c} \subset \mathfrak{d}$, thus $\{\mathfrak{d} \in \Gamma : |\mathfrak{d}| \leqq 3\}$ is a (sub)base for the indiscrete contiguity on X. □

Let us observe for what it is worth that, on the other hand, any contiguity has a subbase consisting of strong covers, because it can be proved by induction on $|\mathfrak{c}|$ that if $\mathfrak{c}$ is a finite cover then there are $m \in \mathbb{N}$ and finite strong covers $\mathfrak{c}_1, \ldots, \mathfrak{c}_m$ such that $\mathfrak{c}$ refines each $\mathfrak{c}_j$, and $(\bigcap_1^m)\mathfrak{c}_j$ refines $\mathfrak{c}$:

Let $\mathfrak{c}$ be a finite cover, $|\mathfrak{c}| = n > 2$. Denote by $\mathfrak{p}$ the partition generated by $\mathfrak{c}$. Assign to each $P \in \mathfrak{p}$ different from $\bigcap\mathfrak{c}$ covers $\mathfrak{d}_P$ and $\mathfrak{e}_P$ as follows:

$$\mathfrak{d}_P = \{X \setminus P, \operatorname{St}(P, \mathfrak{c})\}, \quad \mathfrak{e}_P = \{C \cup (X \setminus \operatorname{St}(P, \mathfrak{c})) : P \subset C \in \mathfrak{c}\}.$$

It follows from $P \neq \bigcap\mathfrak{c}$ that $|\mathfrak{e}_P| < n$, thus covers $\mathfrak{e}_{P,1}, \ldots, \mathfrak{e}_{P,m(P)}$ can be chosen for $\mathfrak{e}_P$ according to the induction hypothesis. Let

$$\mathfrak{f}_{P,j} = \{E \cup P : E \in \mathfrak{e}_{P,j}\} \cup \{X \setminus P\}.$$

Now the covers $\mathfrak{d}_P$ and $\mathfrak{f}_{P,j}$ are the ones we were looking for.

It is clear that the covers are strong, and $\mathfrak{c}$ refines $\mathfrak{d}_P$. $\mathfrak{c}$ refines $\mathfrak{f}_{P,j}$, too, since if $C \in \mathfrak{c}$, $C \not\subset X \setminus P$ then $P \subset C$, thus C is a subset of an element of $\mathfrak{e}_P$, which refines $\mathfrak{e}_{P,j}$, and so $\mathfrak{f}_{P,j}$. Therefore we have only to prove that, given $D_P \in \mathfrak{d}_P$ and $F_{P,j} \in \mathfrak{f}_{P,j}$ for each P and j, the intersection of these sets is contained by some $C \in \mathfrak{c}$. If for each P either $D_P = X \setminus P$ or there is a j with $F_{P,j} = X \setminus P$ then the intersection is contained by $\bigcap\mathfrak{c}$. Otherwise, pick a P such that $D_P = \operatorname{St}(P, \mathfrak{c})$ and $F_{P,j} = E_{P,j} \cup P$ with some $E_{P,j} \in \mathfrak{e}_{P,j}$ $(1 \leqq j \leqq m(P))$. Now

$$\bigcap_j F_{P,j} = P \cup (\bigcap_j E_{P,j}) \subset P \cup C \cup (X \setminus \operatorname{St}(P, \mathfrak{c}))$$

with a $C \in \mathfrak{c}$ satisfying $P \subset C$. Hence $D_P \cap \bigcap_j F_{P,j} \subset C$.

1.3 LEMMA. *A contiguity is proximally fine iff it contains all the strong covers of cardinality* 3.

PROOF. *Necessity.* Any strong cover of cardinality 3 is a $\delta(\Gamma)$-cover, thus Lemma 1.2 can be applied.

Sufficiency. $\Gamma^0(\delta(\Gamma)) \subset \Gamma$, thus Γ contains all the (strong) $\delta(\Gamma)$-covers of cardinality $\leqq 2$, i.e. of cardinality $\leqq 3$ by the assumption; hence $\Gamma = \Gamma^1(\delta(\Gamma))$ follows from Lemma 1.2. □

1.4 LEMMA. a) *Any contiguity finer than a proximally fine contiguity is proximally fine itself.*

b) *If at least one of the contiguities* Γ_i *is proximally fine then so is* $\sup_i \Gamma_i$.

PROOF. a) follows from Lemma 1.3; b) is clear from a). □

THEOREM. *For a non-empty family of proximities* δ_i,

$$\sup_i \Gamma^1(\delta_i) = \Gamma^1(\sup_i \delta_i).$$

REMARK. This theorem is false for the empty family of proximities on a set of cardinality $\geqq 3$.

PROOF. It is well-known (and straightforward from the definitions) that the operation assigning $\delta(\Gamma)$ to Γ commutes with the supremum, thus $\Gamma = \sup_i \Gamma^1(\delta_i)$ is compatible with $\sup_i \delta(\Gamma^1(\delta_i)) = \sup_i \delta_i$; Γ is proximally fine by part b) of the lemma. □

1.5 In view of Lemma 1.2, the following conjectures sound plausible:

(i) given a Riesz proximity δ, the strong δ-covers $\mathfrak{c}$ of cardinality $\leqq 3$ for which int $\mathfrak{c}$ is a cover form a subbase for $\Gamma^1_R(\delta)$;

(ii) given a Lodato proximity δ, the strong open δ-covers of cardinality $\leqq 3$ form a subbase for $\Gamma^1_L(\delta)$.

These statements are false, even if the word "strong" is dropped:

EXAMPLE. Take the sets $S, T \in \{0,1\}^4$ for which $s \in S$ ($s \in T$) iff exactly one of the coordinates of s is equal to 1 (to 0). Let $U = S \cup T$, $X = U \times \mathrm{N}$, $G_k = U_k \times \mathrm{N}$ $(1 \leqq k \leqq 4)$, where U_k consists of those elements of U whose kth coordinate is 0. Consider on X the T_1-topology for which the sets G_k $(1 \leqq k \leqq 4)$ together with all the cofinite sets form an open subbase; denote by c the closure in this topology. Let δ be the finest Lodato proximity (= the finest Riesz proximity) compatible with c, i.e. $A\,\delta\,B$ iff $c(A) \cap c(B) \neq \emptyset$. All the finite open covers form a base for the finest Lodato contiguity (= the finest Riesz contiguity) Γ compatible with c. Clearly, $\Gamma = \Gamma^1_L(\delta) = \Gamma^1_R(\delta)$. Any open cover is now a δ-cover, thus if (i) or (ii) (with or without the word "strong") were true then the open covers of cardinality $\leqq 3$ would form a subbase for Γ; hence there would exist $n \in \mathrm{N}$ and open covers $\mathfrak{c}_j$ $(1 \leqq j \leqq n)$ such that $|\mathfrak{c}_j| \leqq 3$ and $(\bigcap_1^n)\mathfrak{c}_j$ refines $\mathfrak{c} = \{G_k : 1 \leqq k \leqq 4\}$. This assumption will lead to a contradiction.

Each $C \in \bigcup_1^n \mathfrak{c}_j$ can be written as $K_C \cap G_C$ where K_C is cofinite, and G_C is open in the topology $\mathcal{T}$ for which $\mathfrak{c}$ is an open subbase. Choose $m \in \mathrm{N}$ such that

$$U \times \{m\} \subset \bigcap \{K_C : C \in \bigcup_1^n \mathfrak{c}_j\}. \tag{1}$$

For each j, there is a $C_j \in \mathfrak{c}_j$ that meets $T \times \{m\}$ in at least two points; thus $|G_{C_j} \cap (T \times \{m\})| \geqq 2$. As any T-open set with this property contains $S \times \{m\}$, we have $C_j \supset S \times \{m\}$ from (1). Hence $S \times \{m\} \subset \bigcap_1^n C_j \in (\bigcap_1^n)\mathfrak{c}_j$, contradicting the assumption that $(\bigcap_1^n)\mathfrak{c}_j$ refines $\mathfrak{j}$. □

The above example furnishes in fact a little bit more: the open δ-covers of cardinality $\leqq 3$ together with all the finite strong open δ-covers may not constitute a subbase for $\Gamma^1_L(\delta)$: given a finite strong open cover, two arbitrary elements of it make up an open cover, which is in the example necessarily a δ-cover. On the other hand, it follows from the remark at the end of 1.2 that the finite strong δ-covers $\mathfrak{c}$ for which $\operatorname{int}\mathfrak{c}$ is a cover form a subbase for $\Gamma^1_R(\delta)$. (Again by the above example, we cannot require here that $\operatorname{int}\mathfrak{c}$ should be strong.)

1.6 It is now not surprising that the analogues of Theorem 1.4 are false for Γ^1_R and Γ^1_L, even when there are only two proximities:

EXAMPLE. Let $X = \mathbf{N}^2$, and $A\,\overline{\delta}_i\,B$ iff the ith projections of A and B are disjoint ($i = 1, 2$), $\delta = \sup\{\delta_1, \delta_2\}$. Take the cover

$$\mathfrak{c} = \Big\{\{(m,n) \in X : m \leqq n\}, \{(m,n) \in X : m \geqq n\}, \{(m,n) \in X : m \neq n\}\Big\}.$$

$\mathfrak{c}$ is clearly a δ-cover, and it is open, since δ induces the discrete topology. Thus $\mathfrak{c} \in \Gamma^1_L(\delta) \subset \Gamma^1_R(\delta)$. We are going to show that

$$(1) \qquad \mathfrak{c} \notin \sup\{\Gamma^1_R(\delta_1), \Gamma^1_R(\delta_2)\}.$$

This implies that $\mathfrak{c}$ does not belong to the coarser contiguity $\sup\{\Gamma^1_L(\delta_1), \Gamma^1_L(\delta_2)\}$ either.

Assume that (1) is false, and take finite δ_i-covers $\mathfrak{c}_i$ such that $\operatorname{int}_i \mathfrak{c}_i$ is a cover ($i = 1, 2$) and $\mathfrak{c}_1(\cap)\mathfrak{c}_2$ refines $\mathfrak{c}$. There is an infinite set $M \subset \mathbf{N}$ such that $M \times \{1\} \subset \operatorname{int} C_1$ for some $C_1 \in \mathfrak{c}_1$. $c(\delta_1)$ is the closure in the product of the discrete and the indiscrete topology on $\mathbf{N}$, thus $M \times \mathbf{N} \subset C_1$. Similarly, we can pick an infinite set $N \subset M$ and a $C_2 \in \mathfrak{c}_2$ such that $\mathbf{N} \times N \subset C_2$. Now N^2 has to be contained by some element of $\mathfrak{c}$, a contradiction. □

1.7 Let δ be a proximity on X, $f: Y \to X$; one may ask whether Γ^1, Γ^1_R and Γ^1_L commute with the inverse image, i.e. whether $f^{-1}\Gamma^1(\delta) = \Gamma^1(f^{-1}\delta)$, and similarly for Γ^1_R and Γ^1_L. (Let us recall that $A\, f^{-1}\delta\, B$ iff $f[A]\,\delta\, f[B]$, and $\{f^{-1}\mathfrak{c} : \mathfrak{c} \in \Gamma\}$ is a base for $f^{-1}\Gamma$ where $f^{-1}\mathfrak{c} = \{f^{-1}[C] : C \in \mathfrak{c}\}$.) The answer is much simpler than for the supremum: Γ^1 commutes with the inverse image under injections (i.e. with restriction to subsets), while Γ^1_L commutes with the inverse image under surjections; the proofs are straightforward. No other positive statements are valid:

EXAMPLES. a) If $|Y| = 3$, $|X| = 1$, $f\colon Y \to X$, and δ is the (unique) proximity on X then $f^{-1}\Gamma^1(\delta) \neq \Gamma^1(f^{-1}\delta)$.

b) If the topological space X is a convergent sequence, Y is the same without the limit point, $f\colon Y \to X$ is the identical embedding, and δ is the (unique) Efremovich proximity compatible with the topology of X then $f^{-1}\Gamma^1_R(\delta) \neq \Gamma^1_R(f^{-1}\delta)$ and $f^{-1}\Gamma^1_L(\delta) \neq \Gamma^1_L(f^{-1}\delta)$. Indeed, if A_i $(i = 1,2,3)$ are disjoint infinite subsets of Y then $\mathfrak{d} = \{X \setminus A_j\colon i = 1,2,3\} \in \Gamma^1_L(f^{-1}\delta)$, but $\mathfrak{d} \notin f^{-1}\Gamma^1_R(\delta)$. (Assume the contrary. Then $\mathfrak{d}$ is refined by $f^{-1}\mathfrak{c}$ for some $\mathfrak{c} \in \Gamma^1_R(\delta)$. There is a $C \in \mathfrak{c}$ with $X \setminus Y \subset \operatorname{int} C$; therefore C is cofinite, and so is an element of $f^{-1}\mathfrak{c}$, hence of $\mathfrak{d}$, a contradiction.)

c) Let $X = \mathbb{N}$, $Y = \mathbb{N} \times \{1,2\}$, $f\colon Y \to X$ the projection, and take the proximity δ on X for which $A\,\delta\,B$ iff $A \cap B \neq \emptyset$ or both A and B are infinite. Now with P denoting the set of the even numbers we have

$$\Big\{P \times \{1,2\}, Y \setminus (P \times \{1\}), Y \setminus (P \times \{2\})\Big\} \in \Gamma^1_R(f^{-1}\delta) \setminus f^{-1}\Gamma^1_R(\delta). \quad \square$$

REMARK. In the terminology of [3], Example b) can be interpreted as follows: $\Gamma^1_L(\delta\,|\,Y) = \Gamma^1_R(\delta\,|\,Y) = \Gamma^1(\delta\,|\,Y)$ is too fine to have a Riesz or Lodato extension, because the trace filter of the point in $X \setminus Y$ is not Cauchy.

§ 2. Unique compatible contiguities

THEOREM. *For a proximity δ, $\Gamma^0(\delta) = \Gamma^1(\delta)$ iff each δ-compressed filter is the intersection of at most two ultrafilters.*

REMARK. Any ultrafilter is compressed, and conversely, if δ is discrete (i.e. $A\,\delta\,B$ iff $A \cap B \neq \emptyset$) then any compressed filter is an ultrafilter. Thus the theorem states that, roughly speaking, if $\Gamma^0(\delta) = \Gamma^1(\delta)$ then δ is almost discrete.

PROOF. A contiguity is completely determined by the Cauchy filters (see [4] 2.10), thus $\Gamma^0(\delta) = \Gamma^1(\delta)$ iff each δ-compressed filter is $\Gamma^1(\delta)$-Cauchy (because the δ-compressed filters are the same as the $\Gamma^0(\delta)$-Cauchy filters, and any $\Gamma^1(\delta)$-Cauchy filter is clearly $\Gamma^0(\delta)$-Cauchy).

Consequently, it follows from Lemma 1.2 that $\Gamma^0(\delta) = \Gamma^1(\delta)$ iff for each δ-compressed filter $\mathfrak{f}$ and for each strong (δ-)cover $\mathfrak{c}$ of cardinality 3, $\mathfrak{f} \cap \mathfrak{c} \neq \emptyset$ (because it is enough to check the Cauchy property for a subbase, and if $\mathfrak{c}$ is a δ-cover of cardinality < 3 then $\mathfrak{f} \cap \mathfrak{c} \neq \emptyset$ follows from $\mathfrak{f}$ being δ-compressed). The proof can now be completed by the simple observation that a filter is the intersection of at most two ultrafilters iff there are no three disjoint sets in $\sec \mathfrak{f}$, i.e. iff $\mathfrak{f} \cap \mathfrak{c} \neq \emptyset$ for all the covers mentioned above. $\square$

PROBLEM. Characterize those Riesz/Lodato proximities for which there exists only one compatible Riesz/Lodato contiguity.

REFERENCES

[1] Čech, E., *Topological spaces*, Revised by Z. Frolík and M. Katětov, Academia, Prague, and Interscience, London, 1966. *MR* **35** #2254

[2] Császár, Á. and Deák, J., Simultaneous extensions of proximities, semi-uniformities, contiguities and merotopies, I, *Math. Pannonica* **1** (1990), No 2, 67–90. *MR* **92g**:54033

[3] Császár, Á. and Deák, J., Simultaneous extensions of proximities, semi-uniformities, contiguities and merotopies, III, *Math. Pannonica* **2** (1991), No 2, 3–23. *MR* **93d**:54039

[4] Katětov, M., On continuity structures and spaces of mappings, *Comment. Math. Univ. Carolinae* **6** (1965), No 2, 257–278. *MR* **33** #1826

(Received July 11, 1990)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H-1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 9–23

DENSEST BALL PACKINGS BY ORBITS OF THE 10 FIXED POINT FREE EUCLIDEAN SPACE GROUPS

Á. G. HORVÁTH and E. MOLNÁR

0. Introduction

Let G be a fixed point free Euclidean space group, i.e. any group from among the crystallographic groups

1. P1; 2. P2_1; 7. Pb; 9. Bb; 19. P$2_1 2_1 2_1$; 29. Pca2_1; 33. Pna2_1; 76. P4_1; 144. P3_1; 169. P6_1 (see [3], [7]).

Take a point X in the Euclidean space E^3, and consider the G-orbit of X

$$X^G := \{X^g \in E^3 : g \in G\}. \tag{0.1}$$

The radius $r(X^G)$ of the ball packing with centres by the orbit X^G is defined as follows:

$$r(X^G) = (1/2)\inf\{\varrho(X, X^g) : g \in G \setminus \{1\}\} \tag{0.2}$$

where ϱ is the distance function in E^3. We are interested in the optimal ball packing of the group G whose radius is

$$r(G) = \sup\{r(X^G) : X \in E^3\}. \tag{0.3}$$

The optimal density of the ball packing by G is

$$\delta(G) = \frac{\mathrm{Vol}\, B(r(G))}{\mathrm{Vol}\, F(G)} \tag{0.4}$$

where the volume of the optimal ball is related to the volume of the fundamental domain of G. This density $\delta(G)$ depends on the free parameters of G. Finally, we optimize $\delta(G)$ also by the free parameters of G. Thus the optimal density $\delta(G)$ will be determined by the isomorphy class of G as it is natural to expect.

This program generalizes the problem of finding the densest lattice-like ball packing in the Euclidean space E^3, where the group $G = \mathrm{P1}$ generated

1991 *Mathematics Subject Classification*. Primary 52C17; Secondary 52C07.
Key words and phrases. Packing, space group.

by 3 independent translations. The result, due to Gauss is well-known [1]. The optimal packing provides the face centred cubic lattice, spanned by the vectors $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3$ with Gramian matrix

$$(0.5) \qquad (f_{ik}) = (\langle \mathbf{f}_i, \mathbf{f}_k \rangle) = \begin{bmatrix} 1 & 1/2 & 1/2 \\ 1/2 & 1 & 1/2 \\ 1/2 & 1/2 & 1 \end{bmatrix}, \qquad \det(f_{ik}) = 1/2,$$

$$(0.6) \qquad r(\mathrm{P1}) = 1/2; \ \mathrm{Vol}\, F(\mathrm{P1}) = (1/2)^{1/2}; \ \delta(\mathrm{P1}) = \pi(18)^{-1/2} \approx 0.7405.$$

It is clear that congruent orbits of G provide ball packings of the same density. Therefore, those isometries of E^3, which preserve the G-orbits, play important role. These isometries constitute the metric normalizer of the space group G as a supergroup

$$(0.7) \qquad M_G = \{\mu \in \mathrm{Iso}\, E^3 \mid \mu^{-1} G \mu = G\}.$$

A fundamental set of M_G, denoted by $F(M_G)$ has the basic property

$$(0.8) \qquad r(G) = \sup\{r(X^G) \mid X \in F(M_G)\}$$

because $F(M_G)$ consists of points providing all noncongruent G-orbits. Since we know the metric normalizer for each group G considered [2], it is reasonable to assign a suitable $F(M_G)$ as a parameter domain for determining the optimal radius $r(G)$. It will turn out that each above group G provides the optimal ball packing of the same density by (6). For the first 9 groups one extremal arrangement is the same as the lattice-like one. The groups $\mathrm{P}2_12_12_1$; $\mathrm{P}2_1$; Bb; $\mathrm{Pna}2_1$ and $\mathrm{Pca}2_1$ has two different extremal arrangements, one of them is lattice-like. The other optimal ball packing of these five groups coincides with the unique one of $\mathrm{P}6_1$. This non-lattice arrangement is also well-known.

1. The method of the proof, the case $\mathrm{P}2_12_12_1$

In the International Tables [3], for each Euclidean space group G there is given its lattice L_G, related to a primitive Bravais lattice with the corresponding centering. The coordinate presentation of the so-called point-group G_0 is expressed in the basis of L_G. Moreover, a so-called vector-system, associated with the point group G_0 is given which describes the images of the origin O under G.

For instance let the group $\mathrm{P}2_12_12_1$ be fixed as G. The primitive orthorombic lattice L_G is spanned by the orthogonal basis $\{\mathbf{e}_i\}$ with Gramian

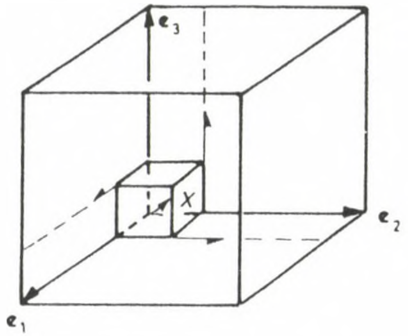
$$(1.1) \qquad (e_{ij}) = (\langle \mathbf{e}_i, \mathbf{e}_j \rangle) = \begin{bmatrix} a^2 & 0 & 0 \\ 0 & b^2 & 0 \\ 0 & 0 & c^2 \end{bmatrix}, \quad 0 < a \leqq b \leqq c, \ abc = 1.$$

The parameters a, b, c are characteristic for G, but we may normalize $abc = 1$.

The point positions are:

(1.2) $$X(x,y,z), X^{S1}(x+1/2, -y, -z+1/2), X^{S2}(-x+1/2, y+1/2, -z), \quad X^{S3}(-x, -y+1/2, z+1/2).$$

S_1, S_2, S_3 mean screw motions indicated in the Figure 1.



$X(x,y,z) \in F(M_G)$.

Fig. 1

The vectors XX^g, $g \in G$, are of the form

$$\mathbf{t}_0, XX^{S1} + \mathbf{t}_1, XX^{S2} + \mathbf{t}_2, XX^{S3} + \mathbf{t}_3 \quad \text{with} \quad \mathbf{t}_0, \mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3 \in L_G.$$

These are in coordinates:

(1.3) $$(0,0,0),\ (1/2, -2y, -2z+1/2),\ (-2x+1/2, 1/2, -2z),$$

plus an integer tripel to any of them, describing a lattice vector from L_G. Here $X(x,y,z)$ runs over a fundamental set of the metric normalizer M_G. Now $M_G = \text{Pmmm}$ is generated by plane reflections in the walls of the brick with edge measures $a/4, b/4, c/4$. The lattice L_{Pmmm} is generated by $(1/2)\mathbf{e}_1$, $(1/2)\mathbf{e}_2$, $(1/2)\mathbf{e}_3$ (see [2]). Hence a fundamental set of M_G is defined by

(1.4) $$F(M_G) = \{X(x,y,z) \in E^3,\ 0 \leqq x \leqq 1/4,\ 0 \leqq y \leqq 1/4,\ 0 \leqq z \leqq 1/4\}.$$

From (1.3) we see that the infimum by the formula (0.2) comes from the length minimum of 4 vectors as follows

(1.5) $$l = \min\Big\{a, [a^2/4 + 4y^2b^2 + (1/2-2z)^2c^2]^{1/2}, \quad [(1/2-2x)^2a^2 + b^2/4 + 4z^2c^2]^{1/2},\ [4x^2a^2 + (1/2-2y)^2b^2 + c^2/4]^{1/2}\Big\}.$$

By (0.3) and (0.4) we look for

(1.6) $$\delta(G) = (2\pi/3)\max\{l^3 \mid 0 \leqq x, y, z \leqq 1/4,\ 0 < a \leqq b \leqq c; abc = 1\}$$

since now Vol $F(G) = 1/4$, Vol $B(l/2) = (1/6)l^3\pi$. To prove $\delta(G) \leqq \pi/(18)^{1/2}$, we shall show that $l \leqq 2^{-1/2}$ in (1.5). So we may assume that

(1.7) $$2^{-1/2} \leqq a \leqq b \leqq c, \quad \text{moreover} \quad 0 \leqq x, y, z \leqq 1/4$$

hold in (1.5) and (1.6). We sketchily prove the following

LEMMA. *If a brick has a unit volume, its side lengths are not less than* $2^{-1/2}$ *and the vertices of a triangle lie on the skew edges of this brick, then there is a side of the triangle whose length is not greater than* $2^{1/2}$ (see Figure 2).

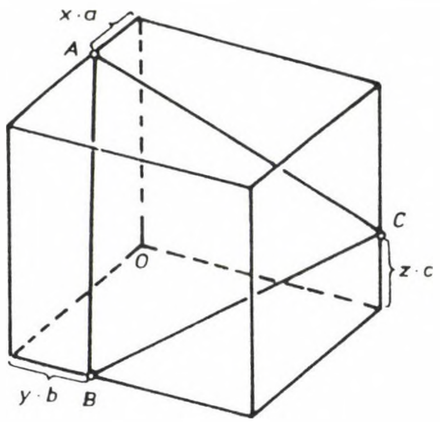


Fig. 2

We may assume that the triangle is regular and prove that the optimal triangle have two common vertices with the brick containing it. Consider the orthogonal projection of the triangle onto any face P of the brick. So we obtain a triangle which has a common vertex A' with the rectangle P and the other vertices B', C' lie on those sides of P which are opposite to A'. A short calculation shows that the indirect assumption, i.e. P has neither B' nor C' as its vertex, may allow choosing a bigger triangle ABC on the skew edges of the brick. So two vertices of the optimal regular triangle are vertices of the brick. From this we derive three optimal cases with the parameters:

1. $a^2 = 1/2, \quad b^2 = 1, \quad c^2 = 2;$
2. $a^2 = 1/2, \quad b^2 = 4/3, \quad c^2 = 3/2;$
3. $a^2 = 1, \quad b^2 = 1, \quad c^2 = 1$

(see Figure 3).

For more details see [4].

Replace x, y, z of Fig. 2 by $1-4x$, $1-4y$, $1-4z$, respectively. Then, e.g., the length $|AB|$ is equal to $2[4x^2a^2 + (1/2 - 2y)^2b^2 + c^2/4]^{1/2}$ and so on, as (1.5) and (1.6) imply. Finally, by the lemma we have obtained three optimal ball systems with the density in (0.6). These are:

1. $a^2 = 1/2, \; b^2 = 1, \; c^2 = 2, \; x = 0, \; y = 1/4, \; z = 1/8;$
2. $a^2 = 1/2, \; b^2 = 4/3, \; c^2 = 3/2, \; x = 1/4, \; y = 1/4, \; z = 1/12;$
3. $a^2 = 1, \; b^2 = 1, \; c^2 = 1, \; x = 1/4, \; y = 1/4, \; z = 1/4.$

The first and third arrangements are lattice-like, the second is not.

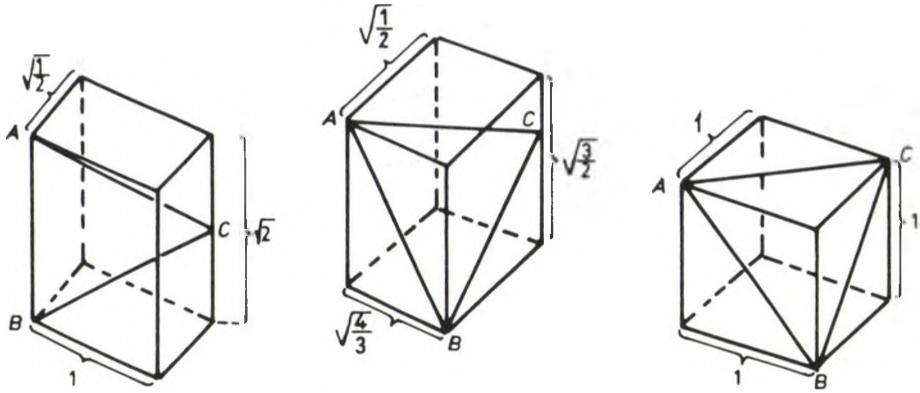


Fig. 3

2. The cases of the groups $P2_1$, Pb, Bb

These three space-groups G generates double lattice-like ball packings, respectively. The optimal ball systems give the density by (0.6) (see [5], [6], [8]). We can prove this in a direct way, too. Namely we regard the metric normalizers M_G and their fundamental sets $F(M_G)$ (see [2]). We have to solve the following problems, respectively:

$$\text{(i) } \delta(P2_1) = \max\left\{ \frac{(l^3\pi)/6}{(abc\sin\tau)/2} \,\middle|\, z = 0 \leqq x, y;\ x + y = 1/2;\ \pi/2 \leqq \tau < 2\pi/3;\ 0 < a, b, c \right\},$$

where

$$l = \min\{a, c, (4x^2a^2 + 2ab\cos\tau(2x)(2y) + 4y^2b^2 + c^2/4)^{1/2}\};$$

$$\text{(ii) } \delta(Pb) = \max\left\{ \frac{(l^3\pi)/6}{(abc\sin\tau)/2} \,\middle|\, x = y = 0 \leqq z \leqq 1/4;\ \pi/2 \leqq \tau < \pi;\ 0 < a \leqq b;\ 0 < c \right\},$$

where

$$l = \min\{a, c, (b^2/4 + 4z^2c^2)^{1/2}\};$$

$$\text{(iii) } \delta(Bb) = \max\left\{ \frac{(l^3\pi)/6}{(abc\sin\tau)/4} \,\middle|\, x = y = 0 \leqq z \leqq 1/4;\ \pi/2 \leqq \tau < \pi;\ 0 < a, b, c \right\},$$

where

$$l = \min\left\{ a, b, c, (a^2+c^2)^{1/2}/2, [(a^2+c^2)/4 + ab\cos\tau + b^2]^{1/2}, (b^2/4 + 4z^2c^2)^{1/2} \right\}.$$

Since these results are treated in [5], [6] as particular cases, we omit here the lengthy calculations which are similar to those at the other groups. The parameters of the optimal ball packings are as follows:

(i) $a=b=c(2)^{-1/2}$, $\tau=\pi/2$, $x=y=1/4$ (lattice-like)
$a=b=c(8/3)^{-1/2}$, $\tau=2\pi/3$, $x=1/3$ $y=1/6$ (nonlattice-like);
(ii) $2^{1/2}a=b=c$, $\tau=3\pi/4$, $z=1/4$ (lattice-like);
(iii) $a=b=c$, $\tau=\pi/2$, $z=1/4$ (lattice-like)
$a=c(3)^{-1/2}$, $\tau=\pi/2$, $z=1/6$ (nonlattice-like).

3. The groups $\mathrm{Pca2_1}$ and $\mathrm{Pna2_1}$

First we take the group $G=\mathrm{Pca2_1}$ and consider the metric normalizer M_G of the group (see [2]). $M_G=\mathrm{Z^1mmm}$ and its fundamental set is $F(M_G)= =\{(x,y,z)\mid z=0\leqq x,y\leqq 1/4\}$. Here L_G is spanned by the orthogonal basis $\{\mathbf{e}_i\}$ where $|\mathbf{e}_1|=a$, $|\mathbf{e}_2|=b$, $|\mathbf{e}_3|=c$. The point positions are:

$X(x,y,z), X^{2_1}(-x,-y,z+1/2), X^c(-x+1/2,y,z+1/2), X^a(x+1/2,-y,z)$

and the vectors XX^g, $g\in G$ are of the form:

$$(3.1)\qquad (0,0,0),\ (-2x,-2y,1/2),\ (1/2-2x,0,1/2),\ (1/2,-2y,0)$$

plus an integer triple from L_G. Since $0\leqq x$, $y\leqq 1/4$ we see from (3.1) that the infimum in (0.2) is nothing but

$$(3.2)\qquad \begin{aligned} l=\min\Big\{&a,b,c,(4x^2a^2+4y^2b^2+c^2/4)^{1/2},\\ &[(1/2-2x)^2a^2+c^2/4]^{1/2},(a^2/4+4y^2b^2)^{1/2}\Big\}\end{aligned}$$

and we look for

$$(3.3)\qquad \delta(G)=(2\pi/3)\max\{l^3\mid z=0\leqq x,y\leqq 1/4,\ 0<a,b,c,\ abc=1\}.$$

Since by (3.2) l is increasing in y, we may suppose that y is equal to $1/4$. Then we see that in the case $x=0$, $y=1/4$, $z=0$ the ball packing is lattice-like since XX^G, $g\in G$ are of the form

$$(3.4)\qquad (0,0,0),\ (0,-1/2,1/2),\ (1/2,0,1/2),\ (1/2,-1/2,0)$$

plus any triple of integers from L_G. We obtain the optimal face centered cubic lattice iff $a=b=c=1$, then $l=2^{-1/2}$ and $\delta=\pi/(18)^{1/2}$.

To prove $l\leqq 2^{-1/2}$ we may assume $x>0$, moreover,

$$(3.5)\qquad \begin{aligned} &1/2\leqq a^2,b^2,c^2=(ab)^{-2},(a^2+b^2)/4\ \text{ and}\\ &d^2=4x^2a^2+(b^2+c^2)/4=(1/2-2x)^2a^2+c^2/4.\end{aligned}$$

From (3.5) we have

$$(3.6)\quad 0<x=\frac{a^2-b^2}{8a^2}<1/4 \text{ and } d^2=(1/2-2x)^2a^2+c^2/4=\frac{(a^2+b^2)^2}{16a^2}+\frac{c^2}{4}.$$

First we assume that $3b^2 = a^2$, hence $l^2 = \min\{1/3b^4, b^2, b^2/3 + 1/12b^4\} = 1/2$ by easy calculations, we get the optimal non-lattice arrangement with the parameters $b^2 = 1/2$, $a^2 = 3/2$, $c^2 = 4/3$, $x = 1/12$, $y = 1/4$, $z = 0$. The vectors XX^G, $g \in G$ will be

$$(0,0,0),\ (-1/16; -1/2; 1/2),\ (1/3; 0; 1/2),\ (1/2; -1/2; 0) \tag{3.4*}$$

plus any integer triple from the lattice L_G. We shall prove that $l^2 \leqq 1/2$ holds also in the other cases. We introduce a new variable u by

$$u^2 = a^2/b^2 > 1. \tag{3.7}$$

1) Assume that

$$1/2 \leqq c^2 \leqq (a^2 + b^2)/4 < b^2 < a^2, \text{ i.e. by substitution, } 1 < u^2 < 3 \tag{3.8}$$

and $1/2 \leqq c^2 \leqq 2^{-4/3}u^{-2/3}(u^2+1)^{2/3}$ for any fixed u. Then $d^2 = \frac{1}{16c}\frac{(u^2+1)^2}{u^3} + \frac{c^2}{4}$ stands by (3.6) and

$$d^2 \leqq \max\{2^{-7/2}u^{-3}(u^2+1)^2 + 2^{-4}; 2^{-10/3}u^{-8/3}(2u^2+1)(u^2+1)^{2/3}\} \tag{3.9}$$

holds for any fixed $u \in (1, 3^{1/2})$. Since both the above functions of u would take their maxima either at 1 or at $3^{1/2}$, hence

$$d^2 < \max\{2^{-3/2} + 2^{-4}; 2^{1/2}3^{-3/2} + 2^{-4}; 3(2^{-8/3}); 7(2^{-2})3^{-4/3}\} < 1/2. \tag{3.10}$$

2) Assume that $1 < u^2 < 3$ and $1/2 \leqq (a^2 + b^2)/4 \leqq c^2$. Introducing $e^2 = (a^2 + b^2)/4$, we express all the variables by e and u. Then we have our assumptions:

$$1 < u^2 < 3 \text{ and } 1/2 \leqq e^2 \leqq 2^{-4/3}u^{-2/3}(u^2+1)^{2/3} \tag{3.11}$$

for any fixed u. By (3.6) we have

$$d^2 = \frac{e^2}{4}\frac{(u^2+1)}{u^2} + \frac{1}{64e^4}\frac{(u^2+1)^2}{u^2} \tag{3.12}$$

and, again by substitution

$$d^2 \leqq \max\{2^{-4}u^{-2}(u^2+1)(u^2+3); 2^{-10/3}u^{-8/3}(2u^2+1)(u^2+1)^{2/3}\} \tag{3.13}$$

holds for any fixed $u \in (1, 3^{1/2})$. Again, taking $u^2 = 1$ and 3 in (3.13) we obtain

$$d^2 < \max\{1/2; 1/2; 3(2)^{-8/3}; 7(2)^{-2}(3)^{-4/3}\} = 1/2. \tag{3.14}$$

3) Assume $1/2 \leqq c^2 \leqq b^2 < (a^2+b^2)/4$. We express all the variables by c and u. Then

(3.15) $$3 < u^2 \text{ and } 1/2 \leqq c^2 \leqq u^{-2/3} \text{ means also } u^2 \leqq 8.$$

Again $d^2 = \frac{1}{16c}\frac{(u^2+1)^2}{u^3} + \frac{c^2}{4}$ and

(3.16) $$d^2 \leqq \max\{2^{-7/2}u^{-3}(u^2+1)^2 + 2^{-4}; 2^{-4}; 2^{-4}u^{-8/3}(u^4+6u^2+1)\}$$

holds for any fixed $u^2 \in (3,8]$. Hence

(3.17) $$d^2 \leqq \max\{2^{1/2}3^{-3/2} + 2^{-4}; 5^2 2^{-8}; 7(2^{-2})3^{-4/3}; 2^{-8}113\} < 1/2.$$

4) Finally, assume $3 < u^2$ and $1/2 \leqq b^2 \leqq c^2$. Then

(3.18) $$1/2 \leqq b^2 \leqq u^{-2/3} \quad \text{and} \quad 3 < u^2 \leqq 8$$

hold. Now

(3.19) $$d^2 = \frac{b^2}{16}\frac{(u^2+1)^2}{u^2} + \frac{1}{4u^2b^4}$$

stands and

(3.20) $$d^2 \leqq \max\{2^{-5}u^{-2}(u^2+1)^2 + u^{-2}; 2^{-4}u^{-8/3}(u^4+6u^2+1)\}$$

holds for any fixed $u^2 \in (3,8]$. Hence

(3.21) $$d^2 < \max\{1/2; 2^{-8}113; 7(2)^{-2}3^{-4/3}; 2^{-8}113\} = 1/2.$$

So, we have two optimal ball systems for $G = \mathrm{Pca}2_1$

(3.22)
1. $a = b = c = 1, \quad x = 0, \quad y = 1/4, \quad z = 0 \qquad (3.4)$
2. $a^2 = 3/2,\ b^2 = 1/2,\ c^2 = 4/3,\ x = 1/12,\ y = 1/4,\ z = 0 \quad (3.4^*)$.

Secondly, we take the group $G = \mathrm{Pna}2_1$. In this case the lattice L_G is also spanned by the orthogonal basis $\{\mathbf{e}_i\}$ where the lengths of these vectors are a, b, c, respectively. The point positions are:

(3.23) $$X(x,y,z),\ X^{2_1}(-x,-y,z+1/2),\ X^n(1/2-x,1/2+y,1/2+z) \text{ and } X^a(1/2+x,1/2-y,z)$$

and the vectors XX^g, $g \in G$ are of the form:

(3.24) $$(0,0,0),\ (-2x,-2y,1/2),\ (1/2-2x,1/2,1/2),\ (1/2,1/2-2y,0)$$

plus an integer triple from L_G.

Since the metric normalizer is $M_G = \mathrm{Z}^1\mathrm{mmm}$ and its fundamental set is $F(M_G) = \{(x,y,z) \mid z = 0 \leqq x, y \leqq 1/4\}$ we get the following:

$$l = \min\Big\{a, b, c, (4x^2a^2 + 4y^2b^2 + c^2/4)^{1/2}, \quad [(1/2-2x)^2a^2 + (b^2+c^2)/4]^{1/2},\ [a^2/4 + (1/2-2y)^2b^2]^{1/2}\Big\} \tag{3.25}$$

and

$$\delta(G) = (2\pi/3)\max\{l^3 \mid 0 \leqq x, y \leqq 1/4,\ 0 < a, b, c, \text{ and } abc = 1\}, \tag{3.26}$$

for this reason we have to prove the inequality $l^2 \leqq 1/2$, too. Certainly we may assume that $a^2, b^2, c^2 \geqq 1/2$ (3.27) and prove

$$\max\Big\{\min\{(4x^2a^2 + 4y^2b^2 + c^2/4), [(1/2-2x)^2a^2 + (b^2+c^2)/4], [a^2/4 + (1/2-2y)^2b^2]\} \mid 0 \leqq x, y \leqq 1/4,\ 1/2 < a^2, b^2, c^2 = (ab)^{-2}\Big\} = 1/2.$$

We have the same geometric statement as in the Lemma at the group $\mathrm{P}2_12_12_1$ (see Figure 4).

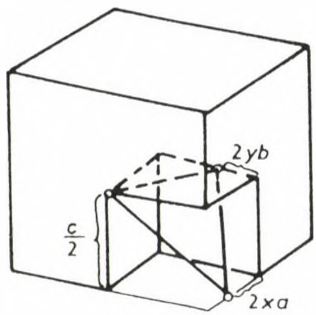


Fig. 4

The optimal arrangements in this case are as follows:

1. $a^2 = 2,\quad b^2 = 1/2,\quad c^2 = 1,\quad x = 1/8,\quad y = 1/4,\quad z = 0;$
2. $a^2 = 3/2,\quad b^2 = 1/2,\quad c^2 = 4/3,\quad x = 1/6,\quad y = z = 0;$
3. $a = b = c = 1,\quad x = 1/4,\quad y = z = 0.$

4. The groups $\mathrm{P}4_1$, $\mathrm{P}3_1$, and $\mathrm{P}6_1$

First we take the group $G = \mathrm{P}4_1$. The translation lattice L_G is spanned by the orthogonal basis $\{\mathbf{e}_i\}$ with the lengths a, a, c, respectively. The point positions are:

$$X(x,y,z),\ X^s(-y,x,z+1/4),\ X^{s^2}(-x,-y,z+1/2) \quad X^{s^3}(y,-x,z+3/4) \tag{4.1}$$

and the vectors XX^g, $g \in G$ are:

$$(4.2)\quad (0,0,0),\ (-y-x, x-y, 1/4),\ (-2x, -2y, 1/2),\ (y-x, -x-y, 3/4)$$

plus an integer triple from L_G. The metric normalizer M_G is the group $Z^1 422$ and its fundamental domain is

$$(4.3)\qquad F(M_G) = \{(x,y,z) \mid z = 0,\ x \geqq 0,\ y - x \geqq 0,\ 1/2 - x - y \geqq 0\}.$$

Let l be the $\inf\{\varrho(XX^g) \mid g \in G\}$. It is easy to see that

$$(4.4)\quad l = \min\{a, c, [2(x^2+y^2)a^2 + c^2/16]^{1/2}, [(4x^2 + (1-2y)^2)a^2 + c^2/4]^{1/2}\}$$

and

$$(4.5)\qquad \delta(G) = (2\pi/3)\max\{l^3/a^2c \mid x, y \in F(M_G),\ a, c > 0\}.$$

Suppose that $a^2c = 1$ and $a, c \geqq 2^{-1/2}$. This means that we have to prove the following inequality:

$$(4.6)\qquad \max\Big\{\min\{2(x^2+y^2)a^2 + c^2/16, (4x^2 + (1-2y)^2)a^2 + c^2/4\} \mid \\ \mid 1/2 \leqq a^2 = 1/c, c^2; x, y \in F(M_G)\Big\} \leqq 1/2.$$

First we assume that $2^{-1/2} \leqq c \leqq (8/3)^{1/3}$. Fixing the value c we look for those points $(x, y, 0) \in F(M_G)$ where

$$(4.7)\qquad 2(x^2+y^2)a^2 + c^2/16 = [4x^2 + (1-2y)^2]a^2 + c^2/4 =: d^2.$$

It is easy to see that from this we get

$$(4.8)\qquad c^3 = (32/3)[-x^2 - (y-1)^2 + 1/2] \leqq 8/3$$

and so

$$x^2 = 1/2 - (y-1)^2 - 3c^3/32.$$

That means

$$(4.9)\qquad d^2 = 2(x^2+y^2)a^2 + c^2/16 = (-1+4y)/c - c^2/8.$$

Since by (4.3) the inequalities $0 \leqq x \leqq 1/2 - y$ hold, from (4.8) we get

$$(4.10)\quad 1/2 - 3c^3/32 = (y-1)^2 + x^2 \leqq (y-1)^2 + (1/2-y)^2 = 2y^2 - 3y + 5/4$$

and so

$$(4.11)\qquad y \leqq 3/4 - (12 - 3c^3)^{1/2}/8.$$

By (4.9) we have

(4.12) $$d^2 \leqq 2/c - c^2/8 - (3/c^2 - 3c/4)^{1/2} =: g(c).$$

A difficult computation shows that $g(c) \leqq 1/2$ on the closed interval $[2^{-1/2}, (8/3)^{1/3}]$.

Consider now $5^{1/2} - 1 \leqq c \leqq 2$. Since the points $(x, y, 0)$ are elements of $F(M_G)$, it is easy to see that $0 \leqq x^2 + y^2 \leqq 1/4$ and so

(4.13) $$2(x^2 + y^2)a^2 + c^2/16 \leqq 1/2c + c^2/16 \leqq 1/2.$$

By (4.6) we have considered each occurring c, and the optimal packing is the following: 1, $a^2 = 1/2$, $c = 2$, $x = z = 0$, $y = 1/2$.

Secondly, we take the group $G = \mathrm{P}3_1$. Now the translation lattice L_G is spanned by the basis $\{\mathbf{e}_i\}$ with Gramian matrix

(4.15) $$(\langle \mathbf{e}_i, \mathbf{e}_j \rangle) = \begin{bmatrix} a^2 & -a^2/2 & 0 \\ -a^2/2 & a^2 & 0 \\ 0 & 0 & c^2 \end{bmatrix}.$$

The point positions are

(4.16) $$X(x, y, z),\ X^s(-y, x - y, z + 1/3),\ X^{s^2}(y - x, -x, z + 2/3)$$

and the vectors XX^g, $g \in G$ are

(4.17) $$(0, 0, 0),\ (-y - x, x - 2y, 1/3),\ (y - 2x, -x - y, 2/3)$$

plus a vector from L_G. The metric normalizer M_G is the group $\mathrm{Z}^1 622$ and its fundamental domain is

(4.18) $$F(M_G) = \{(x, y, z) \mid z = 0,\ 0 \leqq y,\ 0 \leqq 1/3 - x,\ 0 \leqq x/2 - y\}.$$

Let l be the infimum by (0.2) then

(4.19) $$\delta(G) = \frac{\pi l^3}{a^2 c 3^{1/2}}.$$

If $l^3 = a^2 c / 6^{1/2}$ then $\delta(G) = \pi / 18^{1/2}$. Suppose $a^2 c = 1$ then we have to prove that the maximum l^2 is not greater than $6^{-1/3}$ if the number c is on the closed interval $[6^{-1/6}, 6^{1/3}]$. But

(4.20) $$\begin{aligned} l^2 &\leqq (y + x)^2 a^2 + (x - 2y)^2 a^2 - (x + y)(2y - x)a^2 + c^2/9 = \\ &= (3/4)[3x^2 + (x - 2y)^2]a^2 + c^2/9 \leqq a^2/3 + c^2/9 =: d^2, \end{aligned}$$

where the equality stands iff $x = 1/3$, $y = 0$. Taking into consideration $a^2 = c^{-1}$, the function $d^2 := (a^2/3 + c^2/9)^{1/2}$ is not greater than $6^{-1/6}$ on the interval $[6^{-1/6}, 6^{1/3}]$. If $c = 6^{1/3}$ then $d^2 = 6^{-1/6}$. Thus

$$\delta(G) = \pi/18^{1/2} \tag{4.21}$$

holds, too. The optimal arrangement is the following:

$$a^2 = 6^{-1/3}, \quad c = 6^{1/3}, \quad x = 1/3, \quad y = z = 0, \tag{4.22}$$

and the packing is the densest lattice-like one. This is the well-known double lattice-like packing with the density $\pi/18^{1/2}$.

We take now the space group $G = \mathrm{P}6_1$. The lattice L_G is the same as in the previous case of $\mathrm{P}3_1$. Figure 5 shows the projection onto the plane $(0; \mathbf{e}_1, \mathbf{e}_2)$.

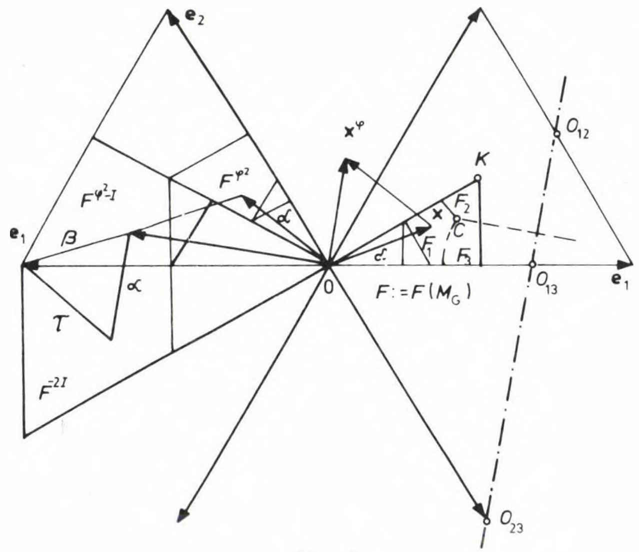


Fig. 5

Denote $X(x, y, z)$ a general point. The vectors XX^g, $g \in G$ are

$$\begin{gathered}(0,0,0),\ (-y, x-y, 1/6),\ (-y-x, x-2y, 1/3),\ (-2x, -2y, 1/2),\\ (y-2x, x-y, 2/3),\ (y-x, -x, 5/6)\end{gathered} \tag{4.23}$$

plus an integer triple from L_G.

The metric normalizer M_G is the group $\mathrm{Z}^1 622$ with the fundamental domain

$$F := F(M_G) = \{X(x,y,z) \mid z = 0,\ 0 \leqq y,\ 0 \leqq y - 2x + 1,\ 0 \leqq x - 2y\}. \tag{4.24}$$

In Figure 5 we see F with the projection vector $\mathbf{x}(x, y)$ of OX. Let $\varphi: \mathbf{x} \longmapsto \mathbf{x}^\varphi$ denote the 6-rotation through $\pi/3$. Then the vectors in (4.23) have components in the $(\mathbf{e}_1, \mathbf{e}_2)$-plane:

$$\mathbf{x} - \mathbf{x} = 0,\ \mathbf{x}^\varphi - \mathbf{x} = \mathbf{x}^{\varphi^2},\ \mathbf{x}^{\varphi^2} - \mathbf{x},\ \mathbf{x}^{\varphi^3} - \mathbf{x} = -2\mathbf{x}, \dots,$$

which determine also the corresponding images of F denoted by

(4.25) $$F^{\varphi^2}, F^{\varphi^2-I}, F^{-2I}, \quad \text{respectively.}$$

We look for the maximal density

$$\delta(G)=\frac{2\pi l^3}{a^2c(3)^{1/2}}$$

where l is the infimum by (0.2). We may assume $a^2c=1$ and conclude that

$$l^2=\min\{d_1^2,d_2^2,d_3^2,c^2,a^2=1/c\} \quad \text{with}$$

$$d_1^2=[(-y)^2+(x-y)^2+y(x-y)]/c+c^2/36=(x^2+y^2-xy)/c+c^2/36=:$$
$$=:\alpha^2/c+c^2/36,$$

(4.26) $$d_2^2=[(-1+y+x)^2+(-x+2y)^2-(-1+y+x)(-x+2y)]/c+c^2/9=$$
$$=(3\alpha^2-3x+1)/c+c^2/9=:\beta^2/c+c^2/9,$$
$$d_3^2=[(-1+2x)^2+(2y)^2-(-1+2x)(2y)]/c+c^2/4=$$
$$=(4\alpha^2-4x+2y+1)/c+c^2/4=:\tau^2/c+c^2/4.$$

To show $\delta(G)\leqq\pi(18)^{-1/2}$, we shall prove that

(4.27) $$\min\{d_1^2,d_2^2,d_3^2\}\leqq(24)^{-1/3} \quad \text{if } c\in[(24)^{-1/6},(24)^{1/3}].$$

We draw three curves $c_{ij}:=\{(x,y)\in F \mid d_i^2=d_j^2\}$ $i\neq j=1,2,3$. These are circle arcs with centres $O_{12}(1,1/2)$, $O_{13}(2/3,0)$, $O_{23}(0,-1)$ as follows:

(4.28) $$c_{12}:(x-1)^2+(y-1/2)^2-(x-1)(y-1/2)=1/4-c^3/24, \quad c^3\leqq 6;$$

(4.29) $$c_{13}:(x-2/3)^2+y^2-(x-2/3)y=1/9-2c^3/27, \quad c^3\leqq 3/2;$$

(4.30) $$c_{23}:x^2+(y+1)^2-x(y+1)=1-5c^3/36, \quad c^3\leqq 36/5.$$

Thus we determine the domain F_i in F, where d_i^2 is the minimum in (4.27). The pencil of circles c_{ij} may have a common point C where the minimum d_i^2 is maximal. This fact depends on the parameter c of the group $G=\mathrm{P}6_1$. A straightforward but awful computation yields the coordinates of the extremal point C. First, C lies on the power line of the circle pencil $\{c_{ij}\}$:

(4.31) $$x+4y=1-7c^3/36.$$

Then we substitute, say, into (4.28) and solve the equation of second degree. We obtain $C(x_c, y_c)$

$$x_c = \frac{5}{7} - \frac{c^3}{36} - \frac{1}{2(3)^3 7}[2^4 3^6 - 2^3 3^3 5(7)c^3 - 3(7)^2 c^6]^{1/2},$$
(4.32)
$$y_c = \frac{1}{14} - \frac{c^3}{24} + \frac{1}{2^3 3^3 7}[2^4 3^6 - 2^3 3^3 5(7)c^3 - 3(7)^2 c^6]^{1/2},$$

whenever $c^3 \leqq (2^2 3^2/7)(2(7)^{1/2} - 5) = 1.4992$, i.e. $c \leqq 1.1445$.

C just lies on F if (x_c, y_c) satisfies (4.24) besides

$$0.5888 = (24)^{-1/6} \leqq c \leqq (24)^{1/3} = 2.8845.$$

We obtain for $C \in F$ the interval

(4.33) $$0.5888 = (24)^{-1/6} \leqq c \leqq 7^{-2/3}(162(57)^{1/2} - 1170)^{1/3} = 1.0270,$$

(4.34) $$d_1^2(C) = \frac{2^2}{7c} - \frac{c^2}{18} - \frac{1}{2^2 3^2 7c}[2^4 3^6 - 2^3 3^3 5(7)c^3 - 3(7)^2 c^6]^{1/2}.$$

A careful computation shows that $d_1^2(C) \leqq (24)^{-1/3}$ holds on the interval (4.33). If C lies out of F, then the intersection of c_{12} and the segment on $y = 2x - 1$, i.e.

(4.35) $$H(3/4 - (9 - 2c^3)^{1/2}/12; 1/2 - (9 - 2c^3)^{1/2}/6)$$

provides the minimum by (4.27), and we obtain

(4.36) $$d_1^2(H) = 5/(8c) - c^2/72 - (9 - 2c^3)^{1/2}/(8c) \text{ for } 1.027 \leqq c \leqq 4^{1/3}.$$

We need again a careful computation to show that $d_1^2(H) \leqq (24)^{-1/3} = 0.34668$ holds on the interval (4.36). If $c^3 = 4$ then we arrive into the vertex $K(2/3, 1/3)$ of F. In the interval

(4.37) $$4 \leqq c^3 \leqq 24, \text{ i.e. } 1.5874 \leqq c \leqq 2.8845$$

the point $K(2/3, 1/3)$ provides the minimum in (4.27) and

(4.38) $$d_1^2(K) = 1/(3c) + c^2/36$$

is a convex function in (4.37), taking its maximum at $c = (24)^{1/3}$. Then $d_1^2 = (24)^{-1/3}$ as we stated at (4.27). We have got the optimal ball arrangement parametrized by

(4.39) $$c = (24)^{1/3} \quad a = (24)^{-1/6} \quad x = 2/3 \quad y = 1/3.$$

This arrangement is not lattice-like.

We remark that the last observation at (4.38) would give the estimate $l^2 \leqq (24)^{-1/3}$ in the interval

$$1.0558 = 3^{5/6} - 3^{1/3} \leqq c \leqq (24)^{-1/3} = 2.8845,$$

but this does not give more and we need such awful computations as we did.

Now the result formulated at the end of the introduction is completely proved.

REFERENCES

[1] HILBERT, D. and COHN-VOSSEN, S., *Anschauliche Geometrie*, Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen, Bd. 37, Springer-Verlag, Berlin, 1932. *Jb. Fortschritte Math.* **58**, 597

[2] KOCH, E. and FISCHER, W., Euclidean and affine normalizers of space groups and their use in crystallography, *International tables for X-ray crystallography*, Vol. A, Ed. by Theo Hahn, Reidel, 1983.

[3] *International tables for X-ray crystallography*, Vol. A, Ed. by Theo Hahn, Reidel, 1983.

[4] HORVÁTH, Á. G., Egy nehezen kezelhető geometriai szélsőérték problémáról, *Mat. Lapok* (to appear).

[5] HOLLAI, M., Doppelgitterförmige Lagerungen inkongruenter Kreise und Kugeln, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **18** (1975), 75–86. *MR* **54** #5978a

[6] HORVÁTH, Á. G., On double lattice-like sphere packing, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **33** (1990), 53–60.

[7] MOLNÁR, E., Minimal presentation of the 10 compact Euclidean space forms by fundamental domains, *Studia Sci. Math. Hungar.* **22** (1987), 19–51. *MR* **88j**: 51029

[8] HORVÁTH, Á. G., Correction to my paper "On double lattice-like sphere packing", *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* (to appear).

(Received July 10, 1990)

BUDAPESTI MŰSZAKI EGYETEM
GÉPÉSZMÉRNÖKI KAR
GEOMETRIA TANSZÉK
EGRI JÓZSEF U. 1
H–1521 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 25–32

AN ILLUMINATION PROBLEM

V. SOLTAN

Abstract

For different types of illumination, the following problem is studied: to determine the maximum natural number m such that any m points in the boundary of a convex body $K \subset E^n$ can be illuminated by an exterior source (point or direction).

Illumination by a point

Let K be a convex body (a proper closed convex set with nonempty interior) in the n-dimensional linear space E^n. The usual abbreviations aff, bd, int are taken for affine hull, boundary, and interior, respectively. The notations $[x, y]$, $]x, y[$, (x, y), $[x, y\rangle$ will mean closed line interval, open line interval, the line passing through different points x, y, and the ray with the apex x passing through a point y. For any (oriented) direction l in E^n, l_x means the ray with the apex x having the direction l, and l'_x means the ray with the apex x having the direction opposite to l.

L. Fejes Tóth [1] introduced the following notion of *illumination:* a point $x \in \mathrm{bd} K$ is called illuminated by a point $y \in E^n \setminus K$ provided $]x, y[\cap K = \emptyset$. This notion was introduced earlier by F. A. Valentine [2] in terms of visibility.

E. Buchman and F. A. Valentine [3] considered another notion of visibility, which we shall name *weak illumination:* a point $y \in E^n \setminus K$ illuminates weakly a point $x \in \mathrm{bd} K$ if $[x, y] \cap \mathrm{int} K = \emptyset$.

There is known a stronger type of illumination, introduced by H. Hadwiger [4] (see also [5]): a point $x \in \mathrm{bd} K$ is called illuminated by a point $y \in E^n \setminus K$ if $]x, y[\cap K = \emptyset$ and the ray $[y, x\rangle$ intersects $\mathrm{int} K$. This type of illumination we shall call *strong.*

Denote by $a(K)$ the maximum natural number $m \geqq 1$ such that any m boundary points of K can be illuminated weakly by a point from $E^n \setminus K$. Put $a(K) = \infty$ if such number m does not exist.

Similarly, define by $b(K)$ and $c(K)$ the corresponding numbers for usual and strong illumination, respectively.

1991 *Mathematics Subject Classification.* Primary 52A37.

Key words and phrases. Weak illumination, strong illumination of several boundary points of compact and unbounded convex bodies, by points and by directions.

THEOREM 1. *If K is compact, then $1 \leqq a(K) \leqq n$. If K is unbounded, then either $1 \leqq a(K) \leqq n-1$ or $a(K) = \infty$.*

For the proof of Theorem 1 we need a lemma.

LEMMA 1. *If $K = M \oplus L$ is a decomposition of an unbounded convex body K into the direct sum of a line-free closed convex set M and a linear subspace L, then $a(K) = a(M)$, where $a(M)$ denotes the respective number for M in the space* aff$M$.

PROOF. Denote by π a linear projection on aff$M$ parallel to L. Then $M = \pi(K)$ and rbd$M = \pi(\text{bd}K)$. The inequality $a(M) \leqq a(K)$ is trivial. Let $a(M) = m$, and $x_1, \dots, x_m \in \text{bd}K$ be any points. By definition, the points $\pi(x_1), \dots, \pi(x_m)$ are illuminated weakly by some point $z \in \text{aff}M \setminus M$. Obviously, $x_1, \dots, x_m$ are illuminated weakly by z. Hence $a(K) \leqq a(M)$.

The case $a(M) = \infty$ is considered by analogy. □

PROOF OF THEOREM 1. Let K be compact, and denote by R the set of all regular points of K. For any point $x \in R$ denote by H_x the hyperplane supporting K at x, and by P_x that closed halfspace determined by H_x which does not contain K. Since K is compact, one has $\cap\{P_x : x \in R\} = \emptyset$. Hence we can choose at most $n+1$ regular points $x_1, \dots, x_{n+1}$ of K such that $\cap\{P_{x_i} : 1 \leqq i \leqq n+1\} = \emptyset$.

Let y be any point in $E^n \setminus K$. One has $y \notin P_{x_i}$ for some number $i = 1, \dots, n+1$. Then $[y, x_i] \cap \text{int}K \neq \emptyset$, i.e. y does not illuminate weakly the point x_i. Thus, any point $y \in E^n \setminus K$ cannot illuminate weakly the whole set $\{x_1, \dots, x_{n+1}\}$. So $a(K) \leqq n$. The inequality $a(K) \geqq 1$ is trivial.

Let K be unbounded. By Lemma 1, K can be considered as line-free. Suppose that $a(K) \geqq n$. Choose any natural number m, and let $x_1, \dots, x_m$ be any boundary points of K. Select some hyperplane H which dissects K in two closed parts with non-empty interiors such that one part, say K_0, is bounded and contains all the points $x_1, \dots, x_m$. Denote by N the unbounded convex body which is the intersection of all closed half-spaces supporting K at all the regular points of K belonging to K_0. Obviously, N can be represented as $N = K_0 + C$, where C is the characteristic cone of N.

We want to show that dim$C = n$. Suppose the contrary. Then we can choose some n closed half-spaces $P_1, \dots, P_n$ supporting N at the regular points $z_1, \dots, z_n \in \text{bd}N$ belonging to $H \cap \text{bd}K$ such that the dimension of $\cap\{P_i - z_i : 1 \leqq i \leqq n\}$ is less than n. In this situation, the closed halfspaces $T_i = E^n \setminus \text{int}Q_i$, $i = 1, \dots, n$, have an empty intersection. If y is any point in $E^n \setminus K$, then $y \notin T_i$ for some T_i. The last means that y does not illuminate weakly the point z_i. Hence the set $\{z_1, \dots, z_n\}$ cannot be illuminated weakly by a point, which is in contradiction with the assumption $a(K) \geqq n$. Therefore dim$C = n$.

Since dim$C = n$, the body $N = K_0 + C$ is contained in some translate $\bar{C}$ of C. If y is the apex of the cone $\bar{C}$, then, obviously, y illuminates weakly the whole boundary of N. In particular, y weakly illuminates the points $x_1, \dots, x_m$. Hence $a(K) \geqq m$. Because the number m is chosen arbitrary, one has $a(K) = \infty$. □

THEOREM 2. *For any convex body $K \subset E^n$, one has $a(K) = b(K)$.*

PROOF OF THEOREM 2 follows obviously from the next statement: a set $X \subset \mathrm{bd} K$ is illuminated by a point from $E^n \setminus K$ if and only if X is illuminated weakly by a point from $E^n \setminus K$.

If a point $y \in E^n \setminus K$ illuminates X, then, trivially, y illuminates X weakly. Conversely, let some point $z \in E^n \setminus K$ illuminate X weakly. Denote by D_z the cone with the apex z generated by K:

$$D_z = \cup\{z + \lambda(K - z) : \lambda \geqq 0\}.$$

Let D'_z denote the cone with the apex z symmetric to $D_z : D'_z = 2z - D_z$. Fix any point $y \in \mathrm{int} D'_z$. An easy verification shows that $[y, x] \cap K = \{x\}$ for each point $x \in X$. In other words, y illuminates X. □

COROLLARY 1. *If K is compact, then $1 \leqq b(K) \leqq n$. If K is unbounded, then either $1 \leqq b(K) \leqq n - 1$ or $b(K) = \infty$.*

We need a lemma, which trivially follows from the definitions.

LEMMA 2. *A point $x \in \mathrm{bd} K$ is illuminated strongly by a point $y \in E^n \setminus K$ if and only if y belongs to the interior of the cone $D'_x = 2x - D_x$, where*

$$D_x = \cup\{x + \lambda(K - x) : \lambda \geqq 0\},$$

is the cone generated by K with the apex x.

THEOREM 3. *If K is compact, then $c(K) = 1$. If K is unbounded, then either $1 \leqq c(K) \leqq n - 1$ or $c(K) = \infty$.*

PROOF. Let K be compact, and H_1, H_2 be two (different) parallel hyperplanes in E^n supporting K. It is easy to see that any two points $x_1 \in K \cap H_1$, $x_2 \in K \cap H_2$ cannot be strongly illuminated by a point from $E^n \setminus K$. At the same time, any point $x \in \mathrm{bd} K$ can be illuminated by a point from $E^n \setminus K$. Hence $c(K) = 1$.

Let K be unbounded. The characteristic cone of K contains some ray l. Denote by H any hyperplane orthogonal to l.

We need the following observation: for any point $x \in \mathrm{bd} K$, the ray l'_x belongs to the closure of the cone D'_x (see Lemma 2).

Suppose that $c(K) \geqq n$, and choose any natural number m. Let $x_1, \dots, x_m$ be any points in $\mathrm{bd} K$. By Lemma 2, any n of the cones $C_i = \mathrm{int} D'_{x_i}$, $i = 1, \dots, m$, have a common point. From the above observation it follows that the closure of any set of the form

(1) $$C_{i_1} \cap \dots \cap C_{i_n}, \quad 1 \leqq i_1 \dots \leqq i_n \leqq m,$$

contains some translate of the ray $-l$. Hence for each set of the form (1), it is possible to choose a translate of H which intersects this set. The number of sets of the form (1) is finite, namely, $\binom{n}{m}$. Therefore, it is possible to choose a hyperplane G parallel to H such that G intersects each set of the form (1). By Helly's theorem for the space G, all the sets $G \cap C_i$, $i = 1, \ldots, m$, have a common point, say, y. From Lemma 2 it follows that all the points $x_1, \ldots, x_m$ are illuminated strongly by y.

Since m is arbitrary, one has $c(K) = \infty$. □

Illumination by a direction

By analogy to the illumination according to L. Fejes Tóth, we introduce the following type of illumination: a direction l in E^n illuminates a boundary point x of a convex body $K \subset E^n$ if the ray l'_x intersects K at the point x only.

In terms of visibility, E. Buchman and F. A. Valentine [6] introduced the following notion of illumination: a direction l in E^n illuminates weakly a boundary point x of a convex body $K \subset E^n$ if the ray l'_x does not intersect $\operatorname{int} K$.

Following V. G. Boltjanskiĭ [7], we shall say that a direction $l \subset E^n$ illuminates strongly a boundary point x of a convex body $K \subset E^n$ if the ray l_x intersects $\operatorname{int} K$.

Denote by $\overline{a}(K)$ the maximum natural number $m \geqq 1$ such that any m boundary points of K can be weakly illuminated by a direction in E^n. Put $\overline{a}(K) = \infty$ if such number m does not exist.

Similarly, denote by $\overline{b}(K)$ and $\overline{c}(K)$ the corresponding numbers for usual and strong illumination by a direction.

THEOREM 4. *If K is compact, then $n \leqq \overline{a}(K) \leqq 2n - 1$. If K is unbounded, then $\overline{a}(K) = \infty$*

PROOF. Let K be compact, and $x_1, \ldots, x_n$, be any boundary points of K. Denote by H_i some hyperplane supporting K at X_i, and by P_i that closed half-space determined by H_i which does not contain K. Let $\bar{P}_i$ denote the half-space obtained from P_i by the translation on the vector $-x_i$ ($\bar{P}_i$ contains the zero vector 0 in its boundary). Obviously, the intersection $\bar{P}_1 \cap \cap \ldots \cap \bar{P}_n$ contains a ray, say, l. Then the direction determined by the opposite ray $-l$ illuminates weakly all the points $x_1, \ldots, x_n$. Hence $\overline{a}(K) \geqq n$.

It is known (see [6]) that if any $2n - 1$ boundary points of some compact convex body $K \subset E^n$ can be illuminated weakly by a direction in E^n, then K is a parallelepiped. If we choose $2n$ points which are the centres of the $(n - - 1)$-dimensional faces of a parallelepiped, then, clearly, these points cannot be weakly illuminated by a direction. Hence $\overline{a}(K) \leqq 2n - 1$ for any compact convex body $K \subset E^n$.

If K is unbounded, and l is a ray contained in K, then, obviously, the direction determined by l sees the whole set bd$K$. Thus $\bar{a}(K) = \infty$. □

THEOREM 5. *If K is compact, then $1 \leqq \bar{b}(K) \leqq n$. If K is unbounded, then either $1 \leqq \bar{b}(K) \leqq n-1$ or $\bar{b}(K) = \infty$.*

For the proof of Theorem 5 we need some auxiliary definitions and results.

Let $B = \{e_1, \ldots, e_m\}$ be any positive basis in E^n. It is well-known that $n+1 \leqq m \leqq 2n$. A subset $B' \subset B$ will be called minimal positive subbasis of B provided the positive hull of B' is a linear subspace in E^n of dimension card$B' - 1$. It is easy to show (see [8]) that each vector e_i belongs to at least one minimal positive subbasis of B. Put $r(B) =$ sup cardB', where the supremum is considered over the family of all minimal positive subbases of B. From the definitions it follows that $2 \leqq r(B) \leqq n+1$ for any positive basis in E^n.

By analogy to Lemma 1 the following lemma can be proved.

LEMMA 3. *If $K = M \oplus L$ is a decomposition of an unbounded convex body K into the direct sum of a line-free closed convex set M and a linear subspace L, then $\bar{b}(K) = \bar{b}(M)$, where $\bar{b}(M)$ denotes the respective number for M in the space* aff$M$.

PROOF OF THEOREM 5. Let K be compact. For any regular boundary point x of K, denote by e_x the unit vector such that $x + e_x$ is the outer unit normal to K at x. Since K is compact, the set $S = \{e_x : x \in \text{bd}K\}$ positively generates E^n. Denote by $\mathcal{U}$ the family of all positive bases of E^n contained in S and having a minimum possible cardinality. Choose in $\mathcal{U}$ the subfamily $\mathcal{W}$ such that each positive basis $B \in \mathcal{W}$ has a maximum value $r(B)$ in $\mathcal{U}$.

Let $B = \{e_1, \ldots, e_m\} \in \mathcal{W}$ be any positive basis. Denote by $x_1, \ldots, x_m$, respectively, those regular points of K for which $x_i + e_i$, $i = 1, \ldots, m$, are the outer unit normals to K. Obviously, the points $x_1, \ldots, x_m$ cannot be illuminated by a direction in E^n. Hence, if $m = n+1$, one has $\bar{b}(K) \leqq n$. Suppose that $m > n+1$. Without the loss of generality, we may suppose that $B_0 = \{e_1, \ldots, e_k\}$ is a minimum positive subbasis of B having cardinality $k = r(B)$. Denote by L the subspace in E^n orthogonal to posB_0. Since $m > n+1$, one has $L \neq \{0\}$.

It is easy to see that some neighbourhood Σ_i of x_i can be chosen in bd$K$ such that for any regular point $z_i \in \Sigma_i$ the unit vectors

$$e_1, \ldots, e_{i-1}, e_{z_i}, e_{i+1}, \ldots, e_m \tag{2}$$

positively generate E^n.

We want to prove that each neighbourhood Σ_i is contained in a cylindrical surface which can be represented as a direct sum of a set in the subspace posB_0 and a set in the subspace L.

Suppose the contrary: let some neighbourhood Σ_i be not of the above mentioned cylindrical form. Then some regular point $z_i \in \Sigma_i$ can be chosen so that the unit vector e_{z_i} does not belong to posB_0. Since the vectors (2) positively generate E^n and B has a minimum possible cardinality, these vectors form a positive basis, say $\bar{B}$, in E^n. The vectors $e_1, \ldots, e_{i-1}$, e_{z_i}, $e_{i+1}, \ldots, e_k$ are linearly independent. Hence they are properly contained in some minimal positive subbasis of $\bar{B}$. The last shows that $r(\bar{B}) > r(B)$, which is in contradiction with the choice of B.

Thus, each Σ_i is a cylindrical surface. Now, it is easy to see that the points $x_1, \ldots, x_k$ cannot be illuminated by a direction in E^n. Hence $\bar{b}(K) \leqq k - 1 \leqq n$.

Let K be unbounded. By Lemma 3, K can be considered as line-free. Suppose that $\bar{b}(K) \geqq n$. Choose any natural number m, and let $x_1, \ldots, x_m$ be any boundary points of K. Select some hyperplane H which dissects K in two closed parts with non-empty interiors such that one part, say K_0, is bounded and contains all the points $x_1, \ldots, x_m$. Denote by N the unbounded convex body which is the intersection of all closed halfspaces supporting K at all the regular points of K belonging to K_0. Obviously, N can be represented as $N = K_0 + C$, where C is the characteristic cone of N.

We want to show that dim$C = n$. Suppose the contrary. Then we can choose some n closed halfspaces $P_1, \ldots, P_n$ supporting N at the regular points $z_1, \ldots, z_n \in$ bd$N$ belonging to $H \cap$ bd$K$ such that the dimension of $\cap\{P_i - z_i : 1 \leqq i \leqq n\}$ is less than n. Obviously, it is possible to select some k ($\leqq n$) of these halfspaces, say, $P_1 - z_1, \ldots, P_k - z_k$ whose intersection is a linear subspace of dimension $< n$. If l is a ray from the characteristic cone of K, then, obviously, the rays $l_1 = v_1 + l, \ldots, l_k = v_k + l$ are contained in bd$K$. It is easy to verify that any points $w_1 \in l_1 \setminus \{v_1\}, \ldots, w_k \in l_k \setminus \{v_k\}$, cannot be illuminated by a direction in E^n, which is in contradiction with the assumption $\bar{b}(K) \geqq n$. Therefore dim$C = n$.

Choose any ray l_0 contained in int$C$. Obviously, the direction determined by l_0 illuminates the whole boundary of N. In particular, this direction illuminates the points $x_1, \ldots, x_m$. Since m is arbitrary, one has $\bar{b}(K) = \infty$. □

The following lemma is analogous to Lemma 2.

LEMMA 4. *A point $x \in$ bd$K$ is strongly illuminated by a direction $l \subset E^n$ if and only if the ray l_x with the apex x having direction l belongs to the interior of the cone D_x generated by K with the apex x:*

$$D_x = \cup\{x + \lambda(K - x) : \lambda \geqq 0\}.$$

THEOREM 6. *If K is compact, then $\bar{c}(K) = 1$. If K is unbounded, then either $1 \leqq \bar{c}(K) \leqq n - 1$ or $\bar{c}(K) = \infty$*

The proof of Theorem 6 is similar to that of Theorem 3.

OBSERVATION. The consideration in E^n of a simplex, of a ball, of a cone, and of an unbounded convex body which is a direct sum of an $(n-1)$-

dimensional simplex and a ray shows that the bounds for the values $a(K) - \overline{c}(K)$ in Theorems 1–6 are sharp.

A special problem

The following problem is connected with a paper of E. O. Buchman and F. A. Valentine [9]. As above, a point $x \in \text{bd}K$ illuminates weakly a point $y \in \text{bd}K$ if $[x, y] \cap \text{int}K = \emptyset$.

Denote by $d(K)$ the maximum natural number $d \geqq 1$ such that any d boundary points of K can be illuminated weakly by a point from $\text{bd}K$. Put $d(K) = \infty$ if such number d does not exist.

The next theorem is a generalization of a respective result from [9].

THEOREM 7. *For a convex body $K \subset E^n$, the following conditions are equivalent:*

1) $d(K) \geqq m$,

2) *every m maximal faces of K have non-empty intersection.*

PROOF. 1) $\Rightarrow$ 2). Suppose that F_i, $i = 1, \ldots, m$, are any m maximal faces of K. Choose some points $x_i \in \text{rint}F_i$. Since $d(K) \geqq m$, there exists a point $z \in \text{bd}K$ which illuminates weakly all points $x_1, \ldots, x_m$. One has $[z, x_i] \subset F_i$ because of the maximality of F_i. Hence $F_1 \cap \ldots \cap F_m \neq \emptyset$.

The implication 2) $\Rightarrow$ 1) is immediate. □

THEOREM 8. *If K is compact, then $d(K) \leqq n$, with $d(K) = n$ only for simplexes. If K is unbounded, then either $1 \leqq d(K) \leqq n - 1$ or $d(K) = \infty$, with $d(K) = \infty$ only for cones.*

PROOF OF THEOREM 8 follows immediately from Theorem 7 and the following result of P. S. Soltan [10, p. 96]: if every n maximal faces of a convex body $K \subset E^n$ have non-empty intersection, then K is either a simplex or a cone. □

REFERENCES

[1] FEJES TÓTH, L., Illumination of convex discs, *Acta Math. Acad. Sci. Hungar.* **29** (1977), 355–360. *MR* **57** # 4002

[2] VALENTINE, F. A., Visible shorelines, *Amer. Math. Monthly* **77** (1970), 146–152. *MR* **41** # 2530

[3] BUCHMAN, E. and VALENTINE, F. A., External visibility, *Pacific J. Math.* **64** (1976), 333–340. *MR* **55** # 11149

[4] HADWIGER, H., Ungelöste Probleme Nr. 38, *Elem. Math.* **15** (1960), 130–131.

[5] VIZITEI, V. N., Some problems on the covering and illumination of unbounded convex figures, *Bull. Acad. Sci. RSSM*, No. 10 (1961), 3–9.

[6] BUCHMAN, E. and VALENTINE, F. A., A characterization of the parallelepiped in E^n, *Pacific J. Math.* **35** (1970), 53–57. *MR* **42** # 5157

[7] BOLTJANSKIĬ, V. G., A problem about the illumination of the boundary of a convex body, *Bull. Moldavian Branch Acad. Sci. USSR*, No. 10 (1960), 79–86.

[8] Davis, C., Theory of positive linear dependence, *Amer. J. Math.* **76** (1954), 733–746. *MR* **16**-211

[9] Buchman, E. O. and Valentine, F. A., A characterization of convex surfaces which are L-sets, *Proc. Amer. Math. Soc.* **40** (1973), 235–239. *MR* **47** # 9420

[10] Soltan, P. S., Extremal problems on convex sets, Izdat. Shtiintsa, Kishinev, 1976 (in Russian). *MR* **58** # 18164

(Received September 5, 1990)

MATHEMATICAL INSTITUTE OF THE
MOLDAVIAN ACADEMY OF SCIENCES
STR. ACADEMIEI 5
KISHINEV 277 028
MOLDOVA

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 33–42

ULJANOV-TYPE IMBEDDING THEOREMS IN THE THEORY OF WEIGHTED POLYNOMIAL APPROXIMATIONS

NGUYEN XUAN KY and NGUYEN QUANG HOA

Introduction

The imbedding of Lipschitz class in the L^p-space into L^q-space was considered firstly by Hardy and Littlewood in the thirties of this century. Later, P. L. Uljanov [1] investigated the imbedding of more general classes, namely the Hölder classes into other classes of functions. This theory was then developed by Storozenko [2], G. Gaimnazorov [3], L. Leindler [4], J. Németh [5], M. Milman [6], [7], A. V. Lapin [8] and others. Recently, a great number of authors concentrated to the theory of approximation by polynomials with weights. Among some of the most interesting results achieved there are direct and converse theorems (Jackson and Bernstein-type theorems). New moduli of continuity of functions were introduced for investigating approximation theorems of that type.

A natural problem therefore is to investigate Uljanov-type imbedding problems for that moduli of continuity and for the corresponding best approximations. The first author [9], [10], [11], [12] considered this problem in the cases of Jacobi and Freud weights. In this paper the case of the weight $W_\alpha(x)=|x|^{\alpha/2}e^{-x^2/2}$ will be considered. In §1 we prove a Nikolskiĭ-type inequality with the weight $W_\alpha(x)$. In §2 some Uljanov-type imbedding theorems will be given.

§ 1. Nikolskiĭ-type inequality with weight $W_\alpha(x)$

Nikolskiĭ-type inequality with weight $W_0(x)=e^{-x^2/2}$ was considered first by C. Markett [14], generalized for the case of exponential weight $e^{-|x|^\alpha}$ ($\alpha>1$) by E. B. Saff [23] and was completed by Névai–Totik [16]. The present authors found that the method applied in [16] can be used to deal with the weight $W_\alpha(x)$, which will be detailed in the following.

1991 *Mathematics Subject Classification*. Primary 41A10; Secondary 41A17.

Key words and phrases. Nikolskiĭ-type inequality, weighted polynomial approximation, imbedding theorems.

First let us give some notations and definitions. Let $L^p(-\infty,\infty)$ be the space of the measurable functions on $(-\infty,\infty)$ with the norm

$$\|f\|_p = \left(\int_{\mathbf{R}} |f(t)|^p dt\right)^{1/p} \qquad (0 < p < \infty) \tag{1}$$
$$\|f\|_\infty = \operatorname*{ess\,sup}_{t\in(-\infty,\infty)} |f(t)|.$$

We denote by $\mathcal{P}_n$ the set of all algebraic polynomials of degree at most n $(n = 0, 1, 2, \ldots)$. Let $L_n^\alpha(x)$ be the orthogonal Laguerre polynomials

$$L_n^\alpha(x) = (n!)^{-1}x^{-\alpha}e^x\left(\frac{d}{dx}\right)^n(x^{\alpha+n}e^{-x}) \qquad (x > 0, n = 0, 1, 2, \ldots) \tag{2}$$

and $\hat{H}_n^\alpha(x)$ be the orthonormal polynomials with respect to the weight $W_\alpha(x)$, that is $\hat{H}_n^\alpha \in \mathcal{P}_n$ and

$$\int_{\mathbf{R}} \hat{H}_n^\alpha(x)\hat{H}_m^\alpha(x)W_\alpha^2(x)dx = \delta_{nm} \qquad (n = 0, 1, 2, \ldots), \tag{3}$$

where δ_{nm} denotes the Kronecker symbol.

In what follow $C(x, y, \ldots)$ always denotes an absolute constant depending only on $x, y, \ldots$.

THEOREM 1. *Let $\alpha > 0$, $1 \leqq q, p \leqq \infty$, then the inequality*

$$\|p_n(x)W_\alpha(x)\|_p \leqq C(p,q)n^{\frac{1}{2}(\frac{1}{p}-\frac{1}{q})}\|p_n(x)W_\alpha(x)\|_q \tag{4}$$

is valid for each $p_n(x) \in \mathcal{P}_n$.

Moreover, the above inequality is sharp in the sense that there exist two sequences $\{\tilde{p}_n(x)\}_{n=1}^\infty$ and $\{p_n^*(x)\}_{n=1}^\infty$ of polynomials satisfying

$$\|\tilde{p}_n(x)W_\alpha(x)\|_p \geqq C_1 n^{\frac{1}{2}(\frac{1}{p}-\frac{1}{q})}\|\tilde{p}_n(x)W_\alpha(x)\|_q, \quad 1 \leqq p < q \leqq \infty \tag{5}$$

and

$$\|p_n^*(x)W_\alpha(x)\|_p \geqq C_2 n^{\frac{1}{2}(\frac{1}{q}-\frac{1}{p})}\|p_n^*(x)W_\alpha(x)\|_q, \quad 1 \leqq q < p \leqq \infty. \tag{6}$$

To prove the theorem, we need the following lemmas.

LEMMA 1. *For $\alpha > -1$ and $p_n(x) \in \mathcal{P}_n$ we have*

$$\int_{\mathbf{R}} (|p_n(x)|W_\alpha(x))^p dx \leqq C \int_{-cn^{1/2}}^{cn^{1/2}} (|p_n(x)|W_\alpha(x))^p dx. \tag{7}$$

PROOF. In [17], Theorem 4.16.2 Névai showed that

$$\int_{\mathbf{R}} |q_n(t)|^p W_\beta(t)dt \leqq 2 \int_{-cn^{1/2}}^{cn^{1/2}} |q_n(t)|^p W_\beta(t)dt,$$

for $\forall\beta > -1$, $q_n \in \mathcal{P}_n$. Put $t := \sqrt{p}x$, $q_n(t) := p_n\left(\frac{t}{\sqrt{p}}\right) = p_n(x)$, $\beta := \alpha p$, then

$$\int_{\mathbf{R}} (|p_n(x)|W_\alpha(x))^p dx = p^{-\frac{(\beta+1)}{2}} \int_{\mathbf{R}} |q_n(t)|^p W_\beta(t)dt \leqq$$

$$\leqq 2p^{-\frac{(\beta+1)}{2}} \int_{-cn^{1/2}}^{cn^{1/2}} |q_n(t)|^p W_\beta(t)dt = C \int_{-cn^{1/2}}^{cn^{1/2}} (|p_n(x)|W_\alpha(x))^p dx.$$

LEMMA 2. *Let $0 < q < p \leqq \infty$, $\alpha > 0$. For any $p_n(x) \in \mathcal{P}_n$ we have* ,

$$\|p_n(x)\chi_{\Delta(1)}(x)\bar{W}_\alpha(x)\|_p \leqq cn^{(\frac{1}{q}-\frac{1}{p})}\|p_n(x)\chi_{\Delta(2)}(x)\bar{W}_\alpha(x)\|_q, \tag{8}$$

where $\Delta(b) := [-b, b]$, $\chi_{\Delta(b)}$ is the characteristic function of $\Delta(b)$, $\bar{W}_\alpha(x) :=$ $= |x|^{\alpha/2}$.

PROOF. In [17], Lemma 6.3.23 and 6.3.24 Névai showed that

$$\inf_{p_n\in\mathcal{P}_n} \frac{1}{|p_n(x)|^q} \int_{\mathbf{R}} |p_n(t)|^q |t|^\beta \chi_{\Delta(2)}(t)dt \geqq c\frac{1}{n}\left(|x| + \frac{1}{n}\right)^\beta, \quad \forall x \in \Delta(1).$$

So, for each $p_n(x) \in \mathcal{P}_n$ we have

$$\int_{\mathbf{R}} |p_n(t)|^q\, |t|^\beta \chi_{\Delta(2)}(t)dt \geqq c\,\frac{1}{n}\,|p_n(x)|^q\left(|x| + \frac{1}{n}\right)^\beta \geqq c\,\frac{1}{n}\,|p_n(x)|^q\,|x|^\beta,$$

$$\forall x \in \Delta(1).$$

Put $\alpha := \beta/2q$ then

$$\int_{\mathbf{R}} |p_n(t)t^{\alpha/2}|^{-q}\chi_{\Delta(2)}(t)dt \geqq c\frac{1}{n}|p_n(x)x^{\alpha/2}|^q, \qquad \forall x \in \Delta(1),$$

that is

$$n^{1/q}\|p_n(x)\bar{W}_\alpha(x)\chi_{\Delta(1)}(x)\|_\infty \leqq C\|p_n(x)\bar{W}_\alpha(x)\chi_{\Delta(2)}(x)\|_q.$$

For $0 < q < p \leqq \infty$

$$\|p_n(x)\chi_{\Delta(1)}(x)\tilde{W}_\alpha(x)\|_p = \||p_n(x)|^{(p-q+q)}\chi_{\Delta(1)}(x)\tilde{W}_\alpha(x)\|_1^{1/p} \leqq$$

$$\leqq \|p_n(x)\chi_{\Delta(1)}(x)\tilde{W}_\alpha(x)\|_\infty^{(p-q)/p}\|p_n(x)\chi_{\Delta(1)}(x)\tilde{W}_\alpha(x)\|_q^{q/p} \leqq$$

$$\leqq cn^{(p-q)/pq}\|p_n(x)\chi_{\Delta(2)}(x)\tilde{W}_\alpha(x)\|_q^{(p-q)/p}\|p_n(x)\chi_{\Delta(1)}(x)\tilde{W}_\alpha(x)\|_q^{q/p} \leqq$$

$$\leqq cn^{(\frac{1}{q}-\frac{1}{p})}\|p_n(x)\chi_{\Delta(2)}(x)\tilde{W}_\alpha(x)\|_q.$$

Proof of Theorem 1. First we prove (4).

Case (a): $1 \leqq p < q \leqq \infty$. Applying Lemma 1 and the Hölder inequality we have

$$\|p_n(x)W_\alpha(x)\|_p = \left(\int_{\mathbf{R}} (|p_n(x)|W_\alpha(x))^p dx\right)^{1/p} \leqq$$

$$\leqq C\left(\int_{-cn^{1/2}}^{cn^{1/2}} (|p_n(x)|W_\alpha(x))^p dx\right)^{1/p} \leqq$$

$$\leqq \left(C\int_{-cn^{1/2}}^{cn^{1/2}} 1dx\right)^{(1/p-1/q)}\left(C\int_{-cn^{1/2}}^{cn^{1/2}} (|p_n(x)|W_\alpha(x))^q dx\right)^{1/q} \leqq$$

$$\leqq cn^{\frac{1}{2}(\frac{1}{p}-\frac{1}{q})}\left(\int_{-\infty}^{\infty} (|p_n(x)|W_\alpha(x))^q dx\right)^{1/q} = cn^{\frac{1}{2}(\frac{1}{p}-\frac{1}{q})}\|p_n(x)W_\alpha(x)\|_q.$$

Case (b): Let now $1 \leqq q < p \leqq \infty$, and

$$\Delta := [-cn^{1/2}, cn^{1/2}], \quad \Delta^* := [-2cn^{1/2}, cn^{1/2}],$$

then by (7)

$$\|p_n(t)W_\alpha(t)\|_p \leqq C\|p_n(t)W_\alpha(t)\chi_\Delta(t)\|_p.$$

It was proved by A. L. Levin and D. S. Lubinsky [21], [22] that for given $C > 0$ there are a positive integer $L = L(a, M)$ and polynomials $Q_{Ln}(x) \in \mathcal{P}_{Ln}$ such that

$$Q_{Ln}(x) \sim e^{-x^2/2} \quad \text{for} \quad -2cn^{1/2} \leqq x \leqq 2cn^{1/2}.$$

Let $t = cn^{1/2}x, \bar{p}_n(x) = p_n(t), \bar{Q}_{Ln}(x) = Q_{Ln}(t), \bar{\tilde{W}}_\alpha(x) = \tilde{W}_\alpha(t)$ thus

$$\|p_n(t)W_\alpha(t)\chi_\Delta(t)\|_p \sim \|p_n(t)Q_{Ln}(t)\tilde{W}_\alpha(t)\chi_\Delta(t)\|_p \leqq$$

$$\leqq cn^{1/2p}\|\bar{p}_n(x)\bar{Q}_{Ln}(x)\bar{\tilde{W}}_\alpha(x)\chi_{\Delta(1)}(x)\|_p \leqq$$

$$\leqq cn^{1/2p}n^{(\frac{1}{q}-\frac{1}{p})}\|\bar{p}_n(x)\bar{Q}_{Ln}(x)\bar{W}_\alpha(x)\chi_{\Delta(2)}(x)\|_q =$$
$$= cn^{1/2p}n^{-1/2q}n^{(\frac{1}{q}-\frac{1}{p})}\|p_n(t)Q_{Ln}(t)\bar{W}_\alpha(t)\chi_{\Delta^*}(t)\|_q \sim$$
$$\sim cn^{\frac{1}{2}(\frac{1}{q}-\frac{1}{p})}\|p_n(t)W_\alpha(t)\chi_{\Delta^*}(t)\|_q \leqq$$
$$\leqq cn^{\frac{1}{2}(\frac{1}{q}-\frac{1}{p})}\|p_n(t)W_\alpha(t)\|_q.$$

Now we return to the proof of sharpness of the inequality (4). Put

$$\bar{p}_n(x) := \{\hat{H}_n^\alpha(x) - \hat{H}_{n-2}^\alpha(x)\}, \quad n \geqq 2.$$

In [15] 2.21 it was showed that

$$\|\bar{p}_n(x)W_\alpha(x)\|_p \sim n^{\frac{1}{2p}-\frac{1}{4}} \quad \text{for each } 1 \leqq p \leqq \infty.$$

Consequently,

$$\frac{\|\bar{p}_n(x)W_\alpha(x)\|_p}{\|\bar{p}_n(x)W_\alpha(x)\|_q} \sim n^{\frac{1}{2}(\frac{1}{p}-\frac{1}{q})},$$

which proves (5). Now put

$$p_n^*(x) := \begin{cases} L_m^{\alpha+3/2}(x^2), & n = 2m, m = 0, 1, 2, \dots \\ L_m^{\alpha+5/2}(x^2)x, & n = 2m+1, m = 0, 1, 2, \dots \end{cases}$$

then

$$\|p_{2m}^*(x)W_\alpha(x)\|_p = \left(\int_{\mathbf{R}} \left| L_m^{\alpha+\frac{3}{2}}(x^2)e^{-x^2/2}|x|^{\alpha/2}\right|^p dx\right)^{1/p} =$$

$$= \left(2\int_0^\infty \left| L_m^{\alpha+\frac{3}{2}}(\xi)e^{-\xi/2}\xi^{\frac{\alpha}{2}-\frac{1}{2p}}\right|^p d\xi\right)^{1/p}.$$

Using (2.9) of [14] we have

$$\|p_n^*(x)W_\alpha(x)\|_p \sim (2m)^{\frac{3}{2}-\frac{1}{2p}} = n^{\frac{3}{2}-\frac{1}{2p}},$$

if n is even. Analogously,

$$\|p_n^*(x)W_\alpha(x)\|_p \sim (2m+1)^{2-\frac{1}{2p}} = n^{2-\frac{1}{2p}},$$

if n is odd. So, for each $1 \leqq p, q \leqq \infty$

$$\frac{\|p_n^*(x)W_\alpha(x)\|_p}{\|p_n^*(x)W_\alpha(x)\|_q} \sim n^{\frac{1}{2}(\frac{1}{q}-\frac{1}{p})},$$

from which (6) follows. This completes the proof of Theorem 1.

By the similar way and using approximating polynomials given by Levin and Lubinsky [21] we can prove Nikolskiĭ-type inequality for the more general weight $W_{\alpha,\beta}(x) = |x|^{\alpha/2}e^{-|x|^\beta}$ $(\alpha, \beta > 0)$. More precisely, we have the following theorem:

THEOREM 2. *Let $0 \leqq p, q \leqq \infty$. Then for each $p_n(x)$ the following inequality is valid:*

$$\|p_n(x)W_{\alpha,\beta}(x)\|_p \leqq C N_n(\beta, p, q)\|p_n(x)W_{\alpha,\beta}(x)\|_q. \tag{9}$$

Here $C = C(p,q)$, $N_n(\beta,p,q)$ is defined by Névai–Totik [16]*:*

$$N_n(\beta,p,q) = \begin{cases} n^{(1-\frac{1}{\beta})(\frac{1}{p}-\frac{1}{q})}, & p \leqq q, \\ n^{(1-\frac{1}{\beta})(\frac{1}{q}-\frac{1}{p})}, & p > q \text{ and } \beta > 1, \\ (\log(n+1))^{(\frac{1}{q}-\frac{1}{p})}, & p > q \text{ and } \beta = 1, \\ 1, & p > q \text{ and } 0 < \beta < 1. \end{cases} \tag{10}$$

§ 2. Imbedding theorems with modulus of continuity

Before presenting imbedding theorems let us give some definitions.

Recalling that $\{\hat{H}_k^\alpha\}_{k=0}^\infty$ is the class of orthonormal polynomials with respect to the weight $W_\alpha^2(x)$, let

$$\mathcal{H}_k^\alpha(x) := \hat{H}_k^\alpha(x)W_\alpha(x), \qquad k = 0,1,2,\ldots.$$

For $f(x) \in L^p(-\infty,\infty)$, let us denote the n-th best approximation of $f(x)$ by the system $\{\mathcal{H}_k^\alpha\}_{k=0}^\infty$ in the following way:

$$E_n(W_\alpha, f)_p := \inf_{c_k \in \mathbf{R}} \|f(x) - \sum_{k=0}^n c_k \mathcal{H}_k^\alpha(x)\|_p.$$

For a given decreasing sequence of real numbers tending to zero $a = (a_n) := (a_n \downarrow 0)$, let

$$E(a, W_\alpha, p) := \{f \in L^p(-\infty,\infty) : E_n(W_\alpha, f) \leqq c(f)a_n, \quad n = 0,1,2,\ldots\}.$$

We shall use the following moduli of continuity and K-functional

$$\omega(f,\delta)_p := \sup_{0<h\leqq\delta} \|f(x+h) - f(x)\|_p, \quad f \in L^p(-\infty,\infty).$$

Furthermore, let

$$0 < A < B < \infty, \quad f(x)W_\alpha(x) \in L^p(-\infty,\infty).$$

Define

$$\tilde{\omega}_{A,B}(f, W_\alpha, \delta)_p := \sup_{0<h\leqq\delta} \|(f(x+h) - f(x))W_\alpha(x)\chi_{(-\infty,B)}(x)\|_p +$$

$$+ \sup_{0<h\leqq\delta} \|(f(x+h)-f(x))W_\alpha(x)\chi_{(A,\infty)}(x)\|_p,$$

and

$$K(W_\alpha, f, \delta)_p := \inf_{g\in W}\{\|(f-g)W_\alpha\|_p + \delta\|g'W_\alpha\|_p\},$$

where W is the set of all locally absolutely continuous even functions g on $(-\infty,\infty)$, for which

$$W_\alpha g \in L^p(-\infty,\infty), \quad W_\alpha g' \in L^p(-\infty,\infty).$$

A nondecreasing continuous function Ω on [0, 1] is called a modulus of continuity if

$$\Omega(0)=0, \quad \Omega(\delta_1+\delta_2)\leqq\Omega(\delta_1)+\Omega(\delta_2) \quad (0\leqq\delta_1\leqq\delta_2\leqq\delta_1+\delta_2\leqq 1).$$

For a modulus of continuity Ω and $1\leqq p\leqq\infty$, $0<A<B<\infty$, let

$$H_p^{\Omega,\tilde\omega} = H_p^{\Omega,\tilde\omega_{A,B}} := \{f\in L^p(-\infty,\infty): \tilde\omega_{A,B}(f,W_\alpha,\delta)_p \leqq c(f)\Omega(\delta), \delta>0\}.$$

By Theorem 2, [10] we have that if $f\in L^p$ is an even function then

$$E_n(W_\alpha, f)_p \leqq c\tilde\omega_{A,B}(f, W_\alpha, n^{-1/2})_p. \tag{11}$$

From Theorem 1 [12] and inequality (4) we have

THEOREM 3. *Let $f(x)\in L^p(-\infty,\infty)$, if for some q $(p<q<\infty)$*

$$\sum_{n=1}^{\infty} n^{\frac{1}{2}(\frac{q}{p}-1)-1} E_n^q(W_\alpha, f)_p < \infty \tag{12}$$

then $f(x)\in L^q(-\infty,\infty)$.

This theorem gives a sufficient condition for imbedding $E(a,W_\alpha,p)$ into $L^q(-\infty,\infty)$. We may ask whether the condition (12) is also necessary. The following theorem gives the affirmative answer to this question in the case of even functions with a complementary condition on the sequence (a_n).

THEOREM 4. *Let $\bar E(a,W_\alpha,p)$ be the set of all even functions $f\in E(a,W_\alpha,p)$. Let $a=(a_k)_{k=0}^\infty$ be a sequence of positive numbers decreasing to zero satisfying*

$$na_n \leqq cma_m, \qquad n\leqq m. \tag{13}$$

Let $1\leqq p<q<\infty$, then the necessary condition for the imbedding $\bar E(a,W_\alpha,p)$ into $L^q(-\infty,\infty)$ is that

$$\sum_{n=1}^{\infty} n^{\frac{1}{2}(\frac{q}{p}-1)-1} a_n^q < \infty.$$

PROOF. Suppose that $\sum_{n=1}^{\infty} n^{\frac{1}{2}(\frac{q}{p}-1)-1} a_n^q = \infty$ then as a special case of [10], Lemma 3, there exists a function $f_0(x) \in L^p[0,1]$ having the following properties:

(14) $$f_0(x) = 0, \qquad \text{if } x \in [2^{-1/2}, 1],$$

(15) $$\int_0^h |f_0(x)|^p dx \leqq c a_{2^k}^p, \qquad \text{if } 0 < h \leqq 2^{-(k+2)/2},\ k = 1, 2, \dots,$$

(16) $$\omega(f_0, 2^{-k/2})_p \leqq c a_{2^k},$$

(17) $$f_0(x) \notin L^q[0,1].$$

We define

$$f_1(x) := \begin{cases} f_0(x), & \text{if } 2^{-1/2} \leqq |x| \leqq 1, \\ 0, & \text{elsewhere,} \end{cases}$$

so $f_1(x)$ is an even function.

Now we estimate $\tilde{\omega}_{2,3}(f_1, W_\alpha, \delta)_p$. For $h \leqq 1 - 2^{-1/2}$ we have

$$I_1(h) := \int_{-\infty}^{3} |f_1(x+h) - f_1(x)|^p W_\alpha^p(x) dx = \int_{-\infty}^{-h} + \int_{-h}^{0} + \int_{0}^{1-h} + \int_{1-h}^{3} =$$

$$= \int_{-h}^{0} + \int_{0}^{1-h} = 2\Big(\int_0^h |f_0(x)|^p W_\alpha^p(x) dx + \int_0^{1-h} |f_0(x+h) - f_0(x)|^p W_\alpha^p(x) dx \Big) \leqq$$

$$\leqq 2d \Big(\int_{2^{-1/2}}^{h} |f_0(x)|^p dx + \int_{2^{-3/2}}^{1-h} |f_0(x+h) - f_0(x)|^p dx \Big),$$

where $d := \max_{2^{-3/2} \leqq x \leqq 1} W_\alpha^p(x)$. By (15) and (16) we get

$$I_1(h) \leqq c a_{2^k}^p \quad \text{if} \quad h \leqq 2^{-(k+2)/2}.$$

Similarly,

$$I_2(h) := \int_2^{\infty} |f_0(x+h) - f_0(x)|^p W_\alpha^p(x) dx \leqq c a_{2^k}^p$$

if $h \leqq 2^{-(k+2)/2}$, so

$$\tilde{\omega}_{2,3}(f_1, W_\alpha, 2^{-k/2})_p \leqq c a_{2^k}.$$

From (11) it follows that

$$E_{2^k}(W_\alpha, f_1)_p \leqq c\tilde{\omega}_{2,3}(f_1, W_\alpha, 2^{-k/2})_p =$$

$$= 2c\tilde{\omega}_{2,3}(f_1, W_\alpha, 2^{-k/2})_p \leqq ca_{2^k}.$$

Since $na_n \leqq cma_m$, $n \leqq m$ we get $E_n(W_\alpha, f_1)_p \leqq ca_n$, that is $f_1(x) \in \in \tilde{E}(a, W_\alpha, p)$, which, together with (17), completes the proof of the theorem.

We remark that not only for $A = 2$, $B = 3$ but also for arbitrary constants A, B we can construct a function $f \notin L^q(-\infty, \infty)$ such that

(18) $$\tilde{\omega}_{A,B}(f, 2^{-k/2}) \leqq ca_{2^k}.$$

THEOREM 5. *Let*

$$\tilde{H}_p^{\Omega,\tilde{\omega}} := \{f \in H_p^{\Omega,\tilde{\omega}} : f \text{ even}\}$$

and let $1 \leqq p < q < \infty$. *Then the necessary and sufficient condition for* $\tilde{H}_p^{\Omega,\tilde{\omega}} \subset \subset L^q(-\infty, \infty)$ *is that*

(19) $$\sum_{n=1}^{\infty} n^{\frac{1}{2}(\frac{q}{p}-1)-1} \Omega^q(n^{-1/2}) < \infty.$$

PROOF. a) Sufficiency. Let $f(x) \in \tilde{H}_p^{\Omega,\tilde{\omega}}$, and suppose that (19) holds, then, by (11), we have

$$\sum_{n=1}^{\infty} n^{\frac{1}{2}(\frac{q}{p}-1)-1} E_n^q(W_\alpha, f)_p < \infty,$$

therefore, from Theorem 4 it follows that $f(x) \in L^q(-\infty, \infty)$.

b) Necessity. Put $a_n := \Omega(n^{-1/2})$, $n = 1, 2, \ldots$. It is easy to see that a_n satisfies Condition (13), so the necessary part can be proved similarly to the proof of the necessary part in Theorem 4.

REFERENCES

[1] UL'JANOV, P. L., The embedding of certain classes H_p^ω of functions, *Izv. Akad. Nauk SSSR Ser. Mat.* **32** (1968), 649–686 (in Russian). *MR* **37** #6749

[2] STOROŽENKO, È. A., Necessary and sufficient conditions for the imbedding of certain classes of functions, *Izv. Akad. Nauk SSSR Ser. Mat.* **37** (1973), 386–398 (in Russian). *MR* **48** #12029

[3] GAĬMNAZAROV, G., Teorema vloženija dlja $L_p(-\infty, \infty)$ klassov funkcij, *Izv. Vysš. Učebn. Zaved. Matematika* **4(119)** (1972), 44–54.

[4] LEINDLER, L., On imbedding of classes of functions, *Anal. Math.* **5** (1979), 51–65. *MR* **81j**:46038

[5] NÉMETH, J., Necessary and sufficient conditions for imbedding of classes of functions, *Acta Sci. Math. (Szeged)* **40** (1978), 317–326. *MR* **81k**:42013

[6] MILMAN, M., Embedding of rearrangement invariant spaces in Lorentz spaces, *Acta Math. Acad. Sci. Hungar.* **30** (1977), 253–258. *MR* **57** #3837

[7] MILMAN, M., An inequality for generalized modulus of continuity, *Notas Mat.* **9** (1977), 7 pp.

[8] LAPIN, A. V., Imbedding theorems for generalized Hölder classes of one variables, *Anal. Math.* **11** (1985), 29–54.

[9] KY, N. X., Approximation of function, Doctorial Thesis, Budapest, 1985.

[10] KY, N. X., On imbedding theorems for weighted polynomial approximation and modulus of continuity of functions (to appear).

[11] KY, N. X., On an imbedding theorem (to appear).

[12] KY, N. X., Necessary and sufficient condition for imbedding in the theory of algebraic approximation (to appear).

[13] JOÓ, I. and KY, N. X., Answer to a problem of Paul Turán, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **31** (1988), 229–241. *MR* **90h**:41006

[14] MARKETT, C., Nikolskii-type inequalities for Laguerre and Hermite expansions, *Functions, series, operators*, Vol. I, II (Budapest, 1980), Coll. Math. Soc. János Bolyai, **35**, North-Holland, Amsterdam–New York, 1983, 811–834. *MR* **85h**: 42031

[15] HOA, N. Q., On a Nikolskiĭ-type inequality, *Publ. Math. Debrecen* (to appear).

[16] NÉVAI, P. and TOTIK, V., Sharp Nikolskiĭ inequalities with exponential weights, *Anal. Math.* **13** (1987), 261–267. *MR* **89h**:41034

[17] NÉVAI, P., Orthogonal polynomials, *Mem. Amer. Math. Soc.* **18** (1979), no. 213, 1–185. *MR* **80k**:42025

[18] NÉVAI, P. and FREUD, G., Orthogonal polynomials and Christoffel functions. A case study, *J. Approx. Theory* **48** (1986), 167 pp. *MR* **88b**:42032

[19] FREUD, G., On direct and converse theorems in the theory of weighted polynomial approximation, *Math. Z.* **126** (1972), 123–134. *MR* **46** #7770

[20] FREUD, G., On weighted L_1-approximation by polynomials, *Studia Math.* **46** (1973), 125–133. *MR* **52** #11421

[21] LEVIN, A. L. and LUBINSKY, D. S., Canonical products and the weights $\exp(-|x|^\alpha)$, $\alpha > 1$, with applications, *J. Approx. Theory* **49** (1987), 149–169. *MR* **88f**: 41011

[22] LEVIN, A. L. and LUBINSKY, D. S., Weights on the real line that admit good relative polynomial approximation, with applications, *J. Approx. Theory* **49** (1987), 170–195. *MR* **88f**:41012

(Received August 2, 1990)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 43–45

AN INEQUALITY RELATED TO MINKOWSKI'S

H. ALZER

The classical Minkowski inequality states: *If* (a_n) *and* (b_n) $(n = 1, \dots, k)$ *are non-negative real numbers, and* $p > 1$, *then*

$$\left[\sum_{n=1}^{k}(a_n+b_n)^p\right]^{1/p} \leqq \left(\sum_{n=1}^{k} a_n^p\right)^{1/p} + \left(\sum_{n=1}^{k} b_n^p\right)^{1/p}. \tag{1}$$

If $0 < p < 1$, *then inequality* (1) *must be reversed. Equality holds if and only if* (a_n) *and* (b_n) *are proportional.* Proofs as well as extensions of this important result can be found, for instance, in [1, pp. 30–33] and [2, pp. 55–57].

In this note we present an inequality for infinite series which is related to Minkowski's theorem. Roughly speaking our aim is to compare the sum $\left(\sum_{n=1}^{\infty} a_n^p\right)^{1/p} + \left(\sum_{n=1}^{\infty} b_n^p\right)^{1/p}$ with $\sum_{n=1}^{\infty}(a_n+b_n)$ instead of $\left[\sum_{n=1}^{\infty}(a_n+b_n)^p\right]^{1/p}$. Under the assumption that there exists a positive number c with

$$ca_n^p \geqq \sum_{\nu=n+1}^{\infty} a_\nu^p \quad \text{and} \quad cb_n^p \geqq \sum_{\nu=n+1}^{\infty} b_\nu^p \quad \text{for all} \quad n \geqq 1$$

we will find the best possible value $K(c,p)$ such that the inequality

$$K(c,p)\sum_{n=1}^{\infty}(a_n+b_n) \leqq \left(\sum_{n=1}^{\infty} a_n^p\right)^{1/p} + \left(\sum_{n=1}^{\infty} b_n^p\right)^{1/p}$$

holds for $p > 1$, and that the reversed inequality holds for $0 < p < 1$.

THEOREM. *Let* p *be a positive real number and let* $\sum_{n=1}^{\infty} a_n^p$ *and* $\sum_{n=1}^{\infty} b_n^p$ *be two convergent series of positive terms. Further, let* $c > 0$ *be a real number such that*

$$ca_n^p \geqq \sum_{\nu=n+1}^{\infty} a_\nu^p \quad \textit{and} \quad cb_n^p \geqq \sum_{\nu=n+1}^{\infty} b_\nu^p \tag{2}$$

1980 *Mathematics Subject Classification* (1985 Revision). Primary 26D15.
Key words and phrases. Minkowski's inequality, infinite series.

for all integers $n \geqq 1$. *Then we have for* $p > 1$:

$$(3) \qquad [(c+1)^{1/p} - c^{1/p}] \sum_{n=1}^{\infty} (a_n + b_n) \leqq \Big(\sum_{n=1}^{\infty} a_n^p\Big)^{1/p} + \Big(\sum_{n=1}^{\infty} b_n^p\Big)^{1/p}.$$

If $0 < p < 1$, *then inequality* (3) *must be reversed. Equality holds in both cases if and only if* $a_n = \left(\frac{c}{1+c}\right)^{(n-1)/p} a_1$ *and* $b_n = \left(\frac{c}{1+c}\right)^{(n-1)/p} b_1$ *for all* $n \geqq 1$.

PROOF. Let $p > 1$ and $y > 0$. Since the function

$$f_p(x, y) = (x + y)^{1/p} - x^{1/p}$$

is strictly decreasing on $(0, \infty)$ with respect to x we conclude from (2) for all $n \geqq 1$:

$$(4) \qquad f_p(ca_n^p, a_n^p) \leqq f_p\Big(\sum_{\nu=n+1}^{\infty} a_\nu^p, a_n^p\Big)$$

and

$$(5) \qquad f_p(cb_n^p, b_n^p) \leqq f_p\Big(\sum_{\nu=n+1}^{\infty} b_\nu^p, b_n^p\Big).$$

Adding (4) and (5) yields for $n \geqq 1$:

$$(6) \qquad \begin{aligned} &[(c+1)^{1/p} - c^{1/p}](a_n + b_n) \leqq \\ &\leqq \Big(\sum_{\nu=n}^{\infty} a_\nu^p\Big)^{1/p} - \Big(\sum_{\nu=n+1}^{\infty} a_\nu^p\Big)^{1/p} + \Big(\sum_{\nu=n}^{\infty} b_\nu^p\Big)^{1/p} - \Big(\sum_{\nu=n+1}^{\infty} b_\nu^p\Big)^{1/p}. \end{aligned}$$

Summing on n leads to

$$(7) \qquad [(c+1)^{1/p} - c^{1/p}] \sum_{n=1}^{\infty} (a_n + b_n) \leqq \Big(\sum_{n=1}^{\infty} a_n^p\Big)^{1/p} + \Big(\sum_{n=1}^{\infty} b_n^p\Big)^{1/p}.$$

If we assume $0 < p < 1$, then $f_p(x, y)$ is strictly increasing on $(0, \infty)$ with respect to x. Hence, in the inequalities (4)–(7) we have to replace the sign "$\leqq$" by "$\geqq$". We remark that for $0 < p < 1$ the convergence of $\sum_{n=1}^{\infty} (a_n + b_n)$ follows from the assumption that the series $\sum_{n=1}^{\infty} a_n^p$ and $\sum_{n=1}^{\infty} b_n^p$ are convergent.

A simple calculation reveals: Let $p > 0$; if $a_n = (\frac{c}{1+c})^{(n-1)/p} a_1$ and $b_n = (\frac{c}{1+c})^{(n-1)/p} b_1$ for $n \geqq 1$, then the sign of equality holds in (2) as well as in (3). Finally, we assume that equality holds in (3). Since $x \mapsto f_p(x,y)$ is strictly monotonic we conclude from (4) and (5):

$$(8) \qquad ca_n^p = \sum_{\nu=n+1}^{\infty} a_\nu^p \quad \text{and} \quad cb_n^p = \sum_{\nu=n+1}^{\infty} b_\nu^p$$

for all $n \geqq 1$. From the first identity of (8) we obtain

$$ca_n^p = a_{n+1}^p + ca_{n+1}^p$$

and therefore

$$\prod_{n=1}^{k-1} \left(\frac{c}{1+c}\right)^{1/p} = \prod_{n=1}^{k-1} \frac{a_{n+1}}{a_n},$$

i.e.

$$a_k = \left(\frac{c}{1+c}\right)^{(k-1)/p} a_1, \quad k \geqq 1.$$

From the second identity of (8) we get

$$b_k = \left(\frac{c}{1+c}\right)^{(k-1)/p} b_1, \quad k \geqq 1.$$

This completes the proof of the Theorem.

REFERENCES

[1] Hardy, G. H., Littlewood, J. E. and Pólya, G., *Inequalities*, 2nd ed., Cambridge Univ. Press, Cambridge, 1952. *MR* **13**–727

[2] Mitrinović, D. S., *Analytic inequalities*, Die Grundlehren der math. Wissenschaften, Band 165, Springer–Verlag, New York–Berlin, 1970. *MR* **43**#448

(Received August 14, 1990)

DEPARTMENT OF MATHEMATICS AND ASTRONOMY
UNIVERSITY OF SOUTH AFRICA
ZA–0001 PRETORIA
SOUTH AFRICA

Present address:

MORSBACHER STR. 10
D–51545 WALDBRÖL
FEDERAL REPUBLIC OF GERMANY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 47–54

BILINEAR FORM OF THE ERROR TERM OF BUCHSTAB'S ITERATION SIEVE

S. SALERNO and A. VITOLO

1. Introduction

Let $\mathcal{A}$ be a finite sequence of integers, and

$$\mathcal{A}_d = \{n \in \mathcal{A} \mid n \equiv 0(\operatorname{mod} d)\}.$$

Choose a suitable approximation X of $|A|$ and suppose that

$$|\mathcal{A}_d| = \frac{\omega(d)}{d} X + r(\mathcal{A}, d), \tag{1.1}$$

where $\omega(d)$ is a multiplicative function such that $\omega(d) < d$ for any d and $R(\mathcal{A}, d)$ is an error term, which has to be small on average.

Under the assumption (1.1) one would get information about the quantity of the primes in $\mathcal{A}$. A way to approach this question is given by the sieve methods. Unfortunately, it has been proved (see [2]) that no sieve method is able to detect primes in $\mathcal{A}$ because of parity phenomenon, but indeed it is a reasonable purpose to search for the almost primes in $\mathcal{A}$, e.g. the integers of $\mathcal{A}$ with few prime factors.

Define the shifting function

$$S(\mathcal{A}, \mathcal{B}, z) = \#\{a \in \mathcal{A} \mid (a, P(z)) = 1\},$$

where $\mathcal{B}$ is a set of primes and

$$P(z) = \prod_{\substack{p<z \\ p \in \mathcal{B}}} p.$$

We assume that

$$-L \leqq \sum_{w \leqq p \leqq z} \frac{\omega(p)}{p} \log p - k \log \frac{z}{w} < A_1, \quad 2 \leqq w \leqq z, \tag{1.2}$$

1991 *Mathematics Subject Classification.* Primary 11N35.

Key words and phrases. Sieve methods, bilinear forms, distribution in arithmetic progressions, representation of almost primes.

0081-6906/94/\$ 4.00 ©1994 *Akadémiai Kiadó, Budapest*

and in this case we say that the sieve has dimensions k.

Also let

$$V(z)=\prod_{\substack{p<z\\ p\in\mathcal{B}}}\Big(1-\frac{\omega(p)}{p}\Big).$$

The typical sieve results are of the form

$$S(\mathcal{A},\mathcal{B},z)\leqq XV(z)\{F(s)+\varepsilon\}+\sum_{\substack{d<D^2\\ d|P(z)}}|R(\mathcal{A},d)|, \tag{1.3}$$

and

$$S(\mathcal{A},\mathcal{B},z)\geqq XV(z)\{f(s)-\varepsilon\}-\sum_{\substack{d<D^2\\ d|P(z)}}|R(\mathcal{A},d)|, \tag{1.4}$$

where $s=(\log D^2)/(\log z)$ and F,f are non-negative functions such that

$$F(s)=1+O(e^{-s}),\quad f(s)=1-O(e^{-s}) \tag{1.5}$$

as $s\longrightarrow+\infty$.

We shall have an effective sieve when we can find a value of D^2 such that

$$f(s)>0 \tag{1.6}$$

and

$$\sum_{\substack{d<D^2\\ d|P(z)}}|R(\mathcal{A},d)|=o(XV(z)). \tag{1.7}$$

The highest value of D^2 for which (1.7) holds is the so-called distribution level of sequence $\mathcal{A}$ in the arithmetical progressions. It turns out that as higher the distribution level is as better sieve results are, because $F(s)$ and $f(s)$ approach their asymptotic value.

However, if one could express the error term in bilinear form, essentially as

$$\sum_{\substack{mn|P(z)\\ m\leqq M\\ m\leqq N}}\alpha_m\beta_n\, r(\mathcal{A},mn) \tag{1.8}$$

where $\alpha(m)$, $\beta(n)$ are suitably bounded coefficients and $MN=D^2$, it may be possible to estimate the above expression by $o(XV(z))$ with a higher D^2 than the distribution level of $\mathcal{A}$ and so to take advantage in the applications.

Iwaniec [6] proved that the bilinear form (1.8) for the error term is available in Rosser's linear sieve, and this has made possible improvements in the twin prime problem (see [4], [9]). In fact, if $\mathcal{A}$ is the sequence

$$\{p+2, p \text{ prime } < x\},$$

the Bombieri–Vinogradov theorem [1] says that the distribution level of $\mathcal{A}$ is $x^{1/2}$, and this is still the best result of this kind, whilst the Bombieri–Friedlander–Iwaniec theorem [3] allows for every A to control the expression (1.8) by $o(x/\log(x)^A)$ with $D^2 = x^{4/7-\epsilon}$.

S. Salerno [8] proved that also in Selberg's sieve we can put the error term of (1.3) and (1.4) in bilinear form. We recall that Selberg's method produces inequalities (1.3) and (1.4) with $F(s)$ and $f(s)$ given by

$$F_0(s) = 1/\sigma(s), \quad f_0(s) = 1 - s^{-k} \int_s^\infty \left\{ \frac{1}{\sigma(t-1)} - 1 \right\} dt^k, \tag{1.9}$$

where $\sigma(s)$ is the continuous solution of the differential-difference equations

$$\begin{aligned} s^{-k}\sigma(s) &= A^{-1} \\ (s^{-k}\sigma(s))' &= -ks^{-k-1}\sigma(s-2) \end{aligned} \tag{1.10}$$

with

$$A = 2^k e^{\gamma k}\Gamma(k+1), \tag{1.11}$$

γ being Euler's constant.

It is well-known that, if we have a pair of functions $F_0(s)$ and $f_0(s)$ in (1.3) and (1.4) satisfying (1.5), we can get a new pair of admissible functions $F_1(s)$, $f_1(s)$ by means of Buchstab's identity

$$S(\mathcal{A}, \mathcal{B}, z) = S(\mathcal{A}, \mathcal{B}, z_1) - \sum_{\substack{z_1 < p < z \\ p \in \mathcal{B}}} S(\mathcal{A}_p, \mathcal{B}, p). \tag{1.12}$$

Such functions $F_1(s)$, $f_1(s)$ could be better than $F_0(s)$ and $f_0(s)$ for some value of s, and so, by letting $F(s) = \min(F_0(s), F_1(s))$ and $f(s) = \max(f_0(s), f_1(s))$ in (1.3) and (1.4), we obtain sharper estimates than the previous ones.

By iterating the above process, we shall arrive to limit functions $F(s)$ and $f(s)$, continuous solutions of the differential-difference equations

$$\begin{aligned} F(s) &= 1/\sigma(s), && s \leq a_k, \\ f(s) &= 0, && s \leq b_k, \\ (s^k F(s))' &= ks^{k-1} f(s-1), && s > a_k, \\ (s^k f(s))' &= ks^{k-1} F(s-1), && s > b_k, \end{aligned} \tag{1.13}$$

where a_k and b_k are called the limits of Selberg's sieve (see [7]).

We shall prove that the bilinear form of the error term is available also for the limit of Selberg's sieve. Next we put a weighted sieve in applicable form with the bilinear expression for the error term. We remark that improvements will be obtained in the applications as soon as techniques able to control an expression of type (1.8) with D^2 higher than the distribution level are developed.

2. Statement of the results

We let $\mathcal{L} = \log X$, c a positive constant which may be different at various occurrences, ε a suitable small positive constant.

Denote by $\tau_q(n)$ the number of representations of the integer n as product of q integer factors.

We will prove the following results.

THEOREM 1. *Let $\mathcal{A}$ verify* (1.1) *and* (1.2). *Let $z < D^2$ and $s = (\log D^2)/\log z$. We have*

$$S(\mathcal{A}, \mathcal{B}, z) \leqq XV(z)\{F(s) + \varepsilon\} + \sum_{l < X^{\varepsilon}} \sum_{\substack{mn | P(z) \\ m \leqq M_l \\ n \leqq N_l}} \alpha_l(m)\beta_l(n)\, r(\mathcal{A}, mn)$$

with $M_l = D/P_l^{1/2}$, $N_l = DP_l^{1/2}$, $\mathcal{L} \leqq P_l \leqq z$, where $|\alpha_l(d)|$, $|\beta_l(d)| < \tau_q(d)$ for some integer q and $F(s)$, $f(s)$ are given by (1.13).

THEOREM 2. *Let $\mathcal{A}$ verify* (1.1) *and* (1.2). *Let $z < D$ and $S = (\log D^2)/\log z$. We have*

$$S(\mathcal{A}, \mathcal{B}, z) \geqq XV(z)\{f(s) - \varepsilon\} - \sum_{l < X^{\varepsilon}} \sum_{\substack{mn | P(z) \\ m \leqq M_l \\ n \leqq N_l}} \alpha_l(m)\beta_l(n)\, r(\mathcal{A}, mn)$$

with $M_l = D/P_l^{1/2}$, $N_l = DP_l^{1/2}$, $\mathcal{L} \leqq P_l \leqq z$, where $|\alpha_l(d)|$, $|\beta_l(d)| < \tau_q(d)$ for some integer q and $F(s)$, $f(s)$ are given by (1.13).

Define

$$g = \sup\{\log(n)/\log(X),\ n \in \mathcal{A}\}$$

the degree of the sequence $\mathcal{A}$.

By $n = P_r$ we mean an integer n with at most r prime factors.

An application to the weighted sieve by Richert's logarithmic weights is given by the following

THEOREM 3. *Assume that $\mathcal{A}$ verify* (1.1) *and* (1.2). *Let* $s = (\log D^2)/\log z$ *and* u, v, α *be real numbers such that*

$$z = X^{1/v} < X^{1/u} < D^2 = X^{\alpha}. \tag{2.1}$$

If r *is a positive integer such that*

$$r > ug - 1 + \frac{k}{f(\alpha v)} \int_u^v F\Big(v\Big(\alpha - \frac{1}{t}\Big)\Big)\Big(1 - \frac{u}{t}\Big)\frac{dt}{t} \tag{2.2}$$

with F *and* f *given by* (1.13), *then we have*

$$\#\{n \in \mathcal{A} \mid n = P_r\} > cXV(z) - \sum_{l<X^{\varepsilon}} \sum_{\substack{mn|P(z)\\ m\leqq M\\ n\leqq N}} \alpha_l(m)\beta_l(n) r(\mathcal{A}, mn)$$

with $MN = D^2$, *where* $|\alpha_l(d)|$, $|\beta_l(d)| < \tau_q(d)$ *for some integer* q.

3. Proof of Theorems

We start from the result of [8], which we write as

$$S(\mathcal{A}, \mathcal{B}, z) \leqq XV(z)\{F_0(s) + \varepsilon\} + \sum_{l<X^{\varepsilon}} \sum_{\substack{mn|P(z)\\ m\leqq D\\ n\leqq D}} \alpha_l(m)\beta_l(n) r(\mathcal{A}, mn) \tag{3.1}$$

with $|\alpha_l(d)|$, $|\beta_l(d)| < \tau_q(d)$ for some integer q and $F_0(s)$ given by (1.9).

We shall show the method which allows us to pass from the upper bound (3.1) to the lower bound of Theorem 2 with $f(s)$ equal to $f_0(s)$ given by (1.9). In order to do this, first we observe that, if z_1 is small, e.g. a power of a logarithm of X, then we have

(3.2)

$$S(\mathcal{A}, \mathcal{B}, z_1) = XV(z_1)\{1 + O(e^{-\tau(\log \tau)})\} + \sum_{l<X^{\varepsilon}} \sum_{\substack{mn|P(z)\\ m\leqq D\\ n\leqq D}} \alpha_l(m)\beta_l(n) r(\mathcal{A}, mn)$$

with $\tau = (\log D^2)/(\log z_1)$ and $|\alpha_l(d)|$, $|\beta_l(d)| < \tau_q(d)$ for some integer q. This is well known as fundamental lemma and can be obtainded for instance by using Salerno's method [8] in the proof of Theorem 7.1 of [5].

Next, we subdivide the interval $[z_1, z]$ in subintervals of the form $[2^{-1}P_i, P_i)$, where $P_i = 2^i z_1$, $1 \leqq i \leqq \log(z/z_1)/\log 2$, and for

$$2^{-1}P_i \leqq p \leqq P_i$$

apply (3.1) to $S(\mathcal{A}_p, \mathcal{B}, p)$ with D^2/P_i instead of D^2.

For such p's we have

$$S(\mathcal{A}_p, \mathcal{B}, p) \leqq \frac{\omega(p)}{p} X V(p)\{F_0(s) + \varepsilon\} + \tag{3.3}$$
$$+ \sum_{l<X^\epsilon} \sum_{\substack{mn|P(p) \\ m\leqq D/\sqrt{P_i} \\ n\leqq D/\sqrt{P_i}}} \alpha_l(m)\beta_l(n) r(\mathcal{A}_p, mn).$$

Then, by application of Buchstab's identity (1.12) with $z_1 = \mathcal{L}$ we deduce from (3.2) and (3.3) that

$$S(\mathcal{A}, \mathcal{B}, z) \geqq X V(z_1)\{1 + O(e^{-\tau(\log\tau)})\} -$$
$$- \sum_{\substack{z_1<p<z \\ p\in\mathcal{B}}} \frac{\omega(p)}{p} X V(p) F_0\Big(\frac{\log D^2/p}{\log p}\Big)\{1+\varepsilon\} -$$
$$- \sum_{l<X^\epsilon} \sum_{\substack{mn|P(z) \\ m\leqq D \\ n\leqq D}} \alpha_l(m)\beta_l(n) r(\mathcal{A}, mn) - \tag{3.4}$$
$$- \sum_{1<i<\mathcal{L}} \sum_{2^{-1}P_i<p<P_i} \sum_{l<X^\epsilon} \sum_{\substack{mn|P(p) \\ m\leqq D/\sqrt{P_i} \\ n\leqq D/\sqrt{P_i}}} \alpha_{l,i}(m)\beta_{l,i}(n) r(\mathcal{A}, mn).$$

First, we deal with the main term of (3.4). We can put it in the following form

$$X\Big\{V(z_1) - \sum_{\substack{z_1<p<z \\ p\in\mathcal{B}}} \frac{\omega(p)}{p} V(p) - \sum_{\substack{z_1<p<z \\ p\in\mathcal{B}}} \frac{\omega(p)}{p} V(p)\Big(F_0\Big(\frac{\log D^2/p}{\log p}\Big) - 1\Big)\Big\}. \tag{3.5}$$

By recalling the recurrence identity

$$V(z_1) - \sum_{\substack{z_1<p<z \\ p\in\mathcal{B}}} \frac{\omega(p)}{p} V(p) = V(z) \tag{3.6}$$

since $z_1 = \mathcal{L}$, from the asymptotic formula

$$V(t)/V(z) = (\log z/\log t)^k (1 + c/\log t) \tag{3.7}$$

we deduce by partial summation that

$$\sum_{\substack{z_1<p<z \\ p\in\mathcal{B}}} \frac{\omega(p)}{p} V(p)\Big(F_0\Big(\frac{\log D^2/p}{\log p}\Big)-1\Big) \cong$$

$$\cong V(z)s^{-k}\int_{z_1}^{z}\Big(F_0\Big(\frac{\log D^2/t}{\log t}\Big)-1\Big)\frac{d}{dt}\Big(\frac{\log D^2/t}{\log t}\Big)^k dt,$$

by changing to variable $\log D^2/\log t$,

$$= V(z)s^{-k}\int_{s}^{T}(F_0(t-1)-1)dt^k$$

and by (1.5)

$$\cong V(z)s^{-k}\int_{s}^{\infty}(F_0(t-1)-1)dt^k. \tag{3.8}$$

Therefore, by (3.5), (3.6) and (3.8), we obtain the right lower bound for the main term

$$XV(z)f_0(s) \tag{3.9}$$

with $f_0(s)$ as in (1.9).

Now we come to the error terms. We have nothing to do about the first one, coming from (3.2), because it is already a bilinear form. What concerns the second one, we observe that

$$\sum_{2^{-1}P_i<p<P_i}\ \sum_{n\leqq D/\sqrt{P_i}} \beta_{l,i}(n)r(\mathcal{A},mnp) = \sum_{n\leqq D/\sqrt{P_i}} \delta_{l,i}(n)r(\mathcal{A},mn) \tag{3.10}$$

where we have set

$$\delta_{l,i} = \chi_i * \beta_{l,i}$$

the convolution product of the characteristic function χ_i of the primes in $[2^{-1}P_i, P_i)$ and the function $\beta_{l,i}$.

Setting $M = D/P_i^{1/2}$ and $N = DP_i^{1/2}$, ν an enumeration of the pairs (l,i), we can put the second error term in the form

$$\sum_{\nu<X^{\epsilon}\mathcal{L}}\ \sum_{\substack{mn|P(z) \\ m\leqq M \\ n\leqq N}} \alpha_\nu(m)\delta_\nu(n)r(\mathcal{A},mn) \tag{3.11}$$

as desired.

From (3.4), (3.9) and (3.11) we deduce the right lower bound with the error term in bilinear form in the first step of Buchstab's iteration on Selberg's sieve. This shows also how the limit process can be carried out in order to prove Theorems 1 and 2 by means of Buchstab's identity.

Finally, Theorem 3 follows from the weighted sieve by Richert's logarithmic weights, as in Theorem 10.2 of [5].

It suffices to observe the inequality

$$\#\{n \in \mathcal{A} \mid n = P_r\} \geqq \sum_{\substack{n \in \mathcal{A} \\ (n.P(z))=1}} \Big[1 - \lambda \sum_{\substack{z \leqq p \leqq y \\ p|n, p \in \mathcal{B}}} \Big(1 - \frac{\log p}{\log y}\Big)\Big] \tag{3.12}$$

when

$$\lambda = 1/(r + 1 - ug),$$

and that the constant which multiplicates $XV(z)$ in the main term of the right-hand sum in (3.12), as it comes out from the estimates of Theorems 1 and 2, is positive if r is an integer satisfying (2.2).

REFERENCES

[1] BOMBIERI, E., Le grand crible dans la théorie analytique des nombres, *Astérisque* **18**, Société Mathématique de France, Paris, 1974. *MR* **51** # 8057

[2] BOMBIERI, E., The asymptotic sieve, *Rend. Accad. Naz. XL* (5) 1/2 (1975/76), 243–269. *MR* **58** # 10799

[3] BOMBIERI, E., FRIEDLANDER, J. B. and IWANIEC, H., Primes in arithmetic progressions to large moduli, *Acta Math.* (to appear).

[4] FOUVRY, E. and GRUPP, F., On the switching principle in sieve theory, *J. Reine Angew. Math.* **370** (1986), 101–126. *MR* **87j**:11092

[5] HALBERSTAM, H. and RICHERT, H. E., *Sieve methods*, London Mathematical Society Monographs, No. 4, Academic Press, London–New York, 1974. *MR* **54** # 12689

[6] IWANIEC, H., A new form of the error term in the linear sieve, *Acta Arith.* **37** (1980), 307–320. *MR* **82d**:10069

[7] IWANIEC, H., LUNE, I. VAN DE and RIELE, H. TE, The limits of Buchstab's iteration sieve, *Nederl. Akad. Wetensch. Indag. Math.* **42** (1980), 409–417. *MR* **82a**:10054

[8] SALERNO, S., Iwaniec's bilinear form of the error term in Selberg's linear sieve.

[9] VITOLO, A. and ZANNIER, U., Sull'esistenza di infiniti primi p tali che $p + 2 = P_2$, *Ricerche Mat.* (to appear).

(Received September 10, 1990)

ISTITUTO DI MATEMATICA FISICA E
INFORMATICA
UNIVERSITÀ DI SALERNO
I–84081 BARONISSI
ITALY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 55–60

ON AN INTERPOLATION THEORETICAL EXTREMAL PROBLEM

P. ERDŐS, J. SZABADOS[1], A. K. VARMA and P. VÉRTESI[1]

Let

$$(1)\qquad (1 \geqq) x_1 > x_2 > \ldots > x_n (\geqq -1)$$

be an arbitrary system of nodes of interpolation, and let

$$l_k(x) = \prod_{\substack{i=1\\ i\neq k}}^{n} \frac{x - x_i}{x_k - x_i} \qquad (k = 1, \ldots, n)$$

be the fundamental polynomials of interpolation. In Erdős [2], the author raised the problem of determining the minimum of

$$(2)\qquad \int_{-1}^{1} \sum_{k=1}^{n} l_k^2(x)\,dx.$$

He conjectured that the minimum is attained when the nodes (1) are the roots of the integral of the Legendre polynomials. This conjecture was disproved in Szabados [6]. A related but somewhat less hopeless problem is to *estimate* the minimum of (2). In [2] it was proved that given an arbitrary $\varepsilon > 0$, for each $n > n_0 = n_0(\varepsilon)$ the value of (2) is always greater than $2 - \varepsilon$. A sharper estimate was announced in Erdős [3] where it was claimed (without proof) that (2) is always greater than $2 - O(\frac{\log n}{n})$.

In this short note we prove a slightly weaker but nevertheless more general estimate. Let

$$w(x) = (1-x)^{\alpha}(1+x)^{\beta} \qquad (|x| < 1, \gamma = \min(\alpha, \beta) > -1)$$

be the Jacobi weight.

THEOREM. *For any system of nodes* (1) *we have*

1991 *Mathematics Subject Classification.* Primary 41A05; Secondary 41A10.
Key words and phrases. Interpolation, fundamental functions, Christoffel function.

[1] Research supported by National Science Foundation of the Hungarian Academy of Sciences Grant No. 1801.

0081-6906/94/$ 4.00 ©1994 *Akadémiai Kiadó, Budapest*

$$(3)\quad \int_{-1}^{1} w(x)\sum_{k=1}^{n} l_k^{2s}(x)dx \geqq \frac{1}{s}\int_{-1}^{1} w(x)dx - \begin{cases} O\Big(\dfrac{\log^2 n}{n}\Big), & if\ \gamma > -1/2, \\ O\Big(\dfrac{\log^3 n}{n}\Big), & if\ \gamma = -1/2, \\ O\Big(\dfrac{\log^6 n}{n^{2+2\gamma}}\Big), & if\ -1 < \gamma < -1/2, \end{cases}$$

$$(s = 1, 2, \dots)$$

where the constant represented by "O" depends only on s.

Hence we get an answer for the original Erdös problem ($s = 1, \alpha = \beta = 0$) with $O(\frac{\log^2 n}{n})$ instead of $O(\frac{\log n}{n})$.

PROOF. We may assume that the nodes (1) are asymptotically uniformly distributed, since otherwise the integral in (3) tends to infinity as $n \to \infty$ (see the proof of Theorem 4 in [2]). More exactly, the following (sharp) estimate holds (with the notation $x_k = \cos \xi_k, k = 1, 2, \dots, n$):

$$(4)\qquad \left|k - \frac{n\xi_k}{\pi}\right| = O(\log^2 n) \qquad (k = 1, 2, \dots, n)$$

(cf. Erdös [1], Theorem 2).

Since $l_k^s(x) \in \Pi_{s(n-1)} \subset \Pi_{sn}$, we have

$$\int_{-1}^{1} w(x) l_k^{2s}(x)dx \geqq \lambda_{sn}(w, x_k) \qquad (k = 1, 2, \dots, n)$$

where $\lambda_{sn}(w, x)$ is the $(sn)^{\text{th}}$ Christoffel function associated with the weight $w(x)$ (see Freud [4], Theorem I.4.1). Hence denoting $y_k = \cos \eta_k$, $k = 1, 2, \dots, sn$ the roots of the $(sn)^{\text{th}}$ Jacobi polynomial associated with $w(x)$ we obtain by the quadrature theorem

$$\int_{-1}^{1} \sum_{k=1}^{n} l_k^{2s}(x)w(x)dx \geqq \sum_{k=1}^{n} \lambda_{sn}(w, x_k) =$$

$$= \frac{1}{s}\int_{-1}^{1} w(x)dx - \frac{1}{s}\sum_{k=1}^{n}\sum_{j=1}^{s}(\lambda_{sn}(w, y_{(k-1)s+j}) - \lambda_{sn}(w, x_k)) \geqq$$

$$\geqq \frac{1}{s}\int_{-1}^{1} w(x)dx - \frac{1}{s}\sum_{k=1}^{n}\sum_{j=1}^{s}|\lambda'_{sn}(w, z_{kj})(y_{(k-1)s+j} - x_k)|$$

$$(z_{kj} \in (y_{(k-1)s+j}, x_k)).$$

Here

$$\left|\eta_i - \frac{i\pi}{n}\right| = O(n^{-1}) \qquad (i = 1, 2, \dots, n)$$

(cf. Szegö [7], Theorem 8.9), whence and by (4)

$$|y_{(k-1)s+j} - x_k| = 2 \sin \frac{|\eta_{(k-1)s+j} - \xi_k|}{2} \sin \frac{\eta_{(k-1)s+j} + \xi_k}{2} =$$

$$= O\Big(\frac{\log^2 n(k + \log^2 n)}{n^2}\Big) \qquad (0 \leqq \xi_k \leqq \pi/2)$$

(of course, a similar estimate holds for $\pi/2 \leqq \xi_k \leqq \pi$). Also, Lemma 2 (formula (23)) of Nevai–Vértesi [5] implies

$$|\lambda_{sn}(z_{kj})| = \begin{cases} O(n^{-2\alpha} \log^{2|2\alpha-1|_+} n) & \text{if } \xi_k \leqq \frac{\log^2 n}{n}, \\ O(n^{-2\alpha} k^{2\alpha-1}) & \text{if } \frac{\log^2 n}{n} \leqq \xi_k \leqq \frac{\pi}{2} \end{cases}$$

where $|a|_+$ is a if $a \geqq 0$ and it is 0 if $a < 0$. Hence

$$\sum_{\xi_k \leqq \pi/2} \sum_{j=1}^{s} |\lambda'_{sn}(w, z_{kj})(y_{(k-1)s+j} - x_k)| =$$

$$= O\Big(\sum_{\xi_k \leqq \frac{\log^2 n}{n}} n^{-2\alpha} \log^{2|2\alpha-1|_+} n \cdot \frac{\log^4 n}{n^2} + \sum_{\frac{\log^2 n}{n} \leqq \xi_k \leqq \frac{\pi}{2}} n^{-2\alpha} k^{2\alpha-1} \frac{k \log^2 n}{n^2}\Big) =$$

$$= O\Big(\frac{\log^{6+2|2\alpha-1|_+} n}{n^{2+2\alpha}} + \frac{\log^2 n}{n^{2+2\alpha}} \sum_{k=1}^{n} k^{2\alpha}\Big) =$$

$$= \begin{cases} O\Big(\frac{\log^2 n}{n}\Big) & \text{if } \alpha > -\frac{1}{2}, \\ O\Big(\frac{\log^3 n}{n}\Big) & \text{if } \alpha = -\frac{1}{2}, \\ O\Big(\frac{\log^6 n}{n^{2+2\alpha}}\Big) & \text{if } -1 < \alpha < -\frac{1}{2}. \end{cases}$$

A similar estimate holds for $\sum\limits_{\pi/2 < \xi_k \leqq \pi}$ but with β instead of α. The Theorem is proved.

REMARKS. For $s = 1$, the main term on the right-hand side of (3) is sharp. This follows from the well-known identity

(5) $$\sum_{k=1}^{n} l_k^2(x) = 1 - \frac{(1 - x^2) P'^2_{n-1}(x)}{n(n-1)} \qquad (n \geqq 2)$$

valid when the nodes (1) are the roots of the integral of the Legendre polynomials (here $P_{n-1}(x)$ is the Legendre polynomial of degree $n-1$ normalized such that $P_{n-1}(1)=1$). Namely, (5) easily yields

$$\int_{-1}^{1} w(x)\sum_{k=1}^{n} l_k^2(x)dx = \int_{-1}^{1} w(x)dx - \begin{cases} O\left(\frac{1}{n}\right), & \text{if } \gamma > -1/2, \\ O\left(\frac{\log n}{n}\right), & \text{if } \gamma = -1/2, \\ O\left(\frac{1}{n^{2+2\gamma}}\right), & \text{if } -1<\gamma<-1/2, \end{cases}$$

and this shows that, apart from log-factors, (3) is sharp when $s=1$.

The situation is not so good when $s \geqq 2$. Namely, in this case taking again the roots of the integrated Legendre polynomials, (5) implies $|l_k(x)| \leqq 1$ $(k=1,2,\dots,n, |x|\leqq 1)$, i.e. $l_k'(x_k)=0$ $(k=2,\dots,n-1)$. Thus using a second degree Taylor expansion about x_k we obtain

$$l_k^{2s}(x) = 1 - s l_k^{2s-1}(y_k) l_k''(y_k)(x-x_k)^2 \geqq$$

$$\geqq 1 - \frac{sn^2}{1-y_k^2}(x-x_k)^2 \geqq 1 - \frac{sn^2}{1-x_k^2}(x-x_k)^2 = 1 - \left(\frac{x-x_k}{\varepsilon_k}\right)^2$$

$$\left(y_k \in (x,x_k),\ |x| \leqq |x_k|,\ \varepsilon_k := \frac{\sqrt{1-x_k^2}}{n\sqrt{s}}, \quad k=2,\dots,n-1\right),$$

by Bernstein's inequality applied to the second derivative of $l_k(x)$. Hence we get by the mean value theorem for integrals

$$\int_{-1}^{1} w(x) l_k^{2s}(x)dx \geqq \int_{I_k} w(x) l_k^{2s}(x)dx \geqq \int_{I_k} w(x)\left[1-\left(\frac{x-x_k}{\varepsilon_k}\right)^2\right]dx =$$

$$= w(z_k)\int_{I_k}\left[1-\left(\frac{x-x_k}{\varepsilon_k}\right)^2\right]dx = \frac{2}{3}w(z_k)\varepsilon_k \geqq c_1 w(x_k)\varepsilon_k$$

$$\left(I_k := \begin{cases} [x_k-\varepsilon_k, x_k] & \text{if } x_k>0, \\ [x_k, x_k+\varepsilon_k] & \text{if } x_k \leqq 0, \end{cases} \quad z_k \in I_k, \quad k=2,\dots,n-1\right)$$

with an absolute constant $c_1>0$. Thus

$$\int_{-1}^{1}\sum_{k=1}^{n} l_k^{2s}(x)dx \geqq \sum_{k=2}^{n-1}\int_{I_k} w(x) l_k^{2s}(x)dx \geqq$$

$$\geqq \frac{c_1}{\sqrt{s}} \sum_{k=2}^{n-1} w(x_k) \frac{\sqrt{1-x_k^2}}{n} \geqq \frac{c_2}{\sqrt{s}} \sum_{k=2}^{n-1} w(x_k)(x_{k-1}-x_{k+1}) \geqq \frac{c_3}{\sqrt{s}} \int_{-1}^{1} w(x)dx,$$

with absolute constants $c_2, c_3 > 0$, which indicates that for $s \geqq 2$, (3) may be far from being sharp.

We conjecture that the above estimates can be carried out for any system of nodes (1), i.e. in the Theorem $1/s$ can be replaced by some $c/\sqrt{s}$. Also, we conjecture that

$$\min \int_{-1}^{1} \sum_{k=1}^{n} l_k^2(x)dx = 2 - \frac{1}{n} - o\left(\frac{1}{n}\right).$$

REMARK. Let

$$\mu_{n,s} = \int_{-1}^{1} \sum_{k=1}^{n} \tilde{l}_k^{2s}(x)dx,$$

where $\bar{l}_k(x)$ are the fundamental polynomials based on the roots of the Legendre polynomials. Then for every $\varepsilon > 0$ there exists an n_0 such that

$$\min \int_{-1}^{1} \sum_{k=1}^{n} l_k^{2s}(x)dx > \mu_{n,s} - \varepsilon.$$

The proof of this statement follows on the lines of [2] (see in particular pp. 243–244).

REFERENCES

[1] ERDŐS, P., On the uniform distribution of the roots of certain polynomials, *Ann. of Math.* (2) **43** (1942), 59-64. *MR* **3**-236

[2] ERDŐS, P., Problems and results on the theory of interpolation. II, *Acta Math. Acad. Sci. Hungar.* **12** (1961), 235-244. *MR* **26** #2779

[3] ERDŐS, P., Problems and results on the convergence and divergence properties of the Lagrange interpolation polynomials and some extremal problems, *Mathematica (Cluj)* **10** (1968), 65-73. *MR* **38** #1437

[4] FREUD, G., *Orthogonal polynomials*, Akadémiai Kiadó–Pergamon Press, Budapest, 1970; Oxford–New York, 1971. (See *MR* **58** #1982 and *Zbl* **169** #80.)

[5] NEVAI, P. and VÉRTESI, P., Mean convergence of Hermite–Fejér interpolation, *J. Math. Anal. Appl.* **105** (1985), 26–58. *MR* 86h:41004

[6] SZABADOS, J., On a problem of P. Erdős, *Acta Math. Acad. Sci. Hungar.* **17** (1966), 155–157. *MR* **33** #1619

[7] SZEGŐ, G., *Orthogonal polynomials*, Fourth edition, American Mathematical Society Colloquium Publications, Vol. 23, American Mathematical Society, Providence, RI, 1975. *MR* **51** #8724

(Received September 25, 1990)

P. Erdős, J. Szabados and P. Vértesi:

MTA MATEMATIKAI KUTATÓINTÉZETE
P.O.BOX 127
H–1364 BUDAPEST
HUNGARY

A. K. Varma:

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF FLORIDA
GAINESVILLE, FL 32611
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 61–66

CHARACTERISING LORENTZ SPACE BY NORM ONE PROJECTIONS

S. FITZPATRICK and B. CALVERT

Abstract

We consider the "normed" version of the Lorentz vector spaces, the indefinite inner product spaces of index or rank of positivity 1. We show the norm is given by an indefinite inner product if (and only if) any two-dimensional subspace containing a timelike vector is the range of a norm one projection. This result corresponds to the Blaschke-Kakutani theorem characterising positive definite inner product spaces among normed linear spaces.

Introduction

In the theory of indefinite inner product spaces, for which the general reference is [4], those of rank of positivity 1 can be considered examples of spaces with a super-additive norm, defined in [2]. In [3] it is shown that the norm is given by an inner product iff it is self polar. In this paper we take the most significant characterisation of definite inner product spaces, in terms of norm one projections [1] and give a corresponding result in this setting. We also give a version in which we suppose only that there are norm one projections onto subspaces containing certain vectors e_i.

DEFINITION. Let X be a real vector space. A timecone C in X is a subset of X which is invariant under addition and multiplication by scalars $k>0$, with $0 \notin C$, $X = C - C$, and C containing no lines $x + \mathbf{R}y$. An element of C is called a timelike vector.

DEFINITION. Let X be a real vector space with timecone C. Let $p: C \to (0,\infty)$ satisfy:

(i) $p(kx) = kp(x)$ for $k > 0$ and $x \in C$, and

(ii) $p(x) + p(y) \leqq p(x+y)$ for x and $y \in C$.

We call p a super-additive (s.a.) norm on C, and (X,p) a super-additive normed linear space, denoted (X,p,C) if we want to label the domain of p.

DEFINITION. Let (X_1, p_1, C_1) and (X_2, p_2, C_2) be s.a. normed linear spaces. Let $L(X_1, X_2)$ denote the vector space of linear operators $A: X_1 \to X_2$. We call $A \in L(X_1, X_2)$ strictly plus [4, p. 154] if there is $\delta > 0$ such that if $p_1(x) \geqq 1$ then $Ax \in C_2$ and $p_2(Ax) \geqq \delta$. Let $B(X_1, X_2)$ be the linear span of the strictly plus operators. Let C be the strictly plus operators.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 46D05.

Key words and phrases. Indefinite inner product, super-additive norm, Lorentz vector space.

PROPOSITION. *$(B(X_1, X_2), p, C)$ is a super-additive normed linear space if $p : C \to (0, \infty)$ is given by $p(A) = \text{glb}\{p_2(Ax) : p_1(x) \geqq 1\}$.*

DEFINITION. Let (X, g) be a real inner product space, i.e. $g : X \times X \to \mathbf{R}$ is symmetric and bilinear. Suppose X has dimension $\geqq 2$, g is nondegenerate and is positive definite on a subspace of dimension 1 but not on any of dimension 2. We will refer to g as having rank of positivity 1, or being a Lorentz inner product [6, p. 140]. We will call (X, g) a Lorentz vector space.

REMARK. Let $v \in X$ with $g(v, v) > 0$, where g is a Lorentz inner product. The set $\{w \in X : g(w, w) > 0, g(w, v) > 0\}$ is a timecone C, as is $-C$. If $p(x) = \sqrt{g(x, x)}$ for $x \in C$, p is a super-additive norm. We will say the s.a. norm $p : C \to (0, \infty)$ is given by g. The following result must be well known.

PROPOSITION. *Let (M, g) be a Lorentz vector space, and let N be a finite dimensional subspace containing a timelike vector. Then there is a linear $P : M \to M$, $P^2 = P$, $P(M) = N$, of super-additive norm 1.*

PROOF. (1) We claim that if $e \in M$ is timelike then $M = \langle e \rangle \oplus \langle e \rangle^\perp$. Here $\langle e \rangle$ is the space spanned by e. For we may assume $g(e, e) = 1$, and then $x = (x - g(x, e)e) + g(x, e)e$ with the first term in $\langle e \rangle^\perp$ and the second in $\langle e \rangle$. If $x \in \langle e \rangle \cap \langle e \rangle^\perp$ then $g(x, x) = 0$ gives $x = 0$.

(2) We claim g is negative definite on $\langle e \rangle^\perp$, e timelike. We cannot have $g(f, f) > 0$ for any $f \in \langle e \rangle^\perp$ for then g would be positive definite on $\langle e, f \rangle$, so g is negative semidefinite on $\langle e \rangle^\perp$. Suppose $x \in \langle e \rangle^\perp$ and $g(x, x) = 0$. Then $0 \geqq g(x - \lambda y, x - \lambda y)$ for all $y \in \langle e \rangle^\perp$ and $\lambda \in \mathbf{R}$. Hence $g(x, y)^2 \leqq g(x, x)g(y, y) = 0$ for all $y \in \langle e \rangle^\perp$. Hence $g(x, z) = 0$ for all $z \in M$ by (1). Hence $x = 0$ since g is nondegenerate.

(3) Let N be k-dimensional, $e_1 \in N$, $g(e_1, e_1) = 1$. We claim there is a basis $(e_1, \ldots, e_k)$ with $g(\sum x_i e_i, \sum x_i e_i) = x_1^2 - (x_2^2 + \ldots + x_k^2)$. Let $(e_2, \ldots, e_k)$ be an orthonormal basis of $\langle e_1 \rangle^\perp \cap N$. Then $g(\sum x_i e_i, \sum x_i e_i) = -\sum_{i=2}^{k} x_i^2 + x_1^2$.

(4) We may define P by $Px = +g(x, e_1)e_1 - \sum_{j=2}^{k} g(x, e_j)e_j$, for $x \in M$. Note P is linear, $P(M) \subseteq N$, and P is the identity on N. By (3), for $x \notin N$, there is a basis $(e_1, \ldots, e_{k+1})$ of $\langle N, x \rangle$ with $g(z, z) = -(z_2^2 + \ldots + z_{k+1}^2) + z_1^2$ for $z = \sum_{i=1}^{k+1} z_i e_i$. Then

$$g(Px, Px) = -\sum_{j=2}^{k} g(x, e_j)^2 + g(x, e_1)^2 \geqq \sum_{j=2}^{k+1} -g(x, e_j)^2 + g(x, e_1)^2 = g(x, x).$$

Hence $p(Px) \geqq p(x)$. □

Main result

By a linear projection we mean a linear function $P: E \to E$ where E is a vector space, with $P^2 = P$, and we say P is onto $P(E)$.

LEMMA 1 [5]. *Let E be a 3-dimensional real vector space. Suppose B is a closed convex subset of E having nonempty interior, with $0 \notin B$ and B containing no lines. Suppose there exists $\{e, f\}$, a linearly independent subset of E, such that for F a two-dimensional subspace containing e or f there is a linear projection P of E onto F with $P(B) \subseteqq B$. Then there is a basis (e_1, e_2, e_3) such that*

$$B = \left\{ \sum x_i e_i : x_1 > 0, x_1^2 - (x_2^2 + x_3^2) \geqq 1 \right\}.$$

COROLLARY 1. *Let (E, p, C) be a 3-dimensional s.a. normed linear space. Suppose there is a linearly independent set $\{e, f\}$ such that for F a two-dimensional subspace containing e or f there is a projection P onto F of s.a. norm 1. Then the s.a. norm on E is given by a unique Lorentz inner product.*

PROOF. Since $E = C - C$, C has nonempty interior as a subset of E, and since $-p$ is convex, p is continuous on the interior of C. Take B to be the closure in E of $\{x \in C : p(x) \geqq 1\}$. Then $\operatorname{int}(B)$ is nonempty since it contains $\operatorname{int}(C)$. B contains no lines since C does not, and likewise $0 \notin B$. By Lemma 1, there is a basis (e_1, e_2, e_3) with $B = \{\sum x_i e_i : x_1 > 0, x_1^2 - x_2^2 - x_3^2 \geqq 1\}$. Let $g(x, y) = x_1 y_1 - x_2 y_2 - x_3 y_3$ where these are their components with respect to (e_1, e_2, e_3). Then for $x \in C$, $g(x, x) = p^2(x)$, while C is one of the two timecones given by g; thus p is given by g. Note g is a Lorentz inner product. Let $\bar{g}$ be an inner product with $\bar{g}(z, z) = p^2(z)$ for z in C. For x, y in C, $\bar{g}(x, y) = p^2(x + y) - p^2(x) - p^2(y) = g(x, y)$. If $x, y \in E$, $x = x_1 - x_2$, $x_i \in C$, and $y = y_1 - y_2$ with $y_i \in C$, and $\bar{g}(x, y) = \bar{g}(x_1 - x_2, y_1 - y_2) = \bar{g}(x_1, y_1) - \bar{g}(x_1, y_2) - \bar{g}(x_2, y_1) + \bar{g}(x_2, y_2) = g(x_1, y_1) - g(x_1, y_2) - g(x_2, y_1) + g(x_2, y_2) = g(x, y)$, and g is unique. □

LEMMA 2. *Let E be a real vector space of dimension $n \geqq 3$. Suppose $L = \{e_2, \dots, e_n\}$ is linearly independent, with $n - 1$ elements, and if F and H are 3-dimensional subspaces each containing at least 2 elements of L then there are inner products g_F on F and g_H on H with $g_F(x, x) = g_H(x, x)$ if $x \in F \cap H$. Then there is a unique inner product g on E whose restriction to a subspace F as above is g_F.*

PROOF. For all $x \in E$ there does exist a three-dimensional space F containing x and two elements of L, and we may define $f: E \to \mathbf{R}$ by $f(x) = g_F(x, x)$.

We expand L to a basis $\{e_1, \dots, e_n\}$. We claim that there exist a_{ij} such that for $x = \sum_{i=1}^{n} x_i e_i$, $f(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$, giving existence and uniqeness by polarization, i.e. $2g(x,y) = g(x+y,x+y) - g(x,x) - g(y,y)$. We suppose $k > 3$ and that for $n < k$ the lemma holds.

For $\{f_1,\dots,f_h\}$ a subset of E, and its linear span $\langle f_1,\dots,f_h\rangle$ having dimension h $(h < k)$, with all but one of the f_i belonging to L, we let $g_{\langle f_1,\dots,f_h\rangle}$ be the inner product on $\langle f_1,\dots,f_h\rangle$ with $g_{\langle f_1,\dots,f_h\rangle}(x,x) = f(x)$ for $x \in \langle f_1,\dots,f_h\rangle$.

Let $x \in E$, $x = \sum_{i=1}^{k} x_i e_i$. Then $x \in \langle x_1e_1 + x_2e_2, e_3, \dots, e_k\rangle$ giving

$$\begin{aligned}
f(x) &= g_{\langle x_1e_1+x_2e_2,e_3,\dots,e_k\rangle}(x,x) = \\
&= g_{\langle x_1e_1+x_2e_2,e_3,\dots,e_k\rangle}(x_1e_1+x_2e_2, x_1e_1+x_2e_2) + \\
&\quad + 2\sum_{h=3}^{k} g_{\langle x_1e_1+x_2e_2,e_3,\dots,e_k\rangle}(x_1e_1+x_2e_2, e_h)x_h + \\
&\quad + \sum_{h,j=3}^{k} g_{\langle x_1e_1+x_2e_2,e_3,\dots,e_k\rangle}(e_j,e_h)x_jx_h = \\
&= g_{\langle x_1e_1+x_2e_2,e_3,e_4\rangle}(x_1e_1+x_2e_2, x_1e_1+x_2e_2) + \\
&\quad + 2\sum_{h=3}^{k} g_{\langle x_1e_1+x_2e_2,e_h,e_i\rangle}(x_1e_1+x_2e_2, e_h)x_h + \\
&\qquad\qquad \text{(where } i \geqq 3,\ i \neq h \text{ for each } h) \\
&\quad + \sum_{h,j=3}^{k} g_{\langle x_1e_1+x_2e_2,e_j,e_h,e_i\rangle}(e_j,e_h)x_jx_h = \\
&\qquad\qquad \text{(where if } j \neq h, \text{ then } i = j \text{ and if } j = h \text{ then } i \geqq 3,\ i \neq j) \\
&= \sum_{i,j=1}^{2} g_{\langle e_1,e_2,e_3,e_4\rangle}(e_i,e_j)x_ix_j + 2\sum_{i=1}^{2}\sum_{h=3}^{k} g_{\langle e_1,e_2,e_h\rangle}(e_i,e_h)x_ix_h + \\
&\quad + \sum_{h,j=3}^{k} g_{\langle e_1,e_2,e_j,e_h\rangle}(e_j,e_h)x_jx_h. \qquad \square
\end{aligned}$$

LEMMA 3. *Let (V,g) be a real inner product space of dimension $\geqq 3$. Suppose $L = \{e_i : i \in I\}$ is a linearly independent subset of V and any 3-dimensional subspace E of V containing two elements of L is a Lorentz vector space. Suppose $g(y,y) > 0$ for some $y \in \langle e_p, e_q\rangle$, some $p, q \in I$. Suppose either L is a basis of V or else V is a topological vector space, the span $\langle L\rangle$ is dense in V and g is continuous. Then V is a Lorentz vector space.*

PROOF. Note g is nondegenerate, and positive definite on a 1-dimensional subspace of V. Suppose, to obtain a contradiction, g is positive definite on a two-dimensional subspace. Then g is positive definite on a two-dimensional subspace E of $\langle L\rangle$. Let W be $\langle E, e_p, e_q\rangle$, and let $\{w_1, \dots, w_n\}$ be an orthonormal basis of W, with $g\left(\sum_{i=1}^{n} x_i w_i, \sum_{i=1}^{n} x_i w_i\right) = \sum_{i=1}^{k} x_i^2 - \sum_{i>k} x_i^2$, $k \geqq 2$. Take $y = \sum_{i=1}^{n} y_i w_i$ with $g(y,y) > 0$. Note $\sum_{i=1}^{k} y_i w_i \neq 0$, and there exists $h \neq 0$, $h \in \langle w_1, w_k\rangle$, with $g(h,y) = 0$. Then g is positive definite on $\langle y, h\rangle$, and hence on a three-dimensional subspace E containing e_p, e_q and h we do not have rank of positivity 1. □

THEOREM 1. *Let (X,p) be a super-additive normed linear space of dimension $\geqq 3$. Suppose that for every two-dimensional subspace N containing a timelike vector there is a projection $P : X \to X$ onto N of super-additive norm 1. Then the s.a. norm p is given by a unique Lorentz inner product on X.*

PROOF. Let C be the timecone in X. Let E be a 3-dimensional space spanned by timelike vectors. By Corollary 1, there is a unique Lorentz inner product g_E on E with $C \cap E$ a timecone given by g_E and $g_E(x,x) = p^2(x)$ for $x \in C \cap E$. By Lemma 2, if F is a finite dimensional space spanned by timelike vectors there is a unique inner product g_F on F with $g_F(x,x) = g_E(x,x)$ for E as above.

Since $\langle C\rangle = X$, there is a unique inner product g on X with $g(x,x) = p^2(x)$ for $x \in C$. By Lemma 3, X is a Lorentz vector space under g. Let $v \in C$ and let $\bar{C}$ be the timecone given by g containing v. For $x \in C$ by (ii) of the definition of a s.a. norm, $g(v,x) \geqq p(x)p(v)$, and hence $x \in \bar{C}$. If $z \in \bar{C}$, then for E a three-dimensional space containing z and v, z is in the timecone given by g_E containing v, hence in C. Thus $C = \bar{C}$, and $g(x,x) = p^2(x)$ for $x \in C$ so that p is given by g. □

THEOREM 2. *Let (X,p,C) be a super-additive normed linear space of dimension $\geqq 3$. Suppose X is also a topological vector space such that $p : C \to (0,\infty)$ is continuous with C open in X. Suppose $L = \{e_i : i \in I\}$ is a set of timelike vectors such that finite sums $\sum \lambda_i e_i$ in C are dense in C. Suppose that for any 2-dimensional subspace N containing an element e_i there is a projection P onto N of s.a. norm 1.*

Then p is given by a unique Lorentz inner product g on X.

PROOF. Let E be spanned by three linearly independent elements of L. There is, by Corollary 1, a unique Lorentz inner product g_E with $C \cap E$ a timecone given by g_E and $g_E(x,x) = p^2(x)$ for $x \in C \cap E$. If F is a finite dimensional space spanned by elements of L there is a unique inner product g_F on F with $g_F(x,c) = p^2(x)$ for $x \in C \cap F$, by Lemma 2. Thus there is a unique inner product $g_{\langle L\rangle}$ on $\langle L\rangle$ with $g_{\langle L\rangle}(x,x) = p^2(x)$ for $x \in C \cap \langle L\rangle$. If we let, for x and y in C,

$$g(x,y) = \frac{1}{2}\left(p^2(x+y) - p^2(x) - p^2(y)\right)$$

then g is continuous and symmetric on $C \times C$. Since $C \cap \langle L \rangle$ is dense in C, $g(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1 g(x_1, y) + \alpha_2 g(x_2, y)$ for $x_i \in C$, $y \in C$, $\alpha_i > 0$. Then g has a unique bilinear extension to $X = C - C$. Note $g(x,x) = p^2(x)$ for x in C and g is the unique inner product on X for which this holds. Also g is symmetric and g is continuous since C is open. By Lemma 3, (X, g) is a Lorentz vector space. Suppose $v \in C$ and $\bar{C} = \{x \in X : g(v,x) > 0, g(x,x) > 0\}$. As in Theorem 1, $C \subseteqq \bar{C}$, and $\bar{C} \subseteqq C$, so that p is given by g. □

REFERENCES

[1] AMIR, D., *Characterizations of inner product spaces,* Operator Theory: Advances and Applications, **20**, Birkhäuser–Verlag, Basel, Boston, Mass., Stuttgart, 1986. *MR* **88m**:46001

[2] BĂLAN, T., Observatii asupra functionalelor supra-aditive, *An. Univ. Craiova Ser. a IV-a* **1** (1970), 45–51. *MR* **48** # 12286

[3] BĂLAN, T. T., The polar of a super-additive norm, *Rev. Roumaine Math. Pures Appl.* **33** (1988), 651–654. *MR* **89j**:46021

[4] BOGNÁR, J., *Indefinite inner product spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 78, Springer–Verlag, New York–Heidelberg, 1974. *MR* **57** # 7125

[5] FITZPATRICK, S. and CALVERT, B., Convex bodies in $\mathbb{R}^3$ invariant under projections, *Math. Chronicle* **20** (1991), 89–108. *MR* **92k**:52007

[6] O'NEILL, B., *Semi-Riemannian geometry with applications to relativity,* Pure and Applied Mathematics, Vol. 103, Academic Press, New York, 1983. *MR* **85f**: 53002

(Received October 25, 1990)

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF AUCKLAND
AUCKLAND
NEW ZEALAND

Current address for S. Fitzpatrick:

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WESTERN AUSTRALIA
NEDLANDS, WA 6009
AUSTRALIA

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 67–70

SOME REMARKS CONNECTED WITH G. CSÓKA'S PAPER "ON AN EXTREMAL PROPERTY OF MINKOWSKI-REDUCED FORMS"

Á. G. HORVÁTH

Denote $f(x) = \sum\limits_{i,j=1}^{n} a_{ij}x_ix_j$ an n-ary positive definite quadratic form. The symmetric matrix $A = [a_{ij}]$ is the matrix of this form. In the famous work of Minkowski [1] the following interesting statement can be found without proof: "*The values*

$$(1) \qquad s_1 = \sum_{i=1}^{n} a_{ii}, \quad s_2 = \sum_{i \neq k} a_{ii}a_{kk}, \quad \dots, \quad s_n = \prod_{i=1}^{n} a_{ii},$$

are minimal for the Minkowski-reduced forms of the equivalent positive definite forms, if $n \leqq 5$."

In the paper [2] G. Csóka proved this theorem for the n-dimensional cases when $n \leqq 6$ and he verified that in the cases $n > 6$ the Minkowski-reduced forms do not have the above mentioned property. Our question is the following: is there such an element in every equivalence class of the n-ary positive definite forms for which the values $s_1, s_2, \dots, s_n$ are at the same time minimal? The answer is negative if $n > 6$, I will give a counter-example in Paragraph 1. This means that minimalizing the values $s_1, s_2, \dots, s_n$ on an equivalence class, after each other, we get in general different forms of this class so we have different types of reductions. In Paragraph 2 we take some further remarks which are connected with this problem.

§ **1.** In the paper [3] C. C. Ryškov gave a set of forms from which one can choose some interesting counter-examples (e.g. such a form which is Hermite-reduced but not Venkov-reduced, Venkov-reduced but not Hermitean one; the suitable definitions can be found in [4] p. 149 and p. 160 and [5]). Consider now the following positive definite form:

$$(2) \qquad f(\mathbf{x}) = \alpha(x_1^2 + \dots + x_5^2) + (1-\alpha)(x_1 + \dots + x_5)^2 + x_6^2 + \beta x_7^2,$$

where $\frac{13}{15} < \alpha < \frac{11}{12}$, and $\beta \geqq 1$. Denote by $\{\mathbf{a}_1, \dots, \mathbf{a}_7\}$ that vector system which corresponds to f, and take the following vectors:

$$\mathbf{e}_1 = \mathbf{a}_1, \dots, \mathbf{e}_6 = \mathbf{a}_6,$$

$$(3) \qquad \mathbf{e}_7 = \frac{1}{3}(\mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3 + \mathbf{a}_4) + \frac{1}{4}\mathbf{a}_5 + \left(\frac{1}{2} - \gamma\right)\mathbf{a}_6 + \frac{\sqrt{11}}{12}\mathbf{a}_7.$$

1991 *Mathematics Subject Classification.* Primary 11H55; Secondary 52C07.
Key words and phrases. Reduced form, euclidean point lattice

These vectors are linearly independent. Let L'_γ be the lattice that is spanned by $\{\mathbf{e}_1, \dots, \mathbf{e}_6\}$. It is obvious that the lattice L_γ that is spanned by the vectors $\{\mathbf{e}_1, \dots, \mathbf{e}_7\}$ is the following: $U\{L'_\gamma + k\mathbf{e}_7 | k \in \mathbf{Z}\}$. From (2) it can be seen that the shortest vectors of L'_γ are the unit vectors $\pm\mathbf{e}_1, \dots, \pm\mathbf{e}_6$, and for an other vector $\mathbf{v} \in L'_\gamma$, $f(\mathbf{v}) > 2\alpha > \frac{26}{15}$. We shall examine the cases of $k \geqq 4$ and $k = 1, 2, 3$, respectively. If $\beta \geqq 1$ and $|k| \geqq 4$ then for a vector $\mathbf{v} \in$ $\in L'_\gamma + k\mathbf{e}_7$ we have $f(\mathbf{v}) \geqq \frac{11}{9}$. Assume that $\mathbf{v} \in L'_\gamma + \mathbf{e}_7$ and $\mathbf{v} \neq \mathbf{e}_7$. From (2) we see that:

$$f(\mathbf{v}) > f(\mathbf{e}_7) = \frac{109}{144} + 2(1-\alpha) + \frac{11}{144}\beta - \gamma + \gamma^2 > 1 \tag{4}$$

if $0 < \gamma$ is sufficiently small. Similarly we get for $\mathbf{v} \in L'_\gamma \pm 2\mathbf{e}_7$ that

$$\mathbf{v} = \sum_{i=1}^{4} \left(-\frac{1}{3} + m_i\right) \mathbf{a}_i + \left(\frac{1}{2} + m_5\right) \mathbf{a}_5 + (m_6 - 2\gamma)\mathbf{a}_6 + \frac{\sqrt{11}}{6}\mathbf{a}_7, \tag{5}$$

where the numbers m_i are integer.

In the case of $\mathbf{v} \neq 2\mathbf{e}_7 - \sum_{i=1}^{4} \mathbf{e}_i - \mathbf{e}_6 \stackrel{\text{def}}{=} \mathbf{e}_6^\star$

$$f(\mathbf{v}) \geqq f(\mathbf{e}_6^\star) = \frac{100}{144} + \frac{44}{144}\beta + 4\gamma^2 > 1. \tag{6}$$

In the last case we suppose that

$$\mathbf{v} \neq \mathbf{e}_7^\star := 3\mathbf{e}_7 - \sum_{i=1}^{6} \mathbf{e}_i = -\frac{1}{4}\mathbf{a}_5 + \left(\frac{1}{2} - 3\gamma\right)\mathbf{a}_6 + \frac{\sqrt{11}}{4}\mathbf{a}_7$$

and so we ha e if $\beta > 10\gamma + 1$ that

$$f(\mathbf{v}) > f(\mathbf{e}_7^\star) = \frac{45}{144} + \frac{99}{144}\beta - 3\gamma + 9\gamma^2 > 1. \tag{7}$$

It is clear that the vector systems $\{\mathbf{e}_1, \dots, \mathbf{e}_7\}$ and $\{\mathbf{e}_1, \dots, \mathbf{e}_6^\star, \mathbf{e}_7^\star\}$ are the bases of the same lattice L_γ so the corresponding forms are equivalent. If we take now the following parameters:

$$\alpha = \frac{263}{288}, \quad \beta = \frac{133}{132}, \quad \gamma = 10^{-10}, \tag{8}$$

then

$$\begin{aligned} &\mathbf{e}_1^2 = \dots = \mathbf{e}_6^2 = 1 \quad \mathbf{e}_7^2 = 1 + \frac{1}{144} + \frac{11}{132 \cdot 144} - 10^{-10} + 10^{-20}, \\ &(\mathbf{e}_6^\star)^2 = 1 + \frac{44}{132 \cdot 144} + 4 \cdot 10^{-20}, \\ &(\mathbf{e}_7^\star)^2 = 1 + \frac{99}{132 \cdot 144} - 3 \cdot 10^{-10} + 9 \cdot 10^{-20}, \end{aligned} \tag{9}$$

and so

$$\mathbf{e}_6^{\star 2} + \mathbf{e}_7^{\star 2} < \mathbf{e}_6^2 + \mathbf{e}_7^2, \tag{10}$$

thus the basis of L_γ for which the value s_1 is minimal is the basis $\{\mathbf{e}_1, \dots, \mathbf{e}_6^\star, \mathbf{e}_7^\star\}$, but

$$\begin{aligned} \mathbf{e}_6^{\star 2} \cdot \mathbf{e}_7^{\star 2} &> 1 + \frac{143}{132 \cdot 144} + \frac{44 \cdot 99}{(132 \cdot 144)^2} - 3 \cdot 10^{-10} \left[1 + \frac{44}{132 \cdot 144}\right] > \\ &> 1 + \frac{143}{132 \cdot 144} - 10^{-10} + 10^{-20} = \mathbf{e}_6^2 \cdot \mathbf{e}_7^2, \end{aligned} \tag{11}$$

for this reason the values $s_2, \dots, s_7$ are not minimal for the basis $\{\mathbf{e}_1, \dots, \mathbf{e}_6^\star, \mathbf{e}_7^\star\}$. So we have proved the following

THEOREM. *If $n \geqq 7$ then there exists such an equivalence class of the n-ary positive forms for which the functions $s_1, \dots, s_n$ take their minima on different elements.*

§ **2.** First we note that in the above mentioned example the basis $\{\mathbf{e}_1, \dots, \mathbf{e}_6, \mathbf{e}_7\}$ of L_γ is a Hermite-reduced one and the basis $\{\mathbf{e}_1, \dots, \mathbf{e}_6^\star, \mathbf{e}_7^\star\}$ of L_γ is a Venkov-reduced form with respect to the form $\varphi = x_1^2 + \dots + x_n^2$. Secondly, it can be seen that in L_γ there is no basis $\{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ for which $|\mathbf{f}_1| \leqq |\mathbf{f}_2| \leqq \dots \leqq |\mathbf{f}_n|$ and if $\{\mathbf{g}_1, \dots, \mathbf{g}_n\}$ is a basis of L_γ for which $|\mathbf{g}_1| \leqq \dots \leqq |\mathbf{g}_n|$ then $|\mathbf{f}_i| \leqq |\mathbf{g}_i|$, $i = 1, \dots, n$. (Such a basis $\{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ is a Hermitean one, but in this lattice the length of the last vector of a Hermite basis is greater than $|\mathbf{e}_7^\star|$.) At third we remark that for the linearly independent vector system of a lattice L the following is true:

STATEMENT. *Let L be an n-lattice and $\{\mathbf{a}_1, \dots, \mathbf{a}_n\} \subset L$ a linearly independent vector system for which $|\mathbf{a}_1| \leqq \dots \leqq |\mathbf{a}_n|$. Then the following statements are equivalent:*

(i) *$\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ is a successive minimum system of L;*

(ii) *if the vectors of the system $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ are independent and $|\mathbf{b}_1| \leqq \dots \leqq |\mathbf{b}_n|$ then $|\mathbf{a}_i| \leqq |\mathbf{b}_i|$;*

(iii) *the system $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ is the common minimum of the functions $s_1, \dots, s_n$ on the set of the independent systems containing n elements of L;*

(iv) *the system $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ is the minimum of the function s_1.*

This statement follows, for example, from the Rado-Edmonds theorem for matroids (see [6]).

REFERENCES

[1] MINKOWSKI, H., Diskontinuitätsbereich für arithmetische Äquivalenz, *J. Reine Angew. Math.* **129** (1905), 220–274. *Jahrbuch Fortschritte Math.* **37**, 251

[2] Csóka, G., On an extremal property of Minkowski-reduced frames, *Studia Sci. Math. Hungar.* **13** (1978), 469–475 (in Russian). *MR* **83e**:10042

[3] Ryškov, C. C., The reduction of positive quadratic forms of n variables in the sense of Hermite, Minkowski and Venkov, *Dokl. Akad. Nauk SSSR* **207** (1972), 1054–1056 (in Russian). *MR* **47** #3316

[4] Gruber, P. M. and Lekkerkerker, C. G., *Geometry of numbers*, Second edition, North-Holland Mathematical Library, **37**, North-Holland Publishing Co., Amsterdam–New York, 1987. *MR* **88j**:11034

[5] Venkov, B. A., Über die Reduktion positiver quadratischer Formen, *Izvestija Akad. Nauk SSSR Ser. Mat.* **4** (1940), 37–52 (in Russian with German summary). *Jahrbuch Fortschritte Math.* **66**, Part 2, 47

[6] Lawler, E. L., *Combinatorial optimization: networks and matroids*, Holt, Rinehart and Winston, New York–Montreal–London, 1976. *MR* **55** #12005

(Received October 15, 1990)

BUDAPESTI MŰSZAKI EGYETEM
GÉPÉSZMÉRNÖKI KAR
GEOMETRIA TANSZÉK
EGRI JÓZSEF U. 1.
H–1521 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 71–83

KRONROD EXTENSION OF TURÁN FORMULA

S. LI

Abstract

In this paper we propose a Kronrod type extension to the well-known Turán formula. It is shown that such an extension exists for any positive measure. For the special Chebyshev measure $d\sigma(t) = (1-t^2)^{-1/2}dt$, some explicit formulas for the weights in the new quadrature formula are obtained.

1. Introduction

In 1950 P. Turán [10] proposed a quadrature formula of the type

$$\int_R f(t)\,d\sigma(t) = \sum_{\nu=1}^{n}\sum_{i=0}^{r-1} a_{i,\nu} f^{(i)}(\tau_\nu) + R_{n,r}(f). \tag{1.1}$$

At that time he treated only the case when $d\sigma(t) = dt$ on $[-1,1]$. In the formula (1.1), τ_ν are the zeros of a polynomial π_n of degree n which satisfies the orthogonality relation

$$\int_R \pi_n^r(t) p_k(t)\,d\sigma(t) = 0, \qquad k = 0,1,\dots,n-1, \tag{1.2}$$

and $a_{i,\nu}$ are determined through interpolation. If τ_ν and $a_{i,\nu}$ are chosen in this way, then the degree of exactness for (1.1) is $(r+1)n-1$. Turán proved that (1.1) always exists and is unique if r is odd. In the case of the first-kind Chebyshev measure, Micchelli and Rivlin [4] have proved the following important result:

If $f \in P_{2(s+1)n-1}$, then

$$\int_{-1}^{1} \frac{f(t)}{\sqrt{1-t^2}}\,dt = \frac{\pi}{n}\left\{\sum_{\nu=1}^{n} f(\tau_\nu) + \sum_{j=1}^{s} \alpha_j f'\left[\tau_1^{2j},\dots,\tau_n^{2j}\right]\right\},$$

1980 *Mathematics Subject Classification* (1985 Revision). Primary 65D30, 65D32; Secondary 33A65.

Key words and phrases. Numerical quadratures, Turán formula, Kronrod extension, Chebyshev weight function.

0081-6906/94/$ 4.00

where

$$\alpha_j = (-1)^j \frac{\binom{-\frac{1}{2}}{j}}{2j4^{(n-1)j}}, \qquad j = 1, 2, \ldots, n.$$

By using this, they derived explicit formulas for the weights in (1.1). Later, Riess [7] and Varma [11], using different methods, found the explicit solution of the Turán problem XXVI for $s = 2$. Recently, Milovanović [5] studied a numerical approach for computing π_n for general measure.

Following Kronrod [2], we propose to extend the formula (1.1) to the following formula

$$(1.3) \qquad \int_R f(t)\,d\sigma(t) = \sum_{\nu=1}^{n} \sum_{i=0}^{r-1} \sigma_{i,\nu} f^{(i)}(\tau_\nu) + \sum_{\mu=1}^{n+1} K_\mu f(\hat{\tau}_\mu) + R_{n,r}(f)$$

where τ_ν are the same nodes as in (1.1), and the new nodes $\hat{\tau}_\nu$ and new weights $\sigma_{i,\nu}$, K_μ are chosen to maximize the degree of exactness for (1.3). It is shown in Section 2 that we can always obtain the maximum degree $(r+2)n+1$ (r is odd) by taking $\hat{\tau}_\mu$ to be the zeros of the polynomial $\hat{\pi}_{n+1}$ satisfying the orthogonality property

$$(1.4) \qquad \int_R \hat{\pi}_{n+1}(t) p(t) \pi_n^r(t)\,d\sigma(t) = 0, \qquad \text{all } p \in P_n.$$

At the same time we show that $\hat{\pi}_{n+1}$ always exists and is unique if it is monic. We devote Section 3 to the special case when $d\sigma(t) = (1-t^2)^{-\frac{1}{2}}dt$. In this case one can determine $\hat{\pi}_{n+1}$ explicitly and go even further to obtain the weights in (1.3) for $r = 3$ and $r = 5$.

2. Existence of $\hat{\pi}_{n+1}$

The orthogonality relation (1.4) is imposed to maximize the degree of exactness for (1.3). A generalization of a theorem in Gautschi [1, p.78] shows that the degree of exactness is $(r+2)n+1$, provided the polynomial $\hat{\pi}_{n+1}$ exists.

THEOREM 2.1. *If $r = 2s+1$ and $d\sigma$ is any positive measure, then $\hat{\pi}_{n+1}$ exists and is unique up to a constant factor.*

PROOF. First we observe that $\{\pi_k\}_{k=0}^{n+1}$ forms a basis in P_{n+1}. To see this, let us assume there are b_k such that

$$\sum_{k=0}^{n+1} b_k \pi_k(t) = 0,$$

then it is easy to show all $b_k = 0$ by multiplying $\pi_k^r(t)$, $k = n+1, n, \dots, 1, 0$, in turn, in the above equation and integrating it with respect to $d\sigma$. Hence it is natural to write

$$\hat{\pi}_{n+1}(t) = \pi_{n+1}(t) + \sum_{j=0}^{n} c_j \pi_j(t).$$

The orthogonality relation (1.4) then becomes

$$\int_R \Big(\pi_{n+1}(t) + \sum_{j=0}^{n} c_j \pi_j(t)\Big) \pi_{n-k}(t) \pi_n^r(t)\, d\sigma(t) = 0, \qquad k = 0, 1, 2, \dots, n.$$

This yields a linear system whose coefficient matrix is upper triangular and the diagonal elements are all equal to $\int_R \pi_n^{r+1}(t)\, d\sigma(t)$, which is positive. Therefore we can conclude that $\hat{\pi}_{n+1}$ is uniquely determined. □

REMARK. In the numerical construction of $\hat{\pi}_{n+1}$, we can also proceed as above. To do this, it is important to compute $\int_R \pi_j(t) \pi_{n-k}(t) \pi_n^r(t)\, d\sigma(t)$ effectively. The latter can be computed by Gauss quadrature.

The weights $\sigma_{i,\nu}$ and K_μ admit various representations, the ones obtained via interpolation [8] having the forms:

(2.1)

$$\sigma_{i,\nu} = \frac{1}{i!} \int_R \frac{\hat{\pi}_{n+1}(t) \pi_n^r(t)}{(t-\tau_\nu)^{r-i}} \sum_{j=0}^{r-i-1} \frac{(t-\tau_\nu)^j}{j!} D^{(j)} \left(\frac{(t-\tau_\nu)^r}{\pi_n^r(t) \hat{\pi}_{n+1}(t)} \right)\Big|_{t=\tau_\nu} d\sigma(t)$$

$$i = 0, 1, 2, \dots, r-1, \qquad \nu = 1, 2, \dots, n,$$

$$K_\mu = \int_R \frac{\pi_n^r(t) \hat{\pi}_{n+1}(t)}{\pi_n^r(\hat{\tau}_\mu) \hat{\pi}'_{n+1}(\hat{\tau}_\mu)(t - \hat{\tau}_\mu)}\, d\sigma(t), \qquad \mu = 1, 2, \dots, n+1,$$

where D is the differential operator. For general i, the first formula above is very complicated. However, in the special case $i = r - 1 = 2s$, one finds the simpler formula

(2.2)
$$\begin{aligned} \sigma_{2s,\nu} &= \frac{1}{(2s)!} \frac{1}{\hat{\pi}_{n+1}(\tau_\nu)(\pi'_n(\tau_\nu))^{2s+1}} \int_R \frac{\hat{\pi}_{n+1}(t) \pi_n^{2s+1}(t)}{t - \tau_\nu}\, d\sigma(t) = \\ &= \frac{1}{(2s)!} \frac{B_{s,\nu}}{(\pi'_n(\tau_\nu))^{2s}}, \end{aligned}$$

where $B_{s,\nu}$ are the weights of the interpolatory quadrature formula

$$\int_R g(t)\, d\mu(t) = \sum_{\nu=1}^{n} B_{s,\nu} g(\tau_\nu) + \sum_{\mu=1}^{n+1} C_{s,\mu} g(\hat{\tau}_\mu) + R_n(f), \tag{2.3}$$

where $d\mu(t) = \pi_n^{2s}(t)d\sigma(t)$.

For the numerical evaluation of $\sigma_{i,\nu}$ and K_μ it is more convenient to solve a linear system resulting from (1.3) by setting f equal to appropriate special polynomials.

3. Chebyshev weight

We now study the quadrature formula (1.3) in detail for the first-kind Chebyshev weight function. In this case it is known, independently of s, that

$$\pi_n(t) = \frac{1}{2^{n-1}} T_n(t). \tag{3.0}$$

This allows us to obtain an explicit formula for $\hat{\pi}_{n+1}$.

THEOREM 3.1. *Let $\hat{\pi}_{n+1}$ be the monic polynomial of degree $n+1$ satisfying the orthogonality relation*

$$\int_{-1}^{1} \hat{\pi}_{n+1}(t)p(t)\pi_n^r(t)(1-t^2)^{-\frac{1}{2}}\, dt = 0 \qquad \text{for all } p \in P_n. \tag{3.1}$$

If $r = 2s+1$, then

$$\hat{\pi}_{n+1}(t) = \begin{cases} \dfrac{1}{2^n}\left(T_{n+1}(t) - T_{n-1}(t)\right) & \text{if } n \geqq 2, \\ \dfrac{1}{2}\left(T_2(t) - \dfrac{s+1}{s+2}T_0(t)\right) & \text{if } n = 1. \end{cases} \tag{3.2}$$

The following lemma is needed in our proof of Theorem 3.1.

LEMMA 3.2. *For any positive integer s, we have*

$$\begin{aligned} (1+T_{2n})^s &= \frac{1}{2^{s-1}}\left[\sum_{k=0}^{s-1}\binom{2s}{k}T_{2(s-k)n} + \frac{1}{2}\binom{2s}{s}\right] = \\ &= \frac{1}{2^{s-1}}\left[\sum_{k=1}^{s}\binom{2s}{s-k}T_{2kn} + \frac{1}{2}\binom{2s}{s}\right]. \end{aligned} \tag{3.3}$$

PROOF. By induction on s. □

PROOF OF THEOREM 3.1. We first consider $n \geqq 2$. Since $\hat{\pi}_{n+1}$ can be written as

$$\hat{\pi}_{n+1}(t) = \frac{1}{2^n}\left(T_{n+1}(t) + \sum_{j=0}^{n} c_j T_j(t)\right),$$

by letting $p(t) = T_k(t)$, $k = 0, 1, 2, \dots, n$, and noting (3.0) and $T_n^2(t) = \frac{1}{2}(1 + T_{2n}(t))$, the conditions (3.1) become

$$(3.4)\quad \int_{-1}^{1} \Big(T_{n+1}(t) + \sum_{j=0}^{n} c_j T_j(t)\Big) T_k(t)(1 + T_{2n}(t))^s T_n(t)(1 - t^2)^{-\frac{1}{2}}\, dt = 0,$$

for $k = 0, 1, 2, \dots, n$.

Applying Lemma 3.2 and observing that

$$\left(\sum_{i=1}^{s} \binom{2s}{s-i} T_{2in} + \frac{1}{2}\binom{2s}{s}\right) T_n =$$
$$= \frac{1}{2}\left[\sum_{i=2}^{s} \binom{2s}{s-i} \left(T_{(2i+1)n} + T_{(2i-1)n}\right) + \binom{2s}{s-1} T_{3n} + \left(\binom{2s}{s-1} + \binom{2s}{s}\right) T_n\right],$$

we obtain from (3.4), by the orthogonality of the T_m,

$$(3.5)\quad \int_{-1}^{1} \Big(T_{n+1}(t) + \sum_{j=0}^{n} c_j T_j(t)\Big) T_k(t) T_n(t)(1 - t^2)^{-\frac{1}{2}}\, dt = 0, \quad k = 0, 1, \dots, n.$$

We thus see that the first factor in (3.5) is equal to the Stieltjes polynomial corresponding to the first kind Chebyshev weight function up to a constant multiple. Therefore (see [6]),

$$\hat{\pi}_{n+1}(t) = \frac{1}{2^n}\left(T_{n+1}(t) - T_{n-1}(t)\right).$$

The proof of the case $n = 1$ is analogous. □

REMARK. It is interesting to note that the new nodes $\hat{\tau}_\mu$ in (1.3) are nothing but the new nodes in the Kronrod extension of Gauss quadrature formula corresponding to the first-kind Chebyshev measure. This probably is the only measure with this property.

In the proof of Theorem 3.1 one can see clearly that if $n + k + 1 < 3n$, then (3.1) is true for any polynomial of degree k. This indicates that $\hat{\pi}_{n+1}$ is orthogonal to all polynomials of degree lower than $2n - 1$ with respect to the measure $d\hat{\sigma}(t) = \pi_n^r(t)(1 - t^2)^{-\frac{1}{2}}\, dt$. We conclude from this that the degree of exactness of the quadrature rule (1.3) for the first-kind Chebyshev weight is as high as $(r + 3)n - 1$ if $n \geqq 2$.

For the time being, we assume $n \geqq 2$. In the following theorem, we develop an explicit formula for the weights K_μ associated with the new nodes for any $r = 2s + 1$.

THEOREM 3.2. *If the nodes $\hat{\tau}_\mu$ are arranged in decreasing order, then the corresponding weights K_μ admit the form*

$$K_\mu = \begin{cases} \dfrac{1}{2^{2s+2}}\dbinom{2s+2}{s+1}\dfrac{\pi}{n}, & \mu = 2,3,\dots,n, \\ \dfrac{1}{2^{2s+2}}\dbinom{2s+2}{s+1}\dfrac{\pi}{2n}, & \mu = 1, n+1. \end{cases} \tag{3.6}$$

PROOF. First we notice that K_μ does not depend on the leading coefficient of π_n and $\hat{\pi}_{n+1}$. Because of this one can rewrite K_μ as

$$K_\mu = \int_{-1}^{1} \frac{T_n^r(t)\,(T_{n+1}(t) - T_{n-1}(t))}{T_n^r(\hat{\tau}_\mu)\,(T'_{n+1}(\hat{\tau}_\mu) - T'_{n-1}(\hat{\tau}_\mu))\,(t-\hat{\tau}_\mu)}(1-t^2)^{-\frac{1}{2}}\,dt. \tag{3.7}$$

To evaluate the above integral, we expand

$$\frac{T_{n+1}(t) - T_{n-1}(t)}{t-\hat{\tau}_\mu} = 2T_n(t) + \sum_{k=0}^{n-1} d_k T_k(t),$$

for some constants d_k. This, together with the s-orthogonality of T_n, leads to

$$\int_{-1}^{1} \frac{T_n^{2s+1}(t)\,(T_{n+1}(t) - T_{n-1}(t))}{t-\hat{\tau}_\mu}(1-t^2)^{-\frac{1}{2}}\,dt = 2\int_{-1}^{1} T_n^{2s+2}(t)(1-t^2)^{-\frac{1}{2}}\,dt.$$

The last integral can be evaluated by using

$$T_n^{2s+2}(t) = \frac{1}{2^{2s+1}}\left(\sum_{k=1}^{s+1}\binom{2s+2}{s+1-k}T_{2kn} + \frac{1}{2}\binom{2s+2}{s+1}\right),$$

which follows from Lemma 3.2. Indeed, we obtain

$$\int_{-1}^{1} T_n^{2s+2}(t)(1-t^2)^{-\frac{1}{2}}\,dt = \frac{\pi}{2^{2s+2}}\binom{2s+2}{s+1},$$

which yields

$$\int_{-1}^{1} T_n^{2s+1}(t)\frac{T_{n+1}(t) - T_{n-1}(t)}{t-\hat{\tau}_\mu}(1-t^2)^{-\frac{1}{2}}\,dt = \frac{\pi}{2^{2s+1}}\binom{2s+2}{s+1}. \tag{3.8}$$

Next we proceed to compute $T_n^r(\hat{\tau}_\mu)$ and $T'_{n+1}(\hat{\tau}_\mu) - T'_{n-1}(\hat{\tau}_\mu)$. Since the nodes $\hat{\tau}_\mu$ are given by $\hat{\tau}_\mu = \cos\frac{(\mu-1)\pi}{n}$, $\mu = 1, 2, \ldots, n+1$, by an elementary computation we find

$$T_n^{2s+1}(\hat{\tau}_\mu) = (-1)^{\mu-1},$$

and

$$T'_{n+1}(\hat{\tau}_\mu) - T'_{n-1}(\hat{\tau}_\mu) = \begin{cases} 2n(-1)^{\mu-1}, & \mu = 2, 3, \ldots, n, \\ 4n(-1)^{\mu-1}, & \mu = 1, n+1. \end{cases}$$

Finally, (3.6) follows from (3.7), (3.8) and the last two equalities. □

By a very similar computation one can show that the weights K_μ in the extension formula with $n = 1$ have the following representations:

$$K_1 = K_2 = \frac{\pi}{2^{2s+3}\left(1 - \dfrac{1}{2(s+2)}\right)^{s+1}} \binom{2s+2}{s+1}.$$

We now turn to the derivation of the weights $\sigma_{i,\nu}$. We succeed in obtaining explicit formulas for $\sigma_{i,\nu}$ when $s = 1$ and $s = 2$. A detailed discussion is presented for the case $s = 1$. For the other case we only state the results.

We will be using some notations and identities from Varma [11]. Let $\tau_k = \cos\theta_k = \cos\frac{(2k-1)\pi}{2n}$, $k = 1, 2, \ldots, n$, be the zeros of T_n. For these nodes we can represent the fundamental polynomials of Lagrange and Hermite interpolation in terms of the first-kind Chebyshev polynomials T_m. The following relations [11] are well known:

$$l_k(t) = \frac{T_n(t)}{(t - \tau_k)T'_n(\tau_k)}, \tag{3.9}$$

$$r_k(t) = \frac{1 - t\tau_k}{1 - \tau_k^2} l_k^2(t) = \frac{1}{n} + \frac{1}{n^2} \sum_{j=1}^{2n-1} (2n - j)T_j(t)T_j(\tau_k), \tag{3.10}$$

$$\rho_k(t) = (t - \tau_k)l_k^2(t) = \frac{\sin\theta_k}{n^2} \sum_{j=1}^{2n-1} \sin j\theta_k T_j(t), \tag{3.11}$$

where l_k are the fundamental Lagrange, r_k and ρ_k the fundamental Hermite interpolation polynomials. Varma [11] has shown

$$\sum_{\nu=1}^{n} (1 - \tau_\nu^2) r''_k(\tau_\nu) = 0, \qquad k = 1, 2, \ldots, n, \tag{3.12}$$

$$\sum_{\nu=1}^{n} (1 - \tau_\nu^2) \rho''_k(\tau_\nu) = \tau_k, \qquad k = 1, 2, \ldots, n. \tag{3.13}$$

With these results we are able to establish the following theorem.

THEOREM 3.3. *If $r=3$ $(s=1)$, then the weights $\sigma_{0,k}$, $\sigma_{1,k}$ and $\sigma_{2,k}$ in the quadrature formula* (1.3) *relative to the measure $d\sigma(t)=(1-t^2)^{-\frac{1}{2}}\,dt$ are given by*

$$\begin{aligned}\sigma_{0,k}&=\frac{5\pi}{8n},\\ \sigma_{1,k}&=-\frac{\pi}{16n^3}\tau_k,\\ \sigma_{2,k}&=\frac{\pi(1-\tau_k^2)}{16n^3},\qquad k=1,2,\ldots,n,\end{aligned}\tag{3.14}$$

where $\tau_k=\cos((2k-1)\pi/2n)$.

PROOF. Letting $d\sigma(t)=(1-t^2)^{-\frac{1}{2}}\,dt$ and $s=1$ in (2.2) yields

$$\begin{aligned}\sigma_{2,k}&=\frac{1}{2\left(T_{n+1}(\tau_k)-T_{n-1}(\tau_k)\right)\left(T_n'(\tau_k)\right)^3}\times\\ &\times\int_{-1}^{1}\frac{\left(T_{n+1}(t)-T_{n-1}(t)\right)T_n^3(t)}{t-\tau_k}(1-t^2)^{-\frac{1}{2}}\,dt.\end{aligned}\tag{3.15}$$

An elementary computation gives

$$\left(T_{n+1}(t)-T_{n-1}(t)\right)T_n^2(t)=\frac{1}{4}\left(T_{n+1}(t)-T_{n-1}(t)+T_{3n+1}(t)-T_{3n-1}(t)\right).$$

To the remaining factor T_n we apply the Christoffel–Darboux formula (see [9], Eq. 3.2.3) to T_n yields

$$\frac{T_n(t)}{t-\tau_k}=\frac{2}{T_{n-1}(\tau_k)}\sum_{j=0}^{n-1}T_j(\tau_k)T_j(t),$$

which then gives

$$\begin{aligned}&\int_{-1}^{1}\frac{\left(T_{n+1}(t)-T_{n-1}(t)\right)T_n^3(t)}{t-\tau_k}(1-t^2)^{-\frac{1}{2}}\,dt=\\ &=-\frac{1}{2}\int_{-1}^{1}T_{n-1}^2(t)(1-t^2)^{-\frac{1}{2}}\,dt=-\frac{\pi}{4}.\end{aligned}\tag{3.16}$$

The last formulae of (3.14) are obtained by applying (3.15) and (3.16), and observing

$$\begin{aligned}T_{n+1}(\tau_k)-T_{n-1}(\tau_k)&=2(-1)^k\sqrt{1-\tau_k^2},\\ T_n'(\tau_k)&=\frac{(-1)^{k-1}n}{\sqrt{1-\tau_k^2}}.\end{aligned}\tag{3.17}$$

To compute $\sigma_{1,k}$, we put $f(t) = \rho_k(t)$ in (1.3) with $r = 3$. Since, by (3.11), $\int_{-1}^{1} \rho_k(t)(1-t^2)^{-\frac{1}{2}}\,dt = 0$, we then obtain:

$$\sigma_{1,k} + \sum_{\nu=1}^{n} \sigma_{2,\nu}\rho_k''(\tau_\nu) + \sum_{\mu=1}^{n+1} K_\mu \rho_k(\hat{\tau}_\mu) = 0. \tag{3.18}$$

It is easy to see that $T_n(\hat{\tau}_\mu) = (-1)^{\mu-1}$. Using this, together with (3.7), (3.11) and the second equality in (3.17), the last sum in (3.18) is found to be

$$-\frac{3\pi}{8n^3}\left[(1-\tau_k^2)\sum_{\mu=1}^{n+1}\frac{1}{\tau_k - \hat{\tau}_\mu} - \frac{1-\tau_k^2}{2}\left(\frac{1}{\tau_k - 1} + \frac{1}{\tau_k + 1}\right)\right].$$

The sum in the last expression, in turn, is computed by using

$$\sum_{\mu=1}^{n+1} \frac{1}{\tau_k - \hat{\tau}_\mu} = \frac{T'_{n+1}(\tau_k) - T'_{n-1}(\tau_k)}{T_{n+1}(\tau_k) - T_{n-1}(\tau_k)} = -\frac{\tau_k}{1-\tau_k^2}. \tag{3.19}$$

This enables us to obtain

$$\sum_{\mu=1}^{n+1} K_\mu \rho_k(\hat{\tau}_\mu) = 0.$$

Therefore it follows from (3.18) that

$$\sigma_{1,k} = -\sum_{\nu=1}^{n} \sigma_{2,\nu}\rho_k''(\tau_\nu) = -\frac{\pi}{16n^3}\sum_{\nu=1}^{n}(1-\tau_\nu^2)\rho_k''(\tau_\nu) = -\frac{\pi}{16n^3}\tau_k.$$

Finally, we proceed to compute $\sigma_{0,k}$. To do this, we begin with setting $f(t) = r_k(t)$ in (1.3) with $r = 3$. This gives

$$\sigma_{0,k} + \sum_{\nu=1}^{n} \sigma_{2,\nu} r_k''(\tau_\nu) + \sum_{\mu=1}^{n+1} K_\mu r_k(\hat{\tau}_\mu) = \frac{\pi}{n}. \tag{3.20}$$

Because of (3.12) and (3.14) the first sum in (3.20) equals zero. A simple observation yields

$$r_k(\hat{\tau}_\mu) = \frac{1 - \hat{\tau}_\mu \tau_k}{n^2(\tau_k - \hat{\tau}_\mu)^2}.$$

Consequently, by using (3.6) with $s = 1$, one has

$$\sum_{\mu=1}^{n+1} K_\mu r_k(\hat{\tau}_\mu) = \frac{3\pi}{8n^3}\left[(1-\tau_k^2)\sum_{\mu=1}^{n+1}\frac{1}{(\tau_k - \hat{\tau}_\mu)^2} + \tau_k \sum_{\mu=1}^{n+1} \tau \right. \tag{3.21}$$

It is not hard to see that the first summation on the right-hand side equals

$$\left(\frac{\hat{\pi}'_{n+1}(\tau_k)}{\hat{\pi}_{n+1}(\tau_k)}\right)^2 - \frac{\hat{\pi}''_{n+1}(\tau_k)}{\hat{\pi}_{n+1}(\tau_k)}.$$

Since

$$\hat{\pi}_{n+1}(t) = \frac{1}{2^n}\left(T_{n+1}(t) - T_{n-1}(t)\right),$$

the function values and the first two derivatives of T_{n+1} and T_{n-1} at $t = \tau_k$ are needed. Using the differential equation satisfied by T_n, one finds

$$T''_{n+1}(\tau_k) = \frac{(-1)^{k-1}(n+1)\left(\tau_k^2 + (n+1)(1-\tau_k^2)\right)}{(1-\tau_k^2)^{\frac{3}{2}}},$$

$$T''_{n-1}(\tau_k) = \frac{(-1)^{k-1}(n-1)\left(\tau_k^2 - (n-1)(1-\tau_k^2)\right)}{(1-\tau_k^2)^{\frac{3}{2}}}.$$

Hence, it follows from the last two equalities and (3.19) that

$$\sum_{\mu=1}^{n+1} \frac{1}{(\tau_k - \hat{\tau}_\mu)^2} = \frac{n^2 + 1 - (n^2-1)\tau_k^2}{(1-\tau_k^2)^2}.$$

Substitution of the last relation and (3.19) in (3.21) gives

$$\sum_{\mu=1}^{n+1} K_\mu r_k(\hat{\tau}_\mu) = \frac{3\pi}{8n},$$

which combined with (3.19) yields

$$\sigma_{0,k} = \frac{5\pi}{8n}. \qquad \square$$

REMARK. Theorem 3.3 is valid only for $n \geq 2$. In the case $n = 1$, the following modifications must be made: $\sigma_{0,1} = \frac{23\pi}{50}$, $\sigma_{1,1} = 0$, $\sigma_{2,1} = \frac{\pi}{40}$.

To conclude this section, we state the result for $r = 5$ without giving a detailed proof.

THEOREM 3.4. *If $r = 5$ $(s = 2)$, then in the quadrature formula* (1.3) *relative to the measure $d\sigma(t) = (1-t^2)^{-\frac{1}{2}}\,dt$, the weights $\sigma_{0,k}$, $\sigma_{1,k}$, $\sigma_{2,k}$, $\sigma_{3,k}$*

and $\sigma_{4,k}$ *admit the representations below:*

$$
\begin{aligned}
\sigma_{0,k} &= \frac{11\pi}{16n}, \\
\sigma_{1,k} &= -\frac{\pi}{384n^5}(40n^2-1)\tau_k, \\
\sigma_{2,k} &= \frac{\pi}{384n^5}(3+(40n^2-7)(1-\tau_k^2)), \\
\sigma_{3,k} &= -\frac{\pi}{64n^5}\tau_k(1-\tau_k^2), \\
\sigma_{4,k} &= \frac{\pi}{384n^5}(1-\tau_k^2)^2,
\end{aligned} \tag{3.22}
$$

where $\tau_k = \cos((2k-1)\pi/2n)$.

The last relation in (3.21) is a direct consequence of a more general identity,

$$\sigma_{2s,k} = \frac{\pi}{(s!)^2(s+1)2^{2s+1}n^{2s+1}}(1-\tau_\nu^2)^s.$$

The penultimate formula of (3.21) is obtained by using the first formula of (2.1) directly. To derive the first three relations of (3.21), we set in turn $f(t) = \frac{T_n^2(t)r_k(t)}{2(T_n'(\tau_k))^2}$, $\rho_k(t)$, and $r_k(t)$ in (1.3). (They are obtained in the order $\sigma_{2,k}, \sigma_{1,k}, \sigma_{0,k}$.) For the computation, equations (3.4) and (3.5) in Varma [11] are needed. Again, in Theorem 3.4, the results only hold for $n \geqq 2$. Explicit formulae of the weights for $n = 1$ can be derived directly.

4. Numerical result

First we make some comments regarding the computational work and accuracy of the quadrature formulae (1.1) and (1.3). The discussion is restricted to the first-kind Chebyshev measure. Since the major computation comes from the evaluations of the function and derivative values, it is reasonable to compare the number of these evaluations. The formula (1.3) needs $n+1$ more function evaluations than (1.1) does, but the degree of exactness of the former is $2n$ times larger. In terms of efficiency, each evaluation of formula (1.1) achieves the degree of exactness $\frac{(r+1)n-1}{nr}$, whereas each evaluation of (1.3) achieves the degree of exactness $\frac{(r+3)n-1}{n(r+1)+1}$. This suggests that (1.3) may be more effective than (1.1).

EXAMPLE. *Evaluate*

$$\int_{-1}^{1} \frac{e^{Pt}}{\sqrt{1-t^2}}\,dt,$$

for any complex number P.

It is known that

$$\int_{-1}^{1} \frac{e^{Pt}}{\sqrt{1-t^2}}\, dt = \pi I_0(P),$$

where I_0 is the modified Bessel function, which can be computed by Algorithm I in Gautschi [3]. Since the function e^{Pt} is well-behaved when P is a real number or a complex number with small imaginary part, the convergence of (1.1) is already very fast. To see the effectiveness of (1.3) we deliberately let P be a complex number with large imaginary part. The computations were carried out on Vax computer, and the numerical results are shown below:

In Table 4.1, T and ET stand for Turán and Extended Turán formulae, respectively.

n	T	ET
2	$6.6 \cdot 10^{3}$	$1.6 \cdot 10^{3}$
3	$1.1 \cdot 10^{3}$	$2.8 \cdot 10^{2}$
4	$1.1 \cdot 10^{3}$	$2.5 \cdot 10^{2}$
5	$4.2 \cdot 10^{2}$	$5.1 \cdot 10^{1}$
6	$4.7 \cdot 10^{2}$	$9.3 \cdot 10^{1}$
7	$2.3 \cdot 10^{2}$	$3.6 \cdot 10^{1}$
8	$1.2 \cdot 10^{2}$	$1.5 \cdot 10^{0}$
9	$1.4 \cdot 10^{2}$	$1.6 \cdot 10^{-2}$
10	$1.1 \cdot 10^{2}$	$7.1 \cdot 10^{-5}$
11	$2.2 \cdot 10^{1}$	$1.4 \cdot 10^{-7}$
12	$2.2 \cdot 10^{0}$	$1.4 \cdot 10^{-10}$
13	$1.2 \cdot 10^{-1}$	$7.9 \cdot 10^{-14}$
14	$4.4 \cdot 10^{-3}$	machine precision
15	$1.1 \cdot 10^{-4}$	machine precision
16	$1.8 \cdot 10^{-6}$	machine precision
17	$2.2 \cdot 10^{-8}$	machine precision
18	$2.1 \cdot 10^{-10}$	machine precision
19	$1.4 \cdot 10^{-12}$	machine precision
20	$6.5 \cdot 10^{-14}$	machine precision

Table 4.1. Relative errors of the Turán and Extended Turán formula applied to the integral of the example with $P = 40i$

ACKNOWLEDGEMENT. The author thanks to Professor W. Gautschi for helpful suggestions and comments in the preparation of this paper.

REFERENCES

[1] GAUTSCHI, W., A survey of Gauss–Christoffel quadrature formulae, *E. B. Christoffel* (Aachen/Monschau, 1979), (P. L. Butzer and F. Fehér, eds.), Birkhäuser, Basel, 1981, pp. 72–147. *MR* **83g**:41031

[2] Gautschi, W., Gauss-Kronrod quadrature — a survey, *Numerical Methods and Approximation Theory*, III, (Niš, 1987), (G. V. Milovanović, ed.), Faculty of Electronic Engineering, Univ. Niš, Niš, 1988, 39-66. *MR* **89k**:41035

[3] Gautschi, W., Computational aspects of three-term recurrence relations, *SIAM Rev.* **9** (1967), 24–82. *MR* **35** #3927

[4] Micchelli, C. A. and Rivlin, T. J., Turán formulae and highest precision quadrature rules for Chebyshev coefficients. Mathematics of numerical computation, *IBM J. Res. Develop.* **16** (1972), 372–379. *MR* **48** # 12784

[5] Milovanović, G. V., Construction of s-orthogonal polynomials and Turán quadrature formulae, *Numerical Methods and Approximation Theory*, III, (Niš, 1987), Univ. Niš, Niš, 1988, 311–328. *MR* **89g**:65023

[6] Monegato, G., Stieltjes polynomials and related quadrature rules, *SIAM Rev.* **24** (1982), 137–158. *MR* **83d**:65067

[7] Riess, R. D., Gauss-Turán quadratures of Chebyshev type and error formulae, *Computing* **15** (1975), 173–179. *MR* **53** #4496

[8] Stancu, D. D., Asupra unor formule generale de integrare numerica, *Acad. R. P. Romîne Stud. Cerc. Mat.* **9** (1958), 209–216. *MR* **20** #4917

[9] Szegö, G., *Orthogonal polynomials*, 4th ed., American Mathematical Society Colloquium Publications, Vol. 23, American Mathematical Society, Providence, RI, 1975. *MR* **51** #8724

[10] Turán, P., On the theory of the mechanical quadrature, *Acta Sci. Math. (Szeged)* **12A** (1950), 30–37. *MR* **12** – 164

[11] Varma, A. K., On optimal quadrature formulae, *Studia Sci. Math. Hungar.* **19** (1984), 437–446. (Not in *MR*.)

(Received November 13, 1990)

DEPARTMENT OF MATHEMATICS
PURDUE UNIVERSITY
WEST LAFAYETTE, IN 47907
U.S.A.

Current address:

DEPARTMENT OF MATHEMATICS
SOUTHEASTERN LOUISIANA UNIVERSITY
P.O.BOX 687
HAMMOND, LA 70402
U.S.A

e-mail: FMAT1627@ SELU.EDU

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 85–90

ON A PROPERTY OF THE LINEAR n-TH ORDER INHOMOGENEOUS ALGEBRAIC DIFFERENTIAL EQUATION

T. FÉNYES

Let M denote the Mikusinski operator field and D the algebraic derivative which is defined by

$$D(f) = -tf(t), \quad f \in C$$

and for $y = \frac{a}{b}$, $a, b \in C$, $b \neq 0$,

$$D(y) = \frac{D(a)b - D(b)a}{b^2}.$$

Schatte [1] has proved the following theorem. If the homogeneous algebraic differential equation

$$D(y) + ay = 0, \quad a \in M$$

has a non-trivial solution $y_0 \in M$, then the inhomogeneous differential equation

$$(1) \qquad D(y) + ay = f, \quad f \in M$$

has a solution in M if and only if the algebraic integral

$$\int \frac{f}{y_0}$$

exists in M. A particular solution of (1) is of the form

$$x = y_0 \int \frac{f}{y_0}.$$

Fényes [2] has proved the analogous statement for the second order linear inhomogeneous algebraic differential equation in a discrete Mikusinski-type operator field based on the Cauchy product of functions defined on the non-negative integers. However, it can easily be seen from [2] that this statement also holds in the original Mikusinski field M.

In this paper we generalize Schatte's result for the arbitrary n-th order equation of the form

$$(2) \qquad D^n(y) + a_{n-1}D^{n-1}(y) + \ldots + a_1 D(y) + a_0 y = f,$$

where $a_0, a_1, \ldots, a_{n-1}, f \in M$ are given operators. We prove the following

---

1991 *Mathematics Subject Classification.* Primary 44A40.
Key words and phrases. Operational calculus.

Research supported by the Hungarian National Foundation for Scientific Research Grant No. 6032/6319.

0081–6906/94/$ 4.00

THEOREM 1. *Let us consider arbitrary differential equations of the type (2) for which the corresponding homogeneous equation*

$$D^n(y) + a_{n-1}D^{n-1}(y) + \ldots + a_1 D(y) + a_0 y = 0 \tag{3}$$

has n linearly independent solutions $y_1, y_2, \ldots, y_n$ over the field of the complex numbers. Then (2) has a solution in M if and only if the algebraic integrals

$$\int c_\nu, \quad \nu = 1, 2, \ldots n \tag{4}$$

exist in M, where the operators c_ν satisfy the equation system

$$\begin{aligned} c_1 y_1 + c_2 y_2 + \ldots + c_n y_n &= 0 \\ c_1 D(y_1) + c_2 D(y_2) + \ldots + c_n D(y_n) &= 0 \\ c_1 D^2(y_1) + c_2 D^2(y_2) + \ldots + c_n D^2(y_n) &= 0 \\ &\vdots \\ c_1 D^{n-1}(y_1) + c_2 D^{n-1}(y_2) + \ldots + c_n D^{n-1}(y_n) &= f. \end{aligned} \tag{5}$$

PROOF. *Sufficiency.* Trivial. Applying the method of variation of parameters we obtain a particular solution of (2) in the form

$$y = \sum_{\nu=1}^{n} y_\nu \int c_\nu. \tag{6}$$

This expression is meaningful since the Wronskian

$$W = W[y_1, y_2, \ldots, y_n] \neq 0$$

so (5) has a unique solution. (See Kaplansky [3].)

Necessity. We prove this by mathematical induction. For $n = 1$ we obtain Schatte's result. Let $n > 1$. Assuming that the theorem holds for $n - 1$, we show that it holds for n, too.

By applying the substitution $y = y_1 x$ (3) can be reduced to

$$D^n(x) + b_{n-1}D^{n-1}(x) + \ldots + b_1 D(x) = \frac{f}{y_1} \tag{7}$$

with some $b_1, b_2, \ldots, b_{n-1} \in M$. Introducing $D(x) = z$ we have

$$D^{n-1}(z) + b_{n-1}D^{n-2}(z) + \ldots + b_1 z = \frac{f}{y_1}. \tag{8}$$

The fundamental system of the solutions of the corresponding homogeneous equation is

$$z_1 = D\left(\frac{y_2}{y_1}\right), \quad z_2 = D\left(\frac{y_3}{y_1}\right), \ \ldots, \ z_{n-1} = D\left(\frac{y_n}{y_1}\right).$$

If y_p is a particular solution of (2), then $D\left(\frac{y_p}{y_1}\right)$ is a particular solution of (8). Since by the induction assumption the Theorem holds for (8), we have

$$D\left(\frac{y_p}{y_1}\right) = \sum_{\nu=1}^{n-1} D\left(\frac{y_{\nu+1}}{y_1}\right)\int \gamma_\nu, \tag{9}$$

where the operators γ_ν satisfy the equation system

$$\begin{aligned}
\gamma_1 D\left(\frac{y_2}{y_1}\right) + \gamma_2 D\left(\frac{y_3}{y_1}\right) + \ldots + \gamma_{n-1} D\left(\frac{y_n}{y_1}\right) &= 0,\\
\gamma_1 D^2\left(\frac{y_2}{y_1}\right) + \gamma_2 D^2\left(\frac{y_3}{y_1}\right) + \ldots + \gamma_{n-1} D^2\left(\frac{y_n}{y_1}\right) &= 0,\\
&\vdots\\
\gamma_1 D^{n-1}\left(\frac{y_2}{y_1}\right) + \gamma_2 D^{n-1}\left(\frac{y_3}{y_1}\right) + \ldots + \gamma_{n-1} D^{n-1}\left(\frac{y_n}{y_1}\right) &= \frac{f}{y_1}.
\end{aligned} \tag{10}$$

Integrating the equation (9) we get

$$y_p = y_1 \int \left[\sum_{\nu=1}^{n-1} D\left(\frac{y_{\nu+1}}{y_1}\right)\int \gamma_\nu\right]. \tag{11}$$

Since

$$\sum_{\nu=1}^{n-1} D\left(\frac{y_{\nu+1}}{y_1}\right)\int \gamma_\nu = \sum_{\nu=1}^{n-1} D\left[\frac{y_{\nu+1}}{y_1}\int \gamma_\nu\right] - \sum_{\nu=1}^{n-1} \frac{y_{\nu+1}}{y_1}\gamma_\nu, \tag{12}$$

so

$$y_p = \sum_{\nu=1}^{n-1} y_{\nu+1}\int \gamma_\nu - y_1 \int \sum_{\nu=1}^{n-1} \frac{y_{\nu+1}}{y_1}\gamma_\nu. \tag{13}$$

We show that

$$c_1 = -\sum_{\nu=1}^{n-1} \frac{y_{\nu+1}}{y_1}\gamma_\nu, \quad c_2 = \gamma_1, \quad c_3 = \gamma_2, \ldots, \quad c_n = \gamma_{n-1} \tag{14}$$

holds.

Obviously, it is easily seen that the first equation of (5) will be satisfied if we substitute (14) into it. From the first and second equations of the system (5) it follows that

$$c_2 D\left(\frac{y_2}{y_1}\right) + c_3 D\left(\frac{y_3}{y_1}\right) + \ldots + c_n D\left(\frac{y_n}{y_1}\right) = 0. \tag{15}$$

By differentiating we have

$$c_2 D^2\left(\frac{y_2}{y_1}\right) + c_3 D^2\left(\frac{y_3}{y_1}\right) + \ldots + c_n D^2\left(\frac{y_n}{y_1}\right) +$$
$$+ D(c_2) D\left(\frac{y_2}{y_1}\right) + D(c_3) D\left(\frac{y_3}{y_1}\right) + \ldots + D(c_n) D\left(\frac{y_n}{y_1}\right) = 0.$$

Since

$$\sum_{i=2}^{n} D(c_i) D\left(\frac{y_i}{y_1}\right) = \frac{1}{y_1} \sum_{i=2}^{n} D(c_i) D(y_i) - \frac{\sum_{i=2}^{n} D(c_i) y_i D(y_1)}{y_1^2} =$$
$$= -\frac{D(c_1) D(y_1)}{y_1} + \frac{D(c_1) D(y_1) y_1}{y_1^2} = 0$$

we have

$$c_2 D^2\left(\frac{y_2}{y_1}\right) + c_3 D^2\left(\frac{y_3}{y_1}\right) + \ldots + c_n D^2\left(\frac{y_n}{y_1}\right) = 0. \tag{16}$$

By continuing this procedure we see that

$$\begin{aligned} &\sum_{i=2}^{n} c_i D^{\mu}\left(\frac{y_i}{y_1}\right) = 0, \qquad \mu = 1, 2, \ldots, n-2, \\ &\sum_{i=2}^{n} c_i D^{n-1}\left(\frac{y_i}{y_1}\right) = \frac{f}{y_1}. \end{aligned} \tag{17}$$

By comparing (17) to (11) we obviously get

$$c_2 = \gamma_1, \quad c_3 = \gamma_2, \ldots, c_n = \gamma_{n-1}.$$

If we take into account (13) we can easily see that the algebraic integrals $\int c_\nu$, $\nu = 1, 2, \ldots n$ exist, so the Theorem is proved.

Now the following question arises: Can we apply the method of variation of parameters in the case that the homogeneous equation (3) has a non-trivial solution but has not n linearly independent solutions? The answer is affirmative. We shall show this for $n = 2$. There holds the following

THEOREM 2. *Let us consider the differential equation*

$$D^2(y) + aD(y) + by = f, \quad a, b, f \in M, \tag{18}$$

and let us assume that the corresponding homogeneous equation has a non-trivial solution y_0 but has not two linearly independent solutions.

If $a = 0$ then (18) has a solution in M if and only if the algebraic integrals

$$\int f y_0, \quad \int \left[\frac{1}{y_0^2} \int f y_0\right] \tag{19}$$

exist.

If $a \neq 0$ and if the differential equation

$$D(u) + au = 0 \tag{20}$$

has a non-trivial solution u_0, then (18) has a solution in M if and only if the algebraic integrals

$$\int \frac{f y_0}{u_0}, \quad \int \left[\frac{u_0}{y_0^2} \int \frac{f y_0}{u_0}\right] \tag{21}$$

exist. The solution of (18) is of the form

$$\begin{aligned} y &= y_0 \int \left[\frac{1}{y_0^2} \int f y_0\right], \quad a = 0, \\ y &= y_0 \int \left[\frac{u_0}{y_0^2} \int \frac{f y_0}{u_0}\right], \quad a \neq 0. \end{aligned} \tag{22}$$

PROOF. By applying the method of variation of parameters we look for a solution of (22) in the form

$$y = y_0 x.$$

Substituting this into (18) we get

$$D^2(x) + \left(a + \frac{2D(y_0)}{y_0}\right) D(x) = \frac{f}{y_0}. \tag{23}$$

Let us substitute $D(x) = \frac{u}{y_0^2}$ into (23) then we get

$$D(u) + au = f y_0. \tag{24}$$

If $a = 0$ then $u = \int f y_0$, $x = \int [\frac{1}{y_0^2} \int f y_0]$, and the first formula of (22) holds. If $a \neq 0$, then applying again the method of variation of parameters for the equation (24) we obtain a solution of (24) in the form

$$u = u_0 \int \frac{f y_0}{u_0}. \tag{25}$$

Moreover,

$$x = \int \left[\frac{u_0}{y_0^2} \int \frac{f y_0}{u_0} \right]$$

and the second formula of (22) holds.

REMARK. It can easily be seen that there exist only one algebraic integrals Φ_1 and Φ_2 of $f y_0$ and $\frac{f y_0}{u_0}$, respectively, for which the outer integrals $\int \frac{\Phi_1}{y_0^2}$ and $\int \frac{\Phi_2 u_0}{y_0^2}$, respectively, also exist. This follows from the fact that the homogeneous equation corresponding to (18) has not two linearly independent solutions.

REFERENCES

[1] SCHATTE, P., Funktionentheoretische Untersuchungen im Mikusińskischen Operatorenkörper, *Math. Nachr.* **35** (1967), 19-56. *MR* **38** #1521

[2] FÉNYES, T. and KOSIK, P., Az operátortest néhány speciális másodrendű differenciálegyenletéről, *Mat. Lapok* **27** (1976/79), 337–354. *MR* **80k**:44005

[3] KAPLANSKY, I., *An introduction to differential algebra*, Actualités Sci. Ind., No.1251, Hermann, Paris, 1957. *MR* **20** #177

(Received January 7, 1991)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 91–106

GENERALIZATION OF POISSON SINGULAR INTEGRALS AND THEIR CONVERGENCE, ORDER OF APPROXIMATION AND SATURATION PROPERTIES

JIA-DING CAO

Abstract

Let $s>0$ and $f\in C_{2\pi}$. We construct new singular integrals $u_r^{(s)}(f,x)$. If $s=1$, we obtain Poisson singular integrals, if $s=2$ we obtain the Ghermanesco operators. $u_r^{(s)}(f,x)$ $(s>1)$ are called Poisson singular integrals of higher order. We show convergence for $s\geqq\frac{1}{2}$ and show that convergence does not hold in general for $s<\frac{1}{2}$. We investigate the order of approximation for $s\geqq\frac{1}{2}$ and solve the saturation problem for $s\geqq\frac{3}{2}$.

§ 1. New singular integrals

Let $f(t)$ be a 2π-periodic continuous function, we denote this by $f(t)\in\in C_{2\pi}$. Also, let $\|f\|\overset{\text{def}}{=}\max\limits_{0\leqq t\leqq 2\pi}|f(t)|$. Let $f\in C_{2\pi}$ and $0\leqq r<1$,

$$P_r(t)\overset{\text{def}}{=}\frac{1}{2}+\sum_{i=1}^{\infty}r^i\cos it=\frac{1}{2}\frac{1-r^2}{1-2r\cos t+r^2}. \tag{1.1}$$

Poisson singular integrals [1] are given by

$$u_r(f,x)\overset{\text{def}}{=}\frac{1}{\pi}\int\limits_{-\pi}^{\pi}f(x+t)P_r(t)dt. \tag{1.2}$$

Let $s>0$,

$$Q_r(t)\overset{\text{def}}{=}\frac{1}{1-2r\cos t+r^2},\quad I^{(s)}(r)\overset{\text{def}}{=}\int\limits_{-\pi}^{\pi}[Q_r(t)]^s dt>0.$$

We construct new singular integrals by putting

$$u_r^{(s)}(f,x)\overset{\text{def}}{=}\frac{1}{I^{(s)}(r)}\int\limits_{-\pi}^{\pi}f(t+x)[Q_r(t)]^s dt. \tag{1.3}$$

1980 *Mathematics Subject Classification* (1985 Revision). Primary 41A35; Secondary 41A36, 41A25, 41A40.

Key words and phrases. Poisson integrals, Ghermanesco operators, convergence, order of approximation, saturation classes.

0081-6906/94/$ 4.00

If $s=1$, from (1.1) we have

$$I^{(1)}(r)=\int_{-\pi}^{\pi}\frac{1}{1-2r\cos t+r^2}dt=\frac{2}{1-r^2}\int_{-\pi}^{\pi}\frac{1-r^2}{2(1-2r\cos t+r^2)}dt=\frac{2\pi}{1-r^2},$$

hence

$$u_r^{(1)}(f,x)=\frac{(1-r^2)}{2\pi}\int_{-\pi}^{\pi}\frac{f(t+x)}{1-2r\cos t+r^2}dt=u_r(f,x).$$

From (1.1), using Parseval's equality [1] we have

$$\int_{-\pi}^{\pi}P_r^2(t)dt=\pi\Big(\frac{1}{2}+\sum_{i=1}^{\infty}r^{2i}\Big)=\pi\Big(\frac{1}{2}+\frac{r^2}{1-r^2}\Big)=\frac{\pi}{2}\frac{1+r^2}{1-r^2}.$$

If $s=2$, then

$$u_r^{(2)}(f,x)=\frac{1}{I^{(2)}(r)}\int_{-\pi}^{\pi}f(t+x)Q_r^{(2)}(t)dt=$$

$$=\frac{1}{\int_{-\pi}^{\pi}P_r^2(t)dt}\int_{-\pi}^{\pi}f(t+x)P_r^2(t)dt=\frac{2}{\pi}\frac{1-r^2}{1+r^2}\int_{-\pi}^{\pi}f(t+x)P_r^2(t)dt.$$

Thus we obtain the Ghermanesco operators (see [2]).

Matsuoka [3] introduced Jackson singular integrals of higher order, in several joint articles with H. H. Gonska (see [4–10]) we studied several of their properties and applications. If $s>1$, the operators $u_r^{(s)}(f,x)$ are called Poisson singular integrals of higher order. Below we solve the problem of convergence of $u_r^{(s)}(f,x)$ $(s\geqq\frac{1}{2})$, we show that the condition of convergence $s\geqq\frac{1}{2}$ cannot be improved, we investigate their order of approximation $(s\geqq\frac{1}{2})$ and solve the saturation problem for $s\geqq\frac{3}{2}$.

§ 2. Some lemmas

LEMMA 1. *We have* $Q_r(t)\leqq\frac{1}{(1-r)^2}$, $0\leqq t\leqq\pi$, *and* $Q_r(t)\leqq\frac{\pi^2}{t^2}$, $0<t\leqq\pi$.

PROOF. We have

$$Q_r(t)=\frac{1}{1-2r\cos t+r^2}=\frac{1}{(1-r)^2+2r(1-\cos t)}=$$
$$=\frac{1}{(1-r)^2+4r\sin^2\frac{t}{2}}\leqq\frac{1}{(1-r)^2}. \tag{2.1}$$

Since $\frac{\sin u}{u}$ is a decreasing function for $0 < u \leqq \frac{\pi}{2}$, and

$$t^2 Q_r(t) = \left(\frac{(1-r)^2}{t^2} + \frac{r \sin^2 \frac{t}{2}}{(\frac{t}{2})^2} \right)^{-1}, \quad 0 < t \leqq \pi,$$

we obtain that $t^2 Q_r(t)$ is an increasing function for $0 < t \leqq \pi$. Hence

$$t^2 Q_r(t) \leqq \pi^2 Q_r(\pi) = \frac{\pi^2}{1 + 2r + r^2} = \frac{\pi^2}{(1+r)^2} \leqq \pi^2,$$

from where we obtain Lemma 1. □

LEMMA 2. *As $r \to 1-0$, we have*

$$I^{(s)}(r) = \begin{cases} O\left(\frac{1}{(1-r)^{2s-1}}\right), & s > \frac{1}{2}, \\ O(|\ln(1-r)|), & s = \frac{1}{2}, \\ O(1), & 0 < s < \frac{1}{2}. \end{cases}$$

PROOF. From Lemma 1 we get

$$[Q_r(t)]^s \leqq \frac{1}{(1-r)^{2s}}, \quad [Q_r(t)]^s \leqq \frac{\pi^{2s}}{t^{2s}}, \quad 0 < t \leqq \pi.$$

Hence

$$\begin{aligned} I^{(s)}(r) &= 2 \int_0^{\pi} [Q_r(t)]^s dt = 2 \int_0^{1-r} [Q_r(t)]^s dt + 2 \int_{1-r}^{\pi} [Q_r(t)]^s dt \leqq \\ &\leqq 2 \int_0^{1-r} \frac{dt}{(1-r)^{2s}} + 2\pi^{2s} \int_{1-r}^{\pi} \frac{dt}{t^{2s}} \leqq \frac{2}{(1-r)^{2s-1}} + 2\pi^{2s} \int_{1-r}^{\pi} \frac{dt}{t^{2s}}. \end{aligned} \tag{2.2}$$

If $s > \frac{1}{2}$ $(2s > 1)$, then

$$\int_{1-r}^{\pi} \frac{dt}{t^{2s}} \leqq \int_{1-r}^{\infty} t^{-2s} dt = \frac{(1-r)^{-2s+1}}{2s-1} = O\left(\frac{1}{(1-r)^{2s-1}}\right).$$

Thus, from (2.2) we have $I^{(s)}(r) = O(\frac{1}{(1-r)^{2s-1}})$ (see also Duren's book [17, Ch. 4, § 4.6]).

If $s = \frac{1}{2}$ $(2s = 1)$, then

$$\int_{1-r}^{\pi} \frac{dt}{t^{2s}} = \int_{1-r}^{\pi} \frac{dt}{t} = \ln \pi - \ln(1-r) = \ln \pi + |\ln(1-r)| = O(|\ln(1-r)|), \; r \to 1-0.$$

Hence from (2.2) we have

$$I^{(\frac{1}{2})}(r) \leqq 2 + O(|\ln(1-r)|) = O(|\ln(1-r)|), \qquad r \to 1-0.$$

If $0 < s < \frac{1}{2}$ $(2s < 1)$, then

$$\int_{1-r}^{\pi} \frac{dt}{t^{2s}} \leqq \int_{0}^{\pi} \frac{dt}{t^{2s}} = O(1).$$

So from (2.2) it follows that

$$I^{(s)}(r) \leqq 2(1-r)^{1-2s} + O(1) = O(1), \quad r \to 1-0,$$

and the proof of Lemma 2 is complete. □

LEMMA 3. *Let $s > 0$, then*

$$I^{(s)}(r) \geqq \frac{2^{1-s}}{(1-r)^{2s-1}}, \quad 0 < r < 1.$$

PROOF. Since $\sin \frac{t}{2} \leqq \frac{t}{2}$, $0 \leqq t \leqq \pi$, and due to (2.1), we have

$$I^{(s)}(r) = 2\int_0^{\pi} [Q_r(t)]^s dt \geqq 2\int_0^{\pi} \frac{dt}{[(1-r)^2 + rt^2]^s} \geqq$$

(2.3)

$$\geqq 2\int_0^{1-r} \frac{dt}{[(1-r)^2 + r(1-r)^2]^s} = \frac{2(1-r)}{(1+r)^s(1-r)^{2s}} \geqq \frac{2^{1-s}}{(1-r)^{2s-1}}. \quad \square$$

LEMMA 4. *Let $s = \frac{1}{2}$, then*

$$I^{(\frac{1}{2})}(r) > \frac{2}{\sqrt{r}}(|\ln(1-r)|), \quad 0 < r < 1.$$

PROOF. From (2.3) we obtain

(2.4)

$$I^{(\frac{1}{2})}(r) = 2\int_0^{\pi} [Q_r(t)]^{1/2} dt \geqq 2\int_0^{\pi} \frac{dt}{\sqrt{(1-r)^2 + rt^2}} = \frac{2}{\sqrt{r}} \int_0^{\pi\sqrt{r}} \frac{dv}{\sqrt{(1-r)^2 + v^2}}.$$

Using the fact that (see [11])

$$\int \frac{dv}{\sqrt{v^2 + a^2}} = \ln(v + \sqrt{v^2 + a^2}) + c, \quad a > 0,$$

it follows that

$$(2.5)\quad 2\int_0^{\pi}\frac{dt}{\sqrt{(1-r)^2+rt^2}}=\frac{2}{\sqrt{r}}\Big\{\ln(\pi\sqrt{r}+\sqrt{(1-r)^2+\pi^2r})-\ln(1-r)\Big\}.$$

For $0<r<1$, $\sqrt{r}>r$, so that

$$(2.6)\quad \begin{aligned}&\ln(\pi\sqrt{r}+\sqrt{(1-r)^2+\pi^2r})>\ln(\pi\sqrt{r}+1-r)>\ln(\pi r+1-r)=\\&=\ln(1+(\pi-1)r)>0.\end{aligned}$$

Combining (2.4), (2.5), (2.6) we have

$$I^{(\frac{1}{2})}(r)>-\frac{2}{\sqrt{r}}\ln(1-r)=\frac{2}{\sqrt{r}}|\ln(1-r)|. \quad \square$$

§ 3. Convergence and order of approximation

We define $\rho_{1.r}^{(s)}\overset{\text{def}}{=}u_r^{(s)}(\cos t,0)$.

LEMMA 5. *As $r\to 1-0$, we have*

$$1-\rho_{1,r}^{(s)}=\begin{cases}O((1-r)^2), & (s>\frac{3}{2}),\\ O((1-r)^2|\ln(1-r)|), & (s=\frac{3}{2}),\\ O((1-r)^{2s-1}), & (0<s<\frac{3}{2}).\end{cases}$$

PROOF. We have

$$1-\rho_{1,r}^{(s)}=\frac{1}{I^{(s)}(r)}\int_{-\pi}^{\pi}(1-\cos t)[Q_r(t)]^s dt=$$

$$(3.1)\quad =\frac{4}{I^{(s)}(r)}\int_0^{\pi}\sin^2\frac{t}{2}[Q_r(t)]^s dt\leqq\frac{1}{I^{(s)}(r)}\int_0^{\pi}t^2[Q_r(t)]^s dt=$$

$$=\frac{1}{I^{(s)}(r)}\int_0^{1-r}t^2[Q_r(t)]^s dt+\frac{1}{I^{(s)}(r)}\int_{1-r}^{\pi}t^2[Q_r(t)]^s dt\overset{\text{def}}{=}J_1^{(s)}(r)+J_2^{(s)}(r).$$

If $s>0$, using Lemma 1, it can be seen that

$$[Q_r(t)]^s\leqq\frac{1}{(1-r)^{2s}}.$$

From Lemma 3 we get

$$J_1^{(s)}(r) = O((1-r)^{2s-1})\,O\Big(\frac{1}{(1-r)^{2s}}\Big)\Big(\int_0^{1-r} t^2\,dt\Big) = \tag{3.2}$$
$$= O((1-r)^{2s-1-2s})\,O((1-r)^3) = O((1-r)^2).$$

Using Lemma 1 we have

$$[Q_r(t)]^s \leqq \frac{\pi^{2s}}{t^{2s}}, \quad (0 < t \leqq \pi),$$

$$J_2^{(s)}(r) \overset{\text{def}}{=} \frac{1}{I^{(s)}(r)} \int_{1-r}^{\pi} t^2 [Q_r(t)]^s\,dt = \tag{3.3}$$

$$= O((1-r)^{2s-1}) \int_{1-r}^{\pi} \frac{t^2}{t^{2s}}\,dt = O((1-r)^{2s-1}) \int_{1-r}^{\pi} \frac{dt}{t^{2s-2}}.$$

If $s > \frac{3}{2}$ $(2s-2>1)$, then

$$\int_{1-r}^{\pi} \frac{dt}{t^{2s-2}} \leqq \int_{1-r}^{\infty} t^{2-2s}\,dt = \frac{(1-r)^{3-2s}}{2s-3} = O((1-r)^{3-2s}),$$

hence

$$J_2^{(s)}(r) = O((1-r)^{2s-1})\,O((1-r)^{3-2s}) = O((1-r)^2). \tag{3.4}$$

If $s = \frac{3}{2}$ $(2s-2=1)$, then

$$\int_{1-r}^{\pi} \frac{dt}{t^{2s-2}} = \int_{1-r}^{\pi} \frac{dt}{t} = \ln\pi - \ln(1-r) =$$
$$= \ln\pi + |\ln(1-r)| = O(|\ln(1-r)|), \qquad (r \to 1-0),$$

hence

$$J_2^{(\frac{3}{2})}(r) = O((1-r)^{3-1})\,O(|\ln(1-r)|) = \tag{3.5}$$
$$= O((1-r)^2\,|\ln(1-r)|).$$

If $0 < s < \frac{3}{2}$ $(2s-2<1)$, then

$$\int_{1-r}^{\pi} \frac{dt}{t^{2s-2}} \leqq \int_{0}^{\pi} \frac{dt}{t^{2s-2}} = O(1),$$

hence

$$J_2^{(s)}(r) = O((1-r)^{2s-1}). \tag{3.6}$$

Combining (3.1)–(3.6) we obtain Lemma 5. □

LEMMA 6. *If $s = \frac{1}{2}$, then*

$$1 - \rho_{1,r}^{(\frac{1}{2})} = O\left(\frac{1}{|\ln(1-r)|}\right), \quad r \to 1-0.$$

PROOF. If $s = \frac{1}{2}$, using Lemma 1 and Lemma 4 we have $[Q_r(t)]^{1/2} \leqq \frac{\pi}{t}$, $0 < t \leqq \pi$ and

$$I^{(\frac{1}{2})}(r) > \frac{2}{\sqrt{r}}|\ln(1-r)|, \quad (0<r<1).$$

From (3.1) we arrive at

$$J_2^{(\frac{1}{2})}(r) = \frac{1}{I^{(\frac{1}{2})}(r)} \int_{1-r}^{\pi} t^2 [Q_r(t)]^{1/2} dt \leqq$$

$$\leqq \frac{\sqrt{r}}{2} \frac{1}{|\ln(1-r)|} \int_{1-r}^{\pi} t^2 \frac{\pi}{t} dt \leqq \frac{\pi}{2} \frac{1}{|\ln(1-r)|} \int_{1-r}^{\pi} t dt \leqq \tag{3.7}$$

$$\leqq \frac{\pi}{2} \frac{1}{|\ln(1-r)|} \int_{0}^{\pi} t dt = O\left(\frac{1}{|\ln(1-r)|}\right), \quad (0<r<1).$$

Combining (3.1), (3.2), (3.7) we obtain

$$1 - \rho_{1,r}^{(\frac{1}{2})} = O((1-r)^2) + O\left(\frac{1}{|\ln(1-r)|}\right) = O\left(\frac{1}{|\ln(1-r)|}\right), \quad r \to 1-0. \ \square$$

For $f \in C_{2\pi}$, the second order modulus of continuity of f is given by

$$\omega_2(f,\delta) \overset{\text{def}}{=} \sup_{|h| \leqq \delta} \|f(x+h) + f(x-h) - 2f(x)\|.$$

THEOREM 1. *Let* $s \geqq \frac{1}{2}$ *and* $f \in C_{2\pi}$. *As* $r \to 1-0$,

$$\|u_r^{(s)}(f,x) - f(x)\| = \begin{cases} O(\omega_2(f, 1-r)), & (s > \frac{3}{2}), \\ O(\omega_2(f, (1-r)\sqrt{|\ln(1-r)|})), & (s = \frac{3}{2}), \\ O(\omega_2(f, (1-r)^{s-\frac{1}{2}})), & (\frac{3}{2} > s > \frac{1}{2}), \\ O(\omega_2(f, \frac{1}{\sqrt{|\ln(1-r)|}})), & (s = \frac{1}{2}). \end{cases}$$

PROOF. The kernel of the singular integrals $u_r^{(s)}(f,x)$ is $\frac{1}{I^{(s)}(r)}[Q_r(t)]^s$. This kernel is even and non-negative. From the book of Butzer and Nessel (see [12], Theorem 1.5.8) it is known that

$$\|u_r^{(s)}(f,x) - f(x)\| = O\Big(\omega_2\Big(f, \sqrt{1-\rho_{1,r}^{(s)}}\Big)\Big).$$

Thus from Lemma 5 and Lemma 6 we obtain Theorem 1. □

THEOREM 2. *Let* $s \geqq \frac{1}{2}$ *and* $f \in C_{2\pi}$. *Then*

$$\lim_{r\to 1-0} \|u_r^{(s)}(f,x) - f(x)\| = 0.$$

PROOF. The proof is an immediate consequence of Theorem 1. □

For $s = 1$, we obtain a result in Korovkin's book [13]. Let C_i be positive constants only depending on s.

THEOREM 3. *Let* $0 < s < \frac{1}{2}$, *then*

$$1 - u_r^{(s)}(\cos t, 0) \not\to 0,$$

i.e., the convergence condition $s \geqq \frac{1}{2}$ *cannot be improved.*

PROOF. Using the inequality $\sin u \geqq \frac{2}{\pi}u$, $0 \leqq u \leqq \frac{\pi}{2}$, we have

$$\begin{aligned} 1 - u_r^{(s)}(\cos t, 0) &= 1 - \rho_{1,r}^{(s)} = \frac{1}{I^{(s)}(r)} \int_{-\pi}^{\pi} (1-\cos t)[Q_r(t)]^s dt = \\ &= \frac{4}{I^{(s)}(r)} \int_0^{\pi} \sin^2 \frac{t}{2} [Q_r(t)]^s dt \geqq \frac{4}{\pi^2 I^{(s)}(r)} \int_0^{\pi} t^2 [Q_r(t)]^s dt. \end{aligned} \tag{3.8}$$

If $0 < s < \frac{1}{2}$, using Lemma 2 we have that $I^{(s)}(r) = O(1) \leqq C_1$, hence

$$\begin{aligned} 1 - u_r^{(s)}(\cos t, 0) &\geqq \frac{4}{\pi^2 C_1} \int_0^{\pi} \frac{t^2 dt}{[(1-r)^2 + 4r\sin^2 \frac{t}{2}]^s} \geqq \\ &\geqq \frac{4}{\pi^2 C_1} \int_0^{\pi} \frac{t^2 dt}{[(1-r)^2 + rt^2]^s}. \end{aligned} \tag{3.9}$$

Define

$$\varphi(r,t) \overset{\text{def}}{=} \frac{t^2}{[(1-r)^2+rt^2]^s}, \quad 0 \leqq r < 1,\ 0 \leqq t \leqq \pi,$$

$$\varphi(1,t) \overset{\text{def}}{=} t^{2-2s}, \quad (r=1,\ 0 \leqq t \leqq \pi).$$

It is easily seen that $\varphi(r,t)$ is a continuous function of two variables on the rectangle $[0 \leqq r \leqq 1] \times [0 \leqq t \leqq \pi]$. Hence

$$\lim_{r\to 1-0} \int_0^\pi \frac{t^2}{[(1-r)^2+rt^2]^s} dt = \int_0^\pi \varphi(1,t)dt = \int_0^\pi t^{2-2s} dt = \frac{\pi^{3-2s}}{3-2s}.$$

From (3.9) we have

$$\varliminf_{r\to 1-0} (1-u_r^{(s)}(\cos t,0)) \geqq$$

$$\geqq \frac{4}{\pi^2 C_1} \varliminf_{r\to 1-0} \int_0^\pi \frac{t^2}{[(1-r)^2+rt^2]^s} dt \geqq \frac{4}{\pi^2 C_1} \frac{\pi^{3-2s}}{3-2s} \overset{\text{def}}{=} C_2 > 0,$$

hence $1-u_r^{(s)}(\cos t,0) \not\to 0$, and

$$\varliminf_{r\to 1-0} \|\cos x - u_r^{(s)}(\cos t,x)\| \geqq \varliminf_{r\to 1-0} |1-u_r^{(s)}(\cos t,0)| \geqq C_2 > 0.$$

Thus our convergence condition $s \geqq \frac{1}{2}$ cannot be improved. □

§ 4. Some further lemmas

LEMMA 7. *Let $s > \frac{1}{2}$. Then*

$$1-\rho_{1,r}^{(s)} \geqq C_4(1-r)^2, \quad 0<r<1.$$

PROOF. From (3.8) we have

$$1-\rho_{1,r}^{(s)} \geqq \frac{4}{\pi^2 I^{(s)}(r)} \int_0^\pi t^2[Q_r(t)]^s dt = \frac{4}{\pi^2 I^{(s)}(r)} \int_0^\pi \frac{t^2}{[(1-r)^2+4r\sin^2\frac{t}{2}]^s} dt.$$

Using Lemma 2, for $s > \frac{1}{2}$ we have

$$I^{(s)}(r) \leqq \frac{C_3}{(1-r)^{2s-1}},$$

hence

$$1-\rho_{1,r}^{(s)} \geqq \frac{4(1-r)^{2s-1}}{C_3\pi^2}\int_0^{\pi}\frac{t^2}{[(1-r)^2+rt^2]^s}dt \geqq$$

$$\geqq \frac{4}{C_3\pi^2}(1-r)^{2s-1}\int_0^{1-r}\frac{t^2}{[(1-r)^2+rt^2]^s}dt \geqq$$

$$\geqq \frac{4}{C_3\pi^2}(1-r)^{2s-1}\int_0^{1-r}\frac{t^2}{[(1-r)^2+r(1-r)^2]^s}dt = \tag{4.1}$$

$$= \frac{4}{C_3\pi^2}(1-r)^{2s-1}\frac{(1-r)^3}{3}\frac{1}{(1-r)^{2s}(1+r)^s} \geqq$$

$$\geqq \frac{4}{3\cdot 2^s\cdot\pi^2 C_3}(1-r)^{2s-1-2s+3} = C_4(1-r)^2. \qquad \square$$

LEMMA 8. *Let $s=\frac{3}{2}$, then there exists $0<r_1<1$ with*

$$1-\rho_{1,r}^{(\frac{3}{2})} \geqq C_5(1-r)^2|\ln(1-r)|, \quad 0<r_1<r<1.$$

PROOF. From (4.1) for $s=\frac{3}{2}$ we have

$$1-\rho_{1,r}^{(\frac{3}{2})} \geqq \frac{4}{C_3\pi^2}(1-r)^2\int_0^{\pi}\frac{t^2}{[(1-r)^2+rt^2]^{3/2}}dt. \tag{4.2}$$

Furthermore,

$$\int_0^{\pi}\frac{t^2\,dt}{[(1-r)^2+rt^2]^{3/2}} = \frac{1}{r}\left\{\int_0^{\pi}\left[\frac{rt^2+(1-r)^2-(1-r)^2}{(\sqrt{(1-r)^2+rt^2})^3}\right]dt\right\} =$$

$$= \frac{1}{r}\int_0^{\pi}\frac{dt}{\sqrt{(1-r)^2+rt^2}} - \frac{(1-r)^2}{r}\int_0^{\pi}\frac{dt}{(\sqrt{(1-r)^2+rt^2})^3} \stackrel{\text{def}}{=} \tag{4.3}$$

$$\stackrel{\text{def}}{=} h_1(r)-h_2(r).$$

From (2.5) we have

$$h_1(r) = \frac{1}{r\sqrt{r}}\left\{\ln(\pi\sqrt{r}+\sqrt{\pi^2r+(1-r)^2}\,)-\ln(1-r)\right\} =$$

$$= O(1)+\frac{1}{r\sqrt{r}}|\ln(1-r)|, \quad r\to 1-0,$$

whence

$$\lim_{r\to 1-0}\frac{h_1(r)}{|\ln(1-r)|}=1. \tag{4.4}$$

Furthermore

$$h_2(r)=\frac{(1-r)^2}{r}\int_0^{\pi}\frac{dt}{(\sqrt{(1-r)^2+rt^2})^3}=\frac{(1-r)^2}{r}\frac{1}{\sqrt{r}}\int_0^{\pi\sqrt{r}}\frac{dv}{(\sqrt{v^2+(1-r)^2})^3}.$$

Using the equation (see, e.g., [11])

$$\int\frac{dv}{(\sqrt{v^2+a^2})^3}=\frac{v}{a^2\sqrt{v^2+a^2}}+C,\qquad a>0,$$

we get

$$h_2(r)=\frac{(1-r)^2}{r\sqrt{r}}\frac{\pi\sqrt{r}}{(1-r)^2\sqrt{r\pi^2+(1-r)^2}}=\frac{\pi}{r\sqrt{r\pi^2+(1-r)^2}},$$

so that

$$\lim_{r\to 1-0}h_2(r)=\frac{\pi}{\pi}=1,\quad\text{and}\quad\lim_{r\to 1-0}\frac{h_2(r)}{|\ln(1-r)|}=0. \tag{4.5}$$

Combining (4.3), (4.4) and (4.5) we conclude that

$$\lim_{r\to 1-0}\int_0^{\pi}\frac{t^2dt}{[(1-r)^2+rt^2]^{3/2}}\Big/|\ln(1-r)|=1.$$

Hence there exists r_1, $0<r_1<1$, so that for $0<r_1<r<1$ we have

$$\int_0^{\pi}\frac{t^2}{[(1-r)^2+rt^2]^{3/2}}dt\Big/|\ln(1-r)|>\frac{1}{2}.$$

Finally, from (4.2) we have

$$1-\rho_{1,r}^{(\frac{3}{2})}\geqq\frac{4}{C_3\pi^2}\frac{1}{2}(1-r)^2|\ln(1-r)|=C_5(1-r)^2|\ln(1-r)|.\ \square$$

Let

$$\tau_r^{(s)}\stackrel{\text{def}}{=}u_r^{(s)}(\sin^4\frac{t}{2},0)=\frac{1}{I^{(s)}(r)}\int_{-\pi}^{\pi}\sin^4\frac{t}{2}[Q_r(t)]^s dt.$$

LEMMA 9. *As* $r \to 1-0$ *we have*

$$\tau_r^{(s)} = \begin{cases} O((1-r)^4), & (s > \frac{5}{2}), \\ O((1-r)^4 |\ln(1-r)|), & (s = \frac{5}{2}), \\ O((1-r)^{2s-1}), & (0 < s < \frac{5}{2}). \end{cases}$$

PROOF. Let $s > 0$. Then

$$\tau_r^{(s)} = \frac{1}{I^{(s)}(r)} \int_{-\pi}^{\pi} \sin^4 \frac{t}{2} [Q_r(t)]^s dt \leqq$$

$$\leqq \frac{1}{8I^{(s)}(r)} \int_0^{1-r} t^4 [Q_r(t)]^s dt + \frac{1}{8I^{(s)}(r)} \int_{1-r}^{\pi} t^4 [Q_r(t)]^s dt \tag{4.6}$$

$$\stackrel{\text{def}}{=} R_1^{(s)}(r) + R_2^{(s)}(r).$$

Using Lemma 1 we have

$$[Q_r(t)]^s \leqq \frac{1}{(1-r)^{2s}}, \quad 0 \leqq t \leqq \pi, \quad \text{and} \quad [Q_r(t)]^s \leqq \frac{\pi^{2s}}{t^{2s}}, \quad 0 < t \leqq \pi.$$

From Lemma 3 we get

$$R_1^{(s)}(r) \leqq O((1-r)^{2s-1}) \int_0^{1-r} t^4 dt \frac{1}{(1-r)^{2s}} = O((1-r)^4), \tag{4.7}$$

$$R_2^{(s)}(r) \leqq O((1-r)^{2s-1}) \int_{1-r}^{\pi} \frac{dt}{t^{2s-4}}. \tag{4.8}$$

Now, for $s > \frac{5}{2}$ $(2s - 4 > 1)$, we have

$$\int_{1-r}^{\pi} \frac{dt}{t^{2s-4}} \leqq \int_{1-r}^{\infty} t^{4-2s} dt = \frac{(1-r)^{5-2s}}{2s-5},$$

hence

$$R_2^{(s)}(r) = O((1-r)^{2s-1}) O((1-r)^{5-2s}) = O((1-r)^4). \tag{4.9}$$

Combining (4.6)–(4.9) yields

$$\tau_r^{(s)} = O((1-r)^4), \quad s > \frac{5}{2}.$$

If $s = \frac{5}{2}$ $(2s-4=1)$,

$$\int_{1-r}^{\pi} \frac{dt}{t^{2s-4}} = \int_{1-r}^{\pi} \frac{dt}{t} = \ln \pi + |\ln(1-r)| = O(|\ln(1-r)|),$$

so that from (4.8) we have

$$R_2^{(\frac{5}{2})}(r) = O((1-r)^{5-1})O(|\ln(1-r)|) = O((1-r)^4|\ln(1-r)|). \tag{4.10}$$

Combining (4.6)–(4.8) and (4.10) we have

$$\tau_r^{(s)} = O((1-r)^4|\ln(1-r)|), \quad s = \frac{5}{2}.$$

If $0 < s < \frac{5}{2}$ $(2s-4<1)$, then

$$\int_{1-r}^{\pi} \frac{dt}{t^{2s-4}} \leqq \int_{0}^{\pi} \frac{dt}{t^{2s-4}} = O(1),$$

such that from (4.8) we now obtain

$$R_2^{(s)}(r) = O((1-r)^{2s-1})O(1) = O((1-r)^{2s-1}). \tag{4.11}$$

A combination of (4.6)–(4.8) and (4.11) now implies

$$\tau_r^{(s)} = O((1-r)^4) + O((1-r)^{2s-1}) = O((1-r)^{2s-1}), \quad 0 < s < \frac{5}{2}. \ \square$$

§ 5. The saturation problem

Let N be the set of natural numbers. Saturation classes for summation methods of Fourier series were defined by Favard (see [14], [15]). Let (L_r) $(0 \leqq r < 1)$ be a family of linear operators mapping $C_{2\pi}$ into itself. Assume that $\lim_{r\to 1-0} \varphi(r) = 0$, and let K denote a class of functions in $C_{2\pi}$. Assume that $\|f - L_r(f)\| = O(\varphi(r))$ holds if and only if $f \in K$ and that $\|f - L_r(f)\| = o(\varphi(r))$ holds if and only if f is a constant. Then (L_r) is said to be saturated with order $\varphi(r)$ ($\varphi(r)$ is the optimal approximation order) and K is called the saturation class of (L_r).

Let $f \in C_{2\pi}$ and (L_r), $0 \leqq r < 1$, be a family of positive linear convolution operators. That is, we assume that each (L_r) has the form

$$L_r(f, x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t+x)\,d\mu_r(t), \tag{5.1}$$

where $d\mu_r$ is a non-negative, even Borel measure on $[-\pi,\pi)$ with $\frac{1}{\pi}\int_{-\pi}^{\pi} d\mu_r(t)=1$,

$$\lambda_{k,r} \stackrel{\text{def}}{=} L_r(\cos kt, 0) = \frac{1}{\pi}\int_{-\pi}^{\pi} \cos kt\, d\mu_r(t), \quad k \in \mathbf{N}.$$

LEMMA 10. *Let (L_r), $0 \leqq r < 1$, be a family of positive convolution operators of the form* (5.1). *If for each $k \in \mathbf{N}$ we have*

$$\lim_{r\to 1-0} \frac{1-\lambda_{k,r}}{1-\lambda_{1,r}} = k^2, \tag{5.2}$$

then (L_r) is saturated with order $1-\lambda_{1,r}$ and the saturation class is $\{f \in C_{2\pi} \mid f' \in \operatorname{Lip} 1\}$.

PROOF. See DeVore's book [14, Chapter 3, Theorem 3.6], or Tureckiĭ [15]. □

LEMMA 11. *Let (L_r), $0 \leqq r < 1$, be a family of positive convolution operators of the form* (5.1). *Then the following two statements are equivalent:*

(i) $$\lim_{r\to 1-0} \frac{1-\lambda_{k,r}}{1-\lambda_{1,r}} = k^2, \text{ for } k \in \mathbf{N},$$

(ii) $$\int_{-\pi}^{\pi} \sin^4 \frac{t}{2} d\mu_r(t) = o(1-\lambda_{1,r}), \qquad (r \to 1-0).$$

PROOF. See DeVore's book [14, Chapter 3, Theorem 3.8]. □

THEOREM 4. *Let $s > \frac{3}{2}$. As $r \to 1-0$, we have*

$$\|f(x) - u_r^{(s)}(f,x)\| = o((1-r)^2) \leftrightarrow f \equiv \text{const.},$$

and if and only if $f' \in \operatorname{Lip} 1$ we have

$$\|f(x) - u_r^{(s)}(f,x)\| = O((1-r)^2).$$

PROOF. If $s > \frac{3}{2}$, then from Lemma 5 and Lemma 7 we have

$$C_7(1-r)^2 \leqq 1 - \rho_{1,r}^{(s)} \leqq C_6(1-r)^2. \tag{5.3}$$

If $s \geqq \frac{5}{2}$, Lemma 9 shows $\tau_r^{(s)} = o((1-r)^2)$. If $\frac{3}{2} < s < \frac{5}{2}$ $(2s-3>0)$, using Lemma 9 we also get

$$\tau_r^{(s)} = O((1-r)^{2s-1}) = O((1-r)^2(1-r)^{2s-3}) = o((1-r)^2).$$

On using (5.3) , for $s > \frac{3}{2}$, we have

$$\tau_r^{(s)} = o((1-r)^2) = o(1)(1-r)^2 \leqq o(1)\frac{1}{C_7}(1-\rho_{1,r}^{(s)}) = o(1-\rho_{1,r}^{(s)}), \quad r \to 1-0.$$

Using Lemma 10 and Lemma 11 finally we get Theorem 4. □

THEOREM 5. *Let $s = \frac{3}{2}$. As $r \to 1-0$ we have*

$$\|f - u_r^{(\frac{3}{2})}(f)\| = o((1-r)^2|\ln(1-r)|) \leftrightarrow f \equiv \text{const.},$$

and if and only if $f' \in \operatorname{Lip} 1$ we have

$$\|f - u_r^{(\frac{3}{2})}(f)\| = O((1-r)^2|\ln(1-r)|).$$

PROOF. Using Lemma 5 and Lemma 8 for $s = \frac{3}{2}$, we have

$$C_9(1-r)^2|\ln(1-r)| \leqq 1 - \rho_{1,r}^{(\frac{3}{2})} \leqq C_8(1-r)^2|\ln(1-r)|, \tag{5.4}$$

$0 < r_1 < r < 1$. By Lemma 9 and (5.4), we have

$$\tau_r^{(\frac{3}{2})} = O((1-r)^{2s-1}) = O((1-r)^2) = O(1)(1-r)^2|\ln(1-r)|\frac{1}{|\ln(1-r)|} \leqq$$
$$\leqq O(1)\frac{1}{C_9}(1-\rho_{1,r}^{(\frac{3}{2})})\frac{1}{|\ln(1-r)|} = o((1-\rho_{1,r}^{(\frac{3}{2})})), \quad r \to 1-0.$$

Lemma 10 and Lemma 11 now imply Theorem 5. □

For $s = 2$ we obtain a theorem of Stark [16].

PROBLEM 1. The determination of saturation classes of $u_r^{(s)}$ for $\frac{1}{2} \leqq s < \frac{3}{2}$ is an interesting problem. These saturation classes are linear manifolds; see Favard's lecture at the 3rd mathematical conference of the Soviet Union [18] and the author's article [19].

REMARK 1. Shi Shen-Liang introduced the A_α method of summation of series [20], A_α is a generalization of Abel's method of the summation of series ($\alpha = 1$); he also introduced a generalization of u_r.

REFERENCES

[1] ZYGMUND, A., *Trigonometric series*, 2nd edition, Cambridge University Press, New York, 1959. *MR* **21** #6498

[2] STARK, E. L., Erzeugung und strukturelle Verknüpfungen von Kernen singulärer Faltungsintegrale, *Approximation theory* (Proc. Internat. Colloq., Inst. Angew. Math., Univ. Bonn, 1976), Springer, Berlin, 1976, 390–402. *Zbl* **335** #42002 and *MR* **58** #29659

[3] MATSUOKA, Y., On the approximation of functions by some singular integrals, *Tôhoku Math. J.* (2)**18** (1966), 13–43. *MR* **34** #3178

[4] CAO, J. D., Generalization of Timan's theorem, Lehnhoff's theorem and Teljakovskii's theorem, Schriftenreihe des Fachbereichs Mathematik, Universität Duisburg, SM-DU-106 (1986), 1–11.

[5] CAO, J. D. and GONSKA, H. H., Approximation by Boolean sums of positive linear operators, *Rend. Mat. Appl.* (7)**6** (1986), 525–546. *MR* **90**d:41041a

[6] CAO, J. D. and GONSKA, H. H., Approximation by Boolean sums of positive linear operators II. Gopengauz-type estimates, *J. Approx. Theory* **57** (1989), 77–89. *MR* **90d**:41041b

[7] CAO, J. D. and GONSKA, H. H., Pointwise estimates for modified positive linear operators, *Portugal. Math.* **46** (1989), 401–430. *MR* **91a**:41026

[8] CAO, J. D. and GONSKA, H. H., Computation of DeVore–Gopengauz-type approximants, *Approximation theory VI*, Vol. 1 (College Station, TX, 1989), Academic Press, Boston, MA, 1989, 117–120. (See *MR* **91j**:41001.)

[9] CAO, J. D., Pointwise estimates of Groetsch–Shisha type, *Approximation Theory VI*, Vol. 1 (College Station, TX, 1989), Academic Press, Boston, MA, 1989, 113–115. (See *MR* **91j**:41001.)

[10] CAO, J. D. and GONSKA, H. H., Approximation by Boolean sums of positive linear operators III. Estimates for some numerical approximation schemes, *Numer. Funct. Anal. Optim.* **10** (1989), 643–672. *MR* **90**k:41024

[11] XU, Q. F., *Table of integrals*, Press of Science and Technic of Shanghai, 1959.

[12] BUTZER, P. L. and NESSEL, R. J., *Fourier analysis and approximation*, Vol. I, One-dimensional theory, Pure and Applied Mathematics, Vol. 40, Academic Press, New York–London, 1971. *MR* **58** # 23312

[13] KOROVKIN, P. P., *Linear operators and approximation theory*, Russian Monographs and Texts on Advanced Mathematics and Physics, Vol. III, Gordon and Breach, New York; Hindustan Publ. Corp., Delhi, 1960. *MR* **27** #561

[14] DEVORE, R. A., *The approximation of continuous functions by positive linear operators*, Lecture Notes in Mathematics, Vol. 293, Springer-Verlag, Berlin–New York, 1972. *MR* **54** # 8100

[15] TURECKIĬ, A. H., Saturation classes in a space *C*, *Izv. Akad. Nauk. SSSR Ser. Mat.* **25** (1961), 411–442 (in Russian). *MR* **23** # A1988

[16] STARK, E. L., Nikolskii constants for positive singular integrals of perturbed Fejér-type, *Linear operators and approximation* (Proc. Conf., Math. Res. Inst., Oberwolfach, 1971), Internat. Ser. Numer. Math., Vol. 20, Birkhäuser, Basel, 1972, 348–363. *MR* **55** # 6075

[17] DUREN, P. L., *Theory of H^p-spaces*, Pure and Applied Mathematics, Vol. 38, Academic Press, New York–London, 1970. *MR* **42** #3552

[18] FAVARD, J., On the theory of approximation of functions: development of the theory and problems, *Proc. 3th Math. Conf. Soviet Union*, Vol. 4, 168–172 (in Russian).

[19] CAO, J. D., Solution of an approximation problem of J. Favard, *Acta Math. Sinica* **21** (1978), 72–76 (in Chinese). *MR* **81**b:42026

[20] SHI, S. L., Introduction of some new methods of summation to the theory of Fourier series, *J. Hangchow Univ. (Nat. Sci. Edition)* **1** (1963), 59–88 (in Chinese).

(Received November 10, 1990)

DEPARTMENT OF MATHEMATICS
FUDAN UNIVERSITY
SHANGHAI 200433
PEOPLE'S REPUBLIC OF CHINA

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 107–164

EIN KREISÜBERDECKUNGSPROBLEM AUF DER SPHÄRE

G. BLIND und R. BLIND

INHALT

1980 *Mathematics Subject Classification* (1985 Revision). Primary 52A45.
Key words and phrases. Arrangement of congruent circles on the sphere, simply covered area, density.

1. Einleitung

Auf der Einheitssphäre S^2 seien $n \geqq 3$ kongruente, abgeschlossene, sphärische Kreise gegeben. Die zwei klassischen Probleme dazu lauten

1. Die Kreise bilden speziell eine Packung. Wie groß ist die maximale Dichte $\overline{D}_n$ einer Kreispackung aus n kongruenten Kreisen?

2. Die Kreise bilden speziell eine Überdeckung. Wie groß ist die minimale Dichte $\underline{D}_n$ einer Kreisüberdeckung aus n kongruenten Kreisen?

Es sei $\omega_n := \frac{n}{n-2}\frac{\pi}{6}$. Dann gelten die Abschätzungen

$$\overline{D}_n \leqq \frac{n}{2}\left(1 - \frac{1}{2\cos\omega_n}\right) \quad \text{und} \quad \underline{D}_n \geqq \frac{n}{2}\left(1 - \frac{1}{\sqrt{3}\tan\omega_n}\right). \tag{1.1}$$

Diese Abschätzungen sind für $n = 3, 4, 6, 12$ scharf; die Kreismittelpunkte sind dann die Ecken eines regulären Dreiecks, Tetraeders, Oktaeders bzw. Ikosaeders. Für $n \to \infty$ gehen die Schranken in (1.1) gegen die optimalen Dichten der entsprechenden ebenen Probleme (siehe [4, S. 114]). Für andere Werte von n ist $\overline{D}_n$ bzw. $\underline{D}_n$ in einzelnen Fällen bekannt, in manchen Fällen gibt es Vermutungen oder gute Abschätzungen, siehe etwa [5] und [6].

Ein Kreissystem zerlegt die S^2 in mehrfach, einfach und überhaupt nicht überdeckte Bereiche. Beim Problem 1 wird der von den Kreisen einfach überdeckte Teil der S^2 abgeschätzt unter der Voraussetzung, daß es keinen mehrfach überdeckten Teil gibt. In [4, S. 97] wird nun folgendes (ebene) Problem gestellt: Der wievielte Teil der Ebene läßt sich durch *beliebig* gelegene kongruente Kreise *einfach* überdecken? Geht man von der dichtesten Packung kongruenter Kreise aus und vergrößert die Kreise konzentrisch, bis jeder Kreis von den 6 benachbarten in den Ecken eines regulären 12-Ecks geschnitten wird, so überdeckt das entstehende Kreissystem 100 $(\sqrt{48} - 6)$% der Ebene einfach. Es gilt: $\vartheta := \sqrt{48} - 6$ ist die maximale Dichte des von einem beliebigen System kongruenter Kreise einfach überdeckten Bereichs des E^2. Dies wurde unter starken Voraussetzungen an das Kreissystem in [4] bzw. in [1] bewiesen (siehe auch [7]), und schließlich ohne jede Voraussetzung in [2].

Analog läßt sich auf der Sphäre fragen

3. Gegeben seien $n \geqq 3$ kongruente, abgeschlossene, sphärische Kreise $K_1, \ldots, K_n$[1]. $E(K_1, \ldots, K_n)$ sei der davon *einfach* überdeckte Bereich der S^2. Wie groß ist

$$\vartheta_n := \max \frac{|E(K_1, \ldots, K_n)|}{4\pi},$$

wobei sich das Maximum auf *alle* Familien aus n kongruenten Kreisen bezieht?

[1] Für den Radius ϱ der Kreise ist $0 < \varrho < \pi$ zugelassen.

Wir zeigen

SATZ 1. *Es gilt*

$$(1.2)\qquad \begin{aligned} \vartheta_n &\leqq \frac{3}{\pi}(n-2)\Big(\omega_n - \pi + 2\arccos\frac{1}{4\cos\frac{\omega_n}{2}}\Big) =: \\ &=: \frac{3}{\pi}(n-2)F(\omega_n) =: S_n \qquad (n \geqq 3), \end{aligned}$$

mit

$$(1.3)\qquad \omega_n := \frac{n}{n-2}\frac{\pi}{6}.$$

Diese Abschätzung ist genau für $n = 3, 4, 6, 12$ scharf; die Kreismittelpunkte sind dann die Ecken eines regulären Dreiecks, Tetraeders, Oktaeders bzw. Ikosaeders, und die Kreisradien sind (analog zum ebenen Fall) so groß, daß jeder Kreis die k benachbarten in den Ecken eines regulären $2k$-Ecks schneidet.

Die Abschätzung (1.2) läßt sich folgendermaßen interpretieren: Ein Dreiecksnetz auf S^2 mit den n Kreismittelpunkten als Ecken besteht aus $2n-4$ Dreiecken vom durchschnittlichen Inhalt $\frac{4\pi}{2n-4} = 6\omega_n - \pi =: |\Delta_n|$. $2\omega_n$ ist deshalb der Winkel eines gleichseitigen sphärischen Dreiecks Δ_n vom Inhalt $|\Delta_n|$. Drei kongruente Kreise in den Ecken von Δ_n bestimmen in Δ_n einen einfach überdeckten Bereich, dessen maximaler Flächeninhalt $6F(\omega_n)$ ist, wie in Hilfssatz 4.2 gezeigt werden wird; dabei tritt $6F(\omega_n)$ genau für denjenigen Kreisradius auf, bei dem die in Δ_n gelegenen Kreisränder genau zur Hälfte einfach überdeckt sind. Es ist

$$S_n = \frac{(2n-4)6F(\omega_n)}{4\pi}.$$

Deshalb ist die Abschätzung (1.2) für $n = 3, 4, 6, 12$ scharf, wenn die Kreismittelpunkte und Kreisradien sich wie beschrieben verhalten.

S_n ist wachsend in n, weil

$$S_n = \frac{1}{\omega_n - \frac{\pi}{6}}F(\omega_n)$$

in ω_n fallend ist, wie man durch Ableiten sieht. Für $n \to \infty$ geht S_n gegen die maximale Dichte des ebenen Problems. Es ist

n	3	4	5	6	12	$n \to \infty$
S_n	0,8098...	0,8814...	0,8987...	0,9066...	0,9199...	0,9282...

Es sei ϱ der Radius der Kreise $K_1, \ldots, K_n$ $(0 < \varrho < \pi)$. Wegen [3] genügt es Satz 1 zu beweisen unter den Voraussetzungen

(1.4) $$n \geqq 4,$$

(1.5) $$\varrho < \varrho_0 := \arccos \frac{1}{\sqrt{7}} \quad (\varrho_0 \text{ ist der optimale Kreisradius im Fall } n = 3),$$

(1.6) nicht alle Kreismittelpunkte liegen in einer abgeschlossenen Halbsphäre.

Satz 1 wird nach folgendem Grundgedanken bewiesen: Man zerlege S^2 in eine Familie von gleichschenkligen Dreiecken $\triangle$ so, daß die Basisecken von $\triangle$ Mittelpunkte von Kreisen K_i, K_j sind, und daß für die einfach überdeckten Bereiche $E(K_1, \dots, K_n) \cap \triangle \subset E(K_i, K_j) \cap \triangle$ gilt. Dies führt zu einer Abschätzung von $|E(K_1, \dots, K_n) \cap \triangle|$ für jedes $\triangle$. Mit Hilfe der Jensenschen Ungleichung gewinne man daraus eine Abschätzung für $|E(K_1, \dots, K_n)|$. Im allgemeinen gibt es aber weder eine solche Zerlegung, noch gilt die Jensensche Ungleichung, so daß dieses Programm nur sehr bedingt und sehr modifiziert durchführbar ist. Der Beweis von Satz 1 gliedert sich entsprechend in einen numerischen Teil I, die Untersuchung von $\sum_{\triangle} |E(K_i, K_j) \cap \triangle|$, und in einen geometrischen Teil II, in dem es um Zerlegungen der Sphäre S^2 geht. Teil I ist so abgefaßt, daß zum Verständnis von Teil II nur die Punkte 2., ..., 5. benötigt werden.

Eine Liste der verwendeten Bezeichnungen steht am Schluß der Arbeit.

Das Ende eines Beweises ist mit □ markiert.

I. GLEICHSCHENKLIGE DREIECKE UND KONGRUENTE KREISE UM DIE BASISECKEN

2. g-Dreiecke

Gegeben sei ein gleichschenkliges sphärisches Dreieck $\triangle$ mit der Basis $\overline{A_1A_2}$, der Spitze C, dem Basiswinkel α und dem Winkel 2β an der Spitze. Weil $\triangle$ ein sphärisches Dreieck ist, gilt $\alpha + \beta \geqq \frac{\pi}{2}$. Außerdem gelte $r := \frac{1}{2}|\overline{A_1A_2}| \leqq \frac{\pi}{2}$ und $0 \leqq \alpha, \beta \leqq \frac{\pi}{2}$; es werden also *entartete* Dreiecke mit $\alpha = 0$ oder $\beta = 0$ zugelassen. Es gilt

(2.1) $$\cos r = \frac{\cos \beta}{\sin \alpha} \quad \text{für} \quad \alpha + \beta \geqq \frac{\pi}{2}, \quad \alpha \neq 0.$$

Es sei ϱ fest gegeben mit

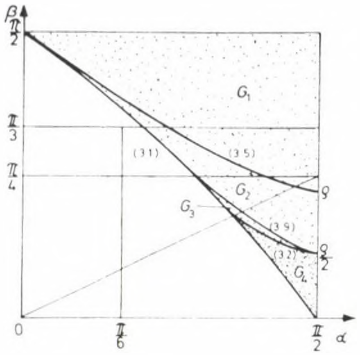


Fig. 1

$0<\varrho<\frac{\pi}{2}$, $G=G_1\cup G_2\cup G_3\cup G_4$

$G_1:\begin{cases} r\geqq\varrho \text{ für } \alpha>0, \\ \text{zusätzlich der Punkt } (0,\frac{\pi}{2})\end{cases}$

$G_2:\begin{cases}\alpha\neq 0 \\ r\leqq\varrho\end{cases}$ und $\begin{cases}\beta\geqq\frac{\pi}{4} \text{ oder} \\ \beta\leqq\frac{\pi}{4} \text{ und } d\geqq\varrho\end{cases}$

$G_3: r>\frac{\varrho}{2}$ und $\beta\leqq\frac{\pi}{4}$ und $d\leqq\varrho$

$G_4: r\leqq\frac{\varrho}{2}$

breite g-Dreiecke

schmale g-Dreiecke

$$0<\varrho<\frac{\pi}{2}. \tag{2.2}$$

$K_\varrho(A_1)$, $K_\varrho(A_2)$ seien sphärische Kreise mit Radius ϱ um A_1 bzw. A_2. Im folgenden werden außer den entarteten nur solche Dreiecke $\triangle$ betrachtet, daß $(K_\varrho(A_1)\cup K_\varrho(A_2))^0$ nicht $\triangle$ überdeckt. In der (α,β)-Ebene entspricht ihnen der Bereich

(2.3)
$$G:=\left\{\left(0,\frac{\pi}{2}\right)\right\}\cup\left\{\left(\frac{\pi}{2},0\right)\right\}\cup\left\{(\alpha,\beta)\,|\,0<\alpha,\beta\leqq\frac{\pi}{2} \text{ und } \cos\varrho\geqq\cot\alpha\cot\beta\right\},$$

vgl. Fig. 1. Gleichschenklige Dreiecke mit $(\alpha,\beta)\in G$ und $r\leqq\frac{\pi}{2}$ heißen *g-Dreiecke (bzgl. ϱ)*. Ein g-Dreieck heißt *breit*, wenn es entartet ist mit $\alpha=0$ oder wenn $r>\frac{\varrho}{2}$, sonst *schmal*.

$E(K_\varrho(A_1),K_\varrho(A_2))\cap\triangle$ ist der von $\{K_\varrho(A_1),K_\varrho(A_2)\}$ in $\triangle$ einfach überdeckte Bereich. Für schmale g-Dreiecke $\triangle$ sei der Punkt $D:=\triangle\cap\partial K_\varrho(A_1)\cap\cap\,\partial K_\varrho(A_2)$. Dann sei $e_\varrho(\triangle)$ die folgendermaßen definierte Funktion, die $|E(K_\varrho(A_1),K_\varrho(A_2))\cap\triangle|$ nach oben abschätzt:

(2.4)
$$e_\varrho(\triangle):=\begin{cases}|E(K_\varrho(A_1),K_\varrho(A_2))\cap\triangle| & \text{für breite g-Dreiecke}\\ |((K_\varrho(A_1)\cup K_\varrho(A_2))\cap\triangle\setminus A_1A_2D| & \text{für schmale g-Dreiecke,}\end{cases}$$

siehe Fig. 5. $e_\varrho(\triangle)$ ist eine von α,β und ϱ abhängige Funktion, und es sei

$$2e_\varrho(\alpha,\beta):=e_\varrho(\triangle) \quad \text{für} \quad (\alpha,\beta)\in G. \tag{2.5}$$

Gesucht wird eine geeignete Abschätzung für $\sum_i e_\varrho(\alpha_i,\beta_i)$ nach oben.

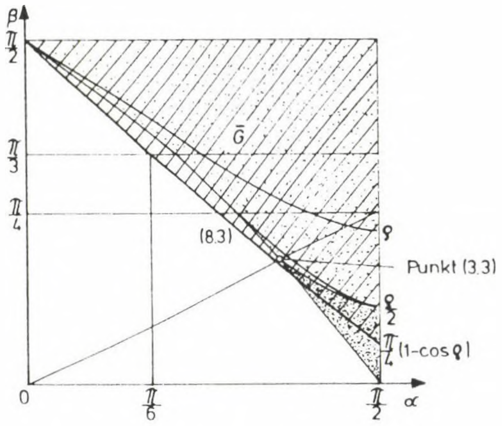


Fig. 2

3. Berechnung der Funktion $e_\varrho(\alpha,\beta)$

In der (α,β)-Ebene betrachten wir in der gesamten Arbeit nur Punkte mit $0 \leqq \alpha \leqq \frac{\pi}{2}$, $0 \leqq \beta \leqq \frac{\pi}{2}$, siehe Fig. 1.

Bei festem ϱ ist $e_\varrho(\alpha,\beta)$ definiert für $(\alpha,\beta) \in G$. Nach (2.3) wird G begrenzt durch die Kurve

$$\tan\alpha\tan\beta = \frac{1}{\cos\varrho}, \tag{3.1}$$

eine konkave Funktion durch $(0,\frac{\pi}{2})$ und $(\frac{\pi}{2},0)$. Schmale und breite Dreiecke werden getrennt durch die Kurve

$$r = \frac{\varrho}{2} \quad \text{oder wegen (2.1)} \quad \cos\beta = \sin\alpha\cos\frac{\varrho}{2}, \tag{3.2}$$

eine konvexe Funktion durch $(\frac{\pi}{2},\frac{\varrho}{2})$, dort waagrechte Tangente. Es gilt

$$\left(2\arcsin\frac{1}{2\cos\frac{\varrho}{2}},\ \arcsin\frac{1}{2\cos\frac{\varrho}{2}}\right) \tag{3.3}$$

ist der Schnittpunkt von (3.1) und (3.2), auf der Geraden $\alpha = 2\beta$ gelegen.

Für die Berechnung von $e_\varrho(\alpha,\beta)$ wird wesentlich sein, ob $r \gtreqqless \varrho$. Es gilt

$$r \gtreqqless \varrho \iff \cos\beta \lesseqqgtr \sin\alpha\cos\varrho, \tag{3.4}$$

und

$$r = \varrho \quad \text{oder} \quad \cos\beta = \sin\alpha\cos\varrho \tag{3.5}$$

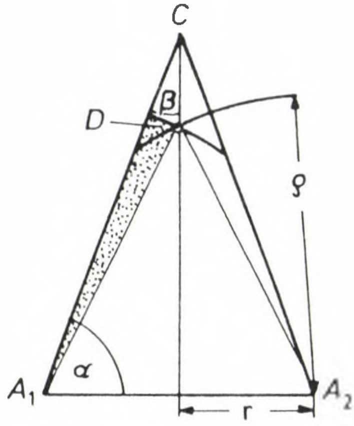


Fig. 5

In G_1 ist $e_\varrho(\alpha,\beta)$ der Flächeninhalt eines Kreisausschnitts, so daß

$$e_\varrho(\alpha,\beta)=\alpha(1-\cos\varrho) \quad \text{in } G_1. \tag{3.10}$$

Wenn g-Dreiecke einem Punkt von G_2 entsprechen, so ist $K_\varrho(A_1)\cap$ $\cap\, K_\varrho(A_2)\neq\emptyset$ und $\partial K_\varrho(A_2)\cap\overline{A_1C}$ enthält höchstens einen Punkt (siehe Fig. 3). Der in Fig. 3 schraffierte Kreisabschnitt hat den Flächeninhalt

$$2\gamma(1-\cos\varrho)-(2\gamma+2\delta-\pi)=\pi-2\gamma\cos\varrho-2\delta.$$

Deshalb gilt

$$\begin{aligned} e_\varrho(\alpha,\beta)&=\alpha(1-\cos\varrho)-\pi+2\gamma\cos\varrho+2\delta=\\ &=\alpha(1-\cos\varrho)-\pi+2\cos\varrho\arccos\frac{\tan r}{\tan\varrho}+2\arcsin\frac{\sin r}{\sin\varrho}\quad \text{in } G_2, \end{aligned} \tag{3.11}$$

mit

$$\cos r=\frac{\cos\beta}{\sin\alpha},\quad \gamma=\arccos\frac{\sqrt{\sin^2\alpha-\cos^2\beta}}{\cos\beta\tan\varrho},\quad \delta=\arcsin\frac{\sqrt{\sin^2\alpha-\cos^2\beta}}{\sin\alpha\sin\varrho}.$$

Wenn g-Dreiecke einem Punkt von G_3 entsprechen, besteht $\partial K_\varrho(A_2)\cap\overline{A_1C}$ aus zwei (eventuell zusammenfallenden) Punkten (siehe Fig. 4). Deshalb ist

$$\begin{aligned} e_\varrho(\alpha,\beta)&=\alpha(1-\cos\varrho)+2\cos\varrho\arccos\frac{\tan r}{\tan\varrho}+2\arcsin\frac{\sin r}{\sin\varrho}-\\ &\quad -2\cos\varrho\arccos\frac{\tan d}{\tan\varrho}-2\arcsin\frac{\sin d}{\sin\varrho}\quad \text{in } G_3, \end{aligned} \tag{3.12}$$

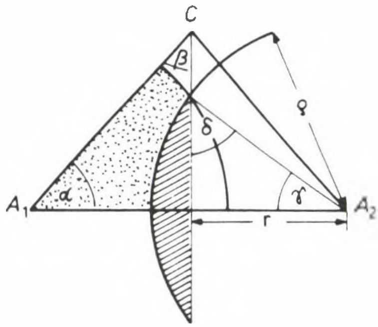


Fig. 3

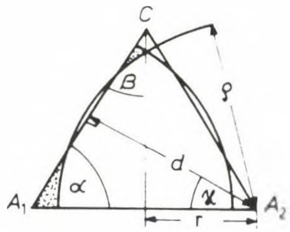


Fig. 4

ist eine konvexe Funktion durch $(0, \frac{\pi}{2})$ und $(\frac{\pi}{2}, \varrho)$, waagrechte Tangente in $(\frac{\pi}{2}, \varrho)$.

Im g-Dreieck $\triangle$ habe das Lot von A_2 aus auf $\overline{A_1C}$ die Länge d und schließe mit $\overline{A_1A_2}$ den Winkel κ ein (vgl. Fig. 4). Es ist

$$\sin d = \sin 2r \sin \alpha = 2 \cos \beta \sin r = 2 \frac{\cos \beta}{\sin \alpha} \sqrt{\sin^2 \alpha - \cos^2 \beta}. \tag{3.6}$$

Es wird wesentlich sein, ob $\kappa \gtreqless \alpha$ und im Fall $\kappa \leqq \alpha$, ob $d \geqq \varrho$. Offensichtlich ist

$$\kappa \gtreqless \alpha \iff \beta \gtreqless \frac{\pi}{4}. \tag{3.7}$$

Nach (3.6) und (2.1) gilt für $\beta \leqq \frac{\pi}{4}$

$$d \gtreqless \varrho \iff \sin \alpha \gtreqless \frac{2 \cos^2 \beta}{\sqrt{4 \cos^2 \beta - \sin^2 \varrho}} \quad \text{für } \beta \leqq \frac{\pi}{4}. \tag{3.8}$$

Es gilt für $\beta \leqq \frac{\pi}{4}$

$$d = \varrho \text{ oder } \sin \alpha = \frac{2 \cos^2 \beta}{\sqrt{4 \cos^2 \beta - \sin^2 \varrho}} \tag{3.9}$$

ist eine konvexe Funktion, schneidet (3.1) genau für $\beta = \frac{\pi}{4}$, schneidet (3.2) genau in $(\frac{\pi}{2}, \frac{\varrho}{2})$, liegt in G oberhalb von (3.2).

Der Definitionsbereich G von $e_\varrho(\alpha, \beta)$ wird nun mit Hilfe von (3.5), (3.9) und (3.2) in 4 Bereiche G_1, G_2, G_3 und G_4 aufgeteilt gemäß Fig. 1. Dabei haben G_1 und G_2 eine gemeinsame Randkurve, ebenso G_2 und G_3.

mit

$$\cos r = \frac{\cos\beta}{\sin\alpha}, \quad \sin d = 2\frac{\cos\beta}{\sin\alpha}\sqrt{\sin^2\alpha - \cos^2\beta}.$$

Für schmale g-Dreiecke gilt nach Definition (2.4) (siehe Fig. 5)

$$e_\varrho(\alpha,\beta) = \left(\alpha - \arccos\frac{\tan r}{\tan\varrho}\right)(1-\cos\varrho) \quad \text{in } G_4, \tag{3.13}$$

mit

$$\cos r = \frac{\cos\beta}{\sin\alpha}.$$

Offensichtlich ist $e_\varrho(\alpha,\beta)$ eindeutig bestimmt.

4. Die Funktion $e_\varrho(\alpha,\frac{\pi}{3})$ bei variablem ϱ, und die Funktion $F(\alpha)$

Um den Definitionsbereich von $e_\varrho(\alpha,\frac{\pi}{3})$ in der (α,ϱ)-Ebene zu bestimmen, beachte man, daß $G = G(\varrho)$, $G_i = G_i(\varrho)$ für $i = 1,\dots,4$ (siehe Fig. 1). $e_\varrho(\alpha,\frac{\pi}{3})$ ist definiert genau für $(G_1(\varrho)\cup G_2(\varrho))\cap\{(\alpha,\frac{\pi}{3})\mid 0\leqq\alpha\leqq\frac{\pi}{2}\}$. Für festes ϱ wird $G_2(\varrho)$ begrenzt durch die Kurven (3.1) und (3.5). Es ist

$$\alpha_1 = \arctan\frac{1}{\sqrt{3}\cos\varrho} \qquad \left(0<\varrho<\frac{\pi}{2}\right) \tag{4.1}$$

der Schnitt von (3.1) mit $\beta = \frac{\pi}{3}$ für festes ϱ, streng monoton wachsend in ϱ, für variables ϱ Kurve in der (α,ϱ)-Ebene durch $(\frac{\pi}{6},0)$ und $(\frac{\pi}{2},\frac{\pi}{2})$, siehe Fig. 6;

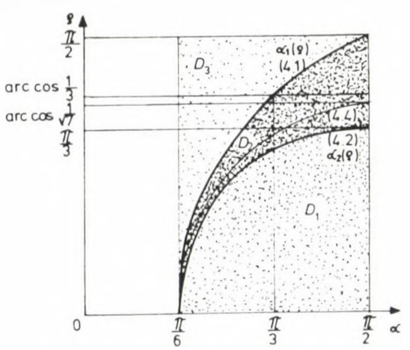


Fig. 6

$0<\varrho<\frac{\pi}{2}$

$D_1:\ \alpha_2(\varrho)\leqq\alpha\leqq\frac{\pi}{2}$

$D_2:\ \begin{cases}\alpha_1(\varrho)\leqq\alpha\leqq\alpha_2(\varrho) & \text{für } \varrho\leqq\frac{\pi}{3}\\ \alpha_1(\varrho)\leqq\alpha\leqq\frac{\pi}{2} & \text{für } \varrho\geqq\frac{\pi}{3}\end{cases}$

$D_3:\ \frac{\pi}{6}\leqq\alpha\leqq\alpha_1(\varrho).$

$$\alpha_2 = \arcsin\frac{1}{2\cos\varrho} \qquad \left(0<\varrho\leqq\frac{\pi}{3}\right) \tag{4.2}$$

ist der Schnitt von (3.5) mit $\beta = \frac{\pi}{3}$ für festes ϱ, streng monoton wachsend in ϱ, für variables ϱ Kurve in der (α,ϱ)-Ebene durch $(\frac{\pi}{6},0)$ und $(\frac{\pi}{2},\frac{\pi}{3})$, siehe Fig. 6. D_1, D_2 und D_3 seien die durch $\alpha_1(\varrho)$ und $\alpha_2(\varrho)$ bestimmten Bereiche wie in Fig. 6. Dann ist $e_\varrho(\alpha,\frac{\pi}{3})$ genau in $D_1\cup D_2$ definiert.

Es gilt

HILFSSATZ 4.1. *Zu jedem α und ϱ aus dem Definitionsbereich von $e_\varrho(\alpha, \frac{\pi}{3})$ ist*

$$e_\varrho\left(\alpha, \frac{\pi}{3}\right) \leqq F(\alpha) = \alpha - \pi + 2\arccos \frac{1}{4\cos\frac{\alpha}{2}}. \tag{4.3}$$

Das Gleichheitszeichen gilt genau dann, wenn (α, ϱ) ein Punkt der folgenden Kurve ist

$$\begin{gathered} \varrho = \arctan \frac{\sqrt{4\sin^2\alpha - 1}}{\cos\frac{\alpha}{2}} \quad \textit{... streng monoton wachsend,} \\ \textit{durch } \left(\frac{\pi}{6}, 0\right) \textit{ und } \left(\frac{\pi}{2}, \arccos\frac{1}{\sqrt{7}}\right), \textit{ siehe Fig. 6.} \end{gathered} \tag{4.4}$$

Dies ist äquivalent dazu, daß $\left(\alpha, \frac{\pi}{3}\right) \in G_2(\varrho)$ und daß für das entsprechende g-Dreieck der Winkel γ in Fig. 3 halb so groß ist wie α.

BEWEIS. $e_\varrho(\alpha, \frac{\pi}{3})$ ist genau in $D_1 \cup D_2$ definiert und dort gilt

$$e_\varrho\left(\alpha, \frac{\pi}{3}\right) = \begin{cases} \alpha(1 - \cos\varrho) \text{ in } D_1 \\ \alpha(1 - \cos\varrho) - \pi + 2\gamma\cos\varrho + 2\delta \text{ mit} \\ \qquad \gamma = \arccos \frac{\sqrt{4\sin^2\alpha - 1}}{\tan\varrho}, \\ \qquad \delta = \arcsin \frac{\sqrt{4\sin^2\alpha - 1}}{2\sin\alpha\sin\varrho} \text{ in } D_2. \end{cases}$$

Nun sei α_0 mit $\frac{\pi}{6} < \alpha_0 \leqq \frac{\pi}{2}$ fest gegeben. Gesucht wird $\sup\{e_\varrho(\alpha_0, \frac{\pi}{3}) \mid (\alpha_0, \varrho) \in D_1 \cup D_2\}$. Weil $e_\varrho(\alpha_0, \frac{\pi}{3})$ in D_1 streng monoton wachsend in ϱ ist, ist

$$\sup\left\{e_\varrho\left(\alpha_0, \frac{\pi}{3}\right) \mid (\alpha_0, \varrho) \in D_1 \cup D_2\right\} = \sup\left\{e_\varrho\left(\alpha_0, \frac{\pi}{3}\right) \mid (\alpha_0, \varrho) \in D_2\right\}.$$

Durch Ableiten sieht man, daß das Supremum angenommen wird genau für $2\gamma(\alpha_0) = \alpha_0$. Hierdurch wird jedem α_0 ein optimales ϱ_0 zugeordnet gemäß (4.4). Weil $(\alpha_0, \varrho_0) \in D_2$ genau dann, wenn $(\alpha_0, \frac{\pi}{3}) \in G_2(\varrho_0)$, sind also (α_0, ϱ_0) genau die in Hilfssatz 4.1 beschriebenen Punktepaare (α, ϱ).

Nun entnimmt man Fig. 3, daß $\cos\delta = \cos r \sin\gamma$. Damit gilt

$$\max_\varrho e_\varrho\left(\alpha_0, \frac{\pi}{3}\right) = \alpha_0 - \pi + 2\arccos\frac{1}{4\cos\frac{\alpha_0}{2}},$$

also (4.3). □

In **1** wurde der folgende Hilfssatz 4.2 zur Interpretation von Satz 1 benützt; er wird im folgenden nicht mehr benötigt.

Hilfssatz 4.2. *Δ_n sei ein gleichseitiges sphärisches Dreieck mit den Winkeln $2\omega_n$ (siehe (1.3)), und es seien 3 kongruente Kreise vom Radius ϱ $(0 < \varrho < \pi)$ um die Ecken von Δ_n gegeben. Der von ihnen in Δ_n einfach überdeckte Bereich hat dann höchstens den Flächeninhalt $6F(\omega_n)$, und dies tritt genau für denjenigen Kreisradius ϱ auf, bei dem die in Δ_n gelegenen Kreisränder genau zur Hälfte einfach überdeckt sind.*

Beweis. Δ_n sei also ein gleichseitiges sphärisches Dreieck mit den Winkeln $2\omega_n$. Kongruente Kreise vom Radius ϱ um die Ecken von Δ_n bestimmen in Δ_n einen einfach überdeckten Bereich vom Flächeninhalt $E_\varrho(\Delta_n)$. Wird $\sup E_\varrho(\Delta_n)$ gesucht, so läßt sich o.B.d.A. annehmen, daß der Schwerpunkt von Δ_n nicht von den Kreisen überdeckt wird. Deshalb ist $\varrho < \frac{\pi}{2}$ und Δ_n läßt sich in drei g-Dreiecke zerlegen, die für jedes ϱ einem Punkt aus $G_1(\varrho) \cup \cup G_2(\varrho)$ entsprechen. Deshalb ist $E_\varrho(\Delta_n) = 6e_\varrho(\omega_n, \frac{\pi}{3})$. Aus Hilfssatz 4.1 folgt dann die Behauptung. □

Man rechnet nach, daß gilt

Hilfssatz 4.3.

$$F(\alpha) = \alpha - \pi + 2 \arccos \frac{1}{4 \cos \frac{\alpha}{2}} \qquad \left(\frac{\pi}{6} \leqq \alpha \leqq \frac{\pi}{2}\right) \tag{4.5}$$

ist streng monoton wachsend, streng konkav, $F(\frac{\pi}{6}) = 0$. □

5. Der im geometrischen Teil benötigte Satz 2

In Hilfssatz 4.1 wurde gezeigt, daß $e_\varrho(\alpha, \frac{\pi}{3}) \leqq F(\alpha)$. Im Hinblick auf Satz 1 wäre wünschenswert, daß die folgende Ungleichung gilt

$$\sum_{i=1}^{k} e_\varrho(\alpha_i, \beta_i) \leqq k e_\varrho\left(\bar{\alpha}, \frac{\pi}{3}\right) \leqq kF(\bar{\alpha}) \quad \text{für } \bar{\alpha} := \frac{\sum \alpha_i}{k},$$

wenn $\bar{\beta} := \frac{\sum \beta_i}{k} = \frac{\pi}{3}$, oder allgemeiner, wenn $\bar{\beta} \leqq \frac{\pi}{3}$. Weil der Definitionsbereich G von $e_\varrho(\alpha, \beta)$ nicht konvex ist, und weil sich $e_\varrho(\alpha, \beta)$ in G_3 und in G_4 als nicht konkav erweisen wird, ist dies nicht ohne weiteres behauptbar, und im allgemeinen auch falsch. Eine derartige Ungleichung läßt sich jedoch zeigen unter geeigneten Bedingungen an die Punkte (α_i, β_i): Die Anzahl und Lage der Punkte aus G_4 wird durch die Lage der Punkte aus $G_1 \cup G_2$ eingeschränkt. Es gilt

Satz 2. *Zu gegebenem ϱ mit*

$$0 < \varrho < \varrho_0 := \arccos \frac{1}{\sqrt{7}} \quad (\textit{siehe } (1.5)) \tag{5.1}$$

sei $G_1, \dots, G_4$ wie in Fig. 1 definiert. Gegeben seien k Punkte (α_i, β_i) mit $\frac{\Sigma \beta_i}{k} \leq \frac{\pi}{3}$. Für ganze Zahlen $\mu, \nu \geqq 0$ mit $\nu \leqq 2\mu$ gelte

$$(\alpha_i, \beta_i) \in G_4 \quad \text{für} \quad 1 \leqq i \leqq \mu \tag{5.2}$$

$$(\alpha_i, \beta_i) \in G_1 \cup G_2 \quad \text{für} \quad \mu + 1 \leqq i \leqq \mu + \nu \tag{5.3}$$

$$(\alpha_i, \beta_i) \in G_1 \cup G_2 \cup G_3 \quad \text{für} \quad \mu + \nu + 1 \leqq i \leqq k, \tag{5.4}$$

und

$$\frac{\sum_{i=1}^{\mu+\nu} \beta_i}{\mu + \nu} \geqq \frac{\pi}{3}. \tag{5.5}$$

Dann gilt

$$\sum_{i=1}^{k} e_\varrho(\alpha_i, \beta_i) \leqq kF\left(\frac{\sum_{i=1}^{k} \alpha_i}{k}\right). \tag{5.6}$$

Dabei gilt das Gleichheitszeichen nur für $\mu = \nu = 0$. Im Fall $\mu = \nu = 0$ gilt das Gleichheitszeichen genau dann, wenn

$$\begin{aligned} &\beta_i = \frac{\pi}{3} \quad \text{für alle } i, \text{ und} \\ &\alpha_1 = \alpha_2 = \dots = \alpha_k, \text{ und } (\alpha_i, \varrho) \text{ ist ein Punkt der Kurve (4.4).} \end{aligned} \tag{5.7}$$

Satz 2 wird in **6**, ... , **12** bewiesen, d.h. im restlichen Teil von I. Zunächst wird $e_\varrho(\alpha, \beta)$ in den einzelnen Teilen seines Definitionsbereiches genauer untersucht und es werden *konkave* Hilfsfunktionen gefunden, die $e_\varrho(\alpha, \beta)$ nach oben abschätzen. Damit läßt sich schließlich in **11** und **12** Satz 2 bewiesen.

6. Die Funktion $e_\varrho(\alpha, \beta)$ in G_3

HILFSSATZ 6.1. *Für $0 < \varrho < \frac{\pi}{2}$ gilt: In G_3 ist $e_\varrho(\alpha, \beta)$ streng monoton wachsend in α und in β, und als Funktion von β konvex.*

BEWEIS. In G_3 ist $e_\varrho(\alpha, \beta)$ streng monoton wachsend in α und in β aus geometrischen Gründen.

Um zu zeigen, daß $e_\varrho(\alpha, \beta)$ in G_3 in β konvex ist, wird G_3 zweckmäßigerweise mit Hilfe der Parameter β, r und d beschrieben (siehe (2.1) und (3.6)). Wegen (2.1) ist (3.1) äquivalent zu $\sin r = \sin \beta \sin \varrho$. G_3 wird also beschrieben durch (vgl. Fig. 1)

$$\left.\begin{array}{lll}
(6.1) & & \sin r \geqq \sin\beta\sin\varrho \\
(6.2) & & \frac{\varrho}{2} \leqq \beta \leqq \frac{\pi}{4} \\
(6.3) & & r > \frac{\varrho}{2} \\
(6.4) & & d \leqq \varrho, \text{ wobei wegen } (3.6) \\
(6.5) & & \sin d = 2\cos\beta\sin r.
\end{array}\right\} G_3$$

In G_3 berechnet sich $e_\varrho(\alpha,\beta)$ nach (3.12), d.h. $e_\varrho(\alpha,\beta) = e_\varrho(\alpha, \sin r(\alpha,\beta), \sin d(\alpha,\beta))$. Bezeichnet man mit $\sin r_\beta$, $\sin r_{\beta\beta}$ bzw. $\sin d_\beta$, $\sin d_{\beta\beta}$ die 1. und 2. partielle Ableitung von $\sin r$ bzw. $\sin d$ nach β, so ist die 2. Ableitung von $e_\varrho(\alpha,\beta)$ nach β gleich

$$\begin{aligned}
&\frac{1}{\cos^4 r\sqrt{\sin^2\varrho-\sin^2 r}}[\cos^2 r(\sin^2\varrho-\sin^2 r)\sin r_{\beta\beta}-\\
&\quad-(\sin^2 r+\cos^2\varrho-\sin^2\varrho)\sin r\,(\sin r_\beta)^2]-\\
&-\frac{1}{\cos^4 d\sqrt{\sin^2\varrho-\sin^2 d}}[\cos^2 d\,(\sin^2\varrho-\sin^2 d)\sin d_{\beta\beta}-\\
&\quad-(\sin^2 d+\cos^2\varrho-\sin^2\varrho)\sin d(\sin d_\beta)^2].
\end{aligned}$$

Setzt man hier explizit die Funktionen $\sin r_\beta$, $\sin r_{\beta\beta}$, $\sin d_\beta$, $\sin d_{\beta\beta}$ als Funktionen von α und β ein, schreibt mit Hilfe von (2.1) alles auf die Parameter r und β um, und kürzt gemeinsame Faktoren der beiden Summanden heraus, so ist also zum Beweis von Hilfssatz 6.1 zu zeigen

$$\begin{aligned}
&\frac{\sin r}{\cos^4 r\sqrt{\sin^2\varrho-\sin^2 r}}[\cos^2 r(\sin^2\varrho-\sin^2 r)(2\cos^2\beta-\cos^2 r\cos^2\beta-1)-\\
&\quad-(\sin^2 r+\cos^2\varrho-\sin^2\varrho)\sin^2 r\cos^2 r\sin^2\beta]-\\
&-\frac{\sin d}{\cos^4 d\sqrt{\sin^2\varrho-\sin^2 d}}\Big[\cos^2 d(\sin^2\varrho-\sin^2 d)(6\cos^2\beta-3-\frac{\cos^2\beta}{\cos^2 r}+\\
&\quad+2\cos^2 r-4\cos^2\beta\cos^2 r)-\\
&-(\sin^2 d+\cos^2\varrho-\sin^2\varrho)\sin^2 d\sin^2\beta\frac{(\cos^2 r-\sin^2 r)^2}{\cos^2 r}\Big]\geqq 0.
\end{aligned}\tag{6.6}$$

Nun ist

$$2\cos^2\beta-\cos^2 r\cos^2\beta-1+\sin^2 r\sin^2\beta=\sin^2 r-\sin^2\beta\geqq-\cos^2\varrho\sin^2\beta$$

nach (6.1).

Wegen (6.5), (6.1) und (6.2) gilt

$$\sin^2 d = 4\cos^2\beta \sin^2 r \geqq 4\cos^2\beta \sin^2\beta \sin^2\varrho \geqq \sin^4\varrho,$$

so daß

$$\sin^2 d + \cos^2\varrho - \sin^2\varrho \geqq \cos^4\varrho \geqq 0$$

gilt. Deshalb kann in (6.6) $\cos^2 r - \sin^2 r$ nach unten abgeschätzt werden. In G_3 gilt wegen (6.5) und (6.2) $d \geqq r$. Deshalb ist $\cos^2 r - \sin^2 r \geqq \cos^2 d$. Außerdem ist $\frac{\cos^2 d}{\cos^2 r} \leqq 1$, so daß weiter gilt

$$\begin{aligned} 6\cos^2\beta - 3 - \frac{\cos^2\beta}{\cos^2 r} + 2\cos^2 r - 4\cos^2\beta\cos^2 r + \sin^2 d \sin^2\beta \frac{\cos^2 d}{\cos^2 r} \leqq \\ \leqq \frac{\sin^2\beta - \sin^2 r}{\cos^2 r} - 2(\sin^2\beta - \sin^2 r) - 4\sin^2 r \sin^4\beta. \end{aligned}$$

Dies ist $\leqq 0$, denn wegen (2.1) ist $\sin^2\beta - \sin^2 r \geqq 0$, und wegen (2.1) und (6.2) ist $\frac{1}{\cos^2 r} - 2 \leqq 0$. Deshalb genügt es zu zeigen, daß

$$\begin{aligned} \frac{\sin r}{\cos^4 r\sqrt{\sin^2\varrho - \sin^2 r}}[-\cos^2 r \sin^2\varrho \cos^2\varrho \sin^2\beta] - \\ - \frac{\sin d}{\cos^4 d\sqrt{\sin^2\varrho - \sin^2 d}}\left[-\cos^2\varrho \sin^2 d \sin^2\beta \frac{\cos^4 d}{\cos^2 r}\right] \geqq 0, \end{aligned}$$

oder wegen (6.5), daß

$$\frac{8\cos^3\beta \sin^2 r}{\sqrt{\sin^2\varrho - 4\cos^2\beta\sin^2 r}} \geqq \frac{\sin^2\varrho}{\sqrt{\sin^2\varrho - \sin^2 r}}. \tag{6.7}$$

Die linke Seite von (6.7) wird für jedes feste r minimal für maximales β gemäß (6.1), so daß zu zeigen bleibt (man beachte, daß $\sin^2\varrho - 2\sin^2 r \geqq 0$ nach (6.4) und (6.2))

$$8\sin^6 r - 16\sin^2\varrho\sin^4 r + 10\sin^4\varrho\sin^2 r - \sin^6\varrho \geqq 0. \tag{6.8}$$

Durch Ableiten nach $\sin^2 r$ sieht man, daß die linke Seite von (6.8) genau dann wachsend in r ist, wenn $(\sin^2\varrho - 2\sin^2 r)(5\sin^2\varrho - 6\sin^2 r) \geqq 0$, was nach oben zutrifft. Außerdem ist (6.8) für $r = \frac{\varrho}{2}$ richtig, womit Hilfssatz 6.1 gezeigt ist. □

Bemerkung. In G_3 ist $e_\varrho(\alpha, \beta)$ als Funktion von α nicht konvex.

7. Die Funktion $e_\varrho(\alpha,\beta)$ in $G_1 \cup G_2$ und die konkave Hilfsfunktion $h_\varrho(\alpha,\beta)$

Die Funktion $e_\varrho(\alpha,\beta)$ in $G_1 \cup G_2$ wird zu einer im ganzen Intervall $0 \leqq \alpha \leqq \frac{\pi}{2}$, $0 \leqq \beta \leqq \frac{\pi}{2}$ definierten konkaven Funktion $h_\varrho(\alpha,\beta)$ fortgesetzt:

HILFSSATZ 7.1. *Zu ϱ mit $0 < \varrho < \frac{\pi}{2}$ sei die Funktion $h_\varrho(\alpha,\beta)$ in $0 \leqq \alpha \leqq \frac{\pi}{2}$, $0 \leqq \beta \leqq \frac{\pi}{2}$ definiert durch*

$$h_\varrho(\alpha,\beta) := \begin{cases} \alpha(1-\cos\varrho) \quad \text{in } G_1 & (7.1) \\ \alpha(1-\cos\varrho)-\pi+2\cos\varrho\arccos\frac{\tan r}{\tan\varrho}+2\arcsin\frac{\sin r}{\sin\varrho} & \\ \quad \text{in } G_2\cup G_3\cup G_4 \quad \text{mit } \cos r = \frac{\cos\beta}{\sin\alpha} & (7.2) \\ \alpha(1+\cos\varrho)-\pi+2\beta \quad \text{sonst.} & (7.3) \end{cases}$$

Dann ist $h_\varrho(\alpha,\beta)$ wohldefiniert und in α und in β nicht fallend, konkav im ganzen Definitionsbereich, streng konkav in $(G_2\cup G_3\cup G_4)^0$, und $h_\varrho(\alpha,\beta) = e_\varrho(\alpha,\beta)$ in $G_1 \cup G_2$.

BEWEIS. Zunächst rechnet man nach, daß $h_\varrho(\alpha,\beta)$ längs (3.1) eindeutig definiert ist, d.h. $h_\varrho(\alpha,\beta)$ ist wohldefiniert.

Die Ableitungen von (7.2) sind

$$\begin{aligned} (h_\varrho)_\alpha &= 1-\cos\varrho+2\frac{\cos\alpha}{\sin\alpha}\frac{V}{W}, \qquad (h_\varrho)_\beta = 2\frac{\sin\beta}{\cos\beta}\frac{V}{W}, \\ (h_\varrho)_{\alpha\alpha} &= -2\frac{V^2W^2+\sin^2\alpha\cos^2\alpha\cos^2\beta\sin^2\varrho}{\sin^2\alpha\, VW^3} \\ (h_\varrho)_{\alpha\beta} &= -2\frac{\sin\alpha\cos\alpha\sin\beta\cos\beta\sin^2\varrho}{VW^3} \\ (h_\varrho)_{\beta\beta} &= 2\frac{V^2W^2-\sin^2\alpha\sin^2\beta\cos^2\beta\sin^2\varrho}{\cos^2\beta\, VW^3}. \end{aligned} \tag{7.4}$$

mit $V := \sqrt{\cos^2\beta-\sin^2\alpha\cos^2\varrho}$, $W := \sqrt{\sin^2\alpha-\cos^2\beta}$. Deshalb ist $h_\varrho(\alpha,\beta)$ in α und in β nicht fallend. In $(G_2\cup G_3\cup G_4)^0$ ist $(h_\varrho)_{\alpha\alpha} < 0$ und $(h_\varrho)_{\alpha\alpha}(h_\varrho)_{\beta\beta} - (h_\varrho)^2_{\alpha\beta} > 0$, so daß h_ϱ dort streng konkav ist. Weiter rechnet man nach, daß das Ebenenstück (7.1) Tangentialebene an (7.2) längs (3.5) ist, und daß das Ebenenstück (7.3) Tangentialebene an (7.2) längs (3.1) ist, so daß $h_\varrho(\alpha,\beta)$ im ganzen Definitionsbereich konkav ist. □

Aus Hilfssatz 7.1 folgt sofort

HILFSSATZ 7.2. *Jeder Punkt der Fläche* $z = h_\varrho(\alpha, \beta)$ *liegt unterhalb der Ebenen*

$$z = \alpha(1 - \cos\varrho) \quad \textit{und} \tag{7.5}$$

$$z = \alpha(1 + \cos\varrho) - \pi + 2\beta \tag{7.6}$$

mit der Schnittgeraden

$$(\alpha, \beta, z) = \left(\left(\frac{\pi}{2} - \beta\right)\frac{1}{\cos\varrho}, \beta, \left(\frac{\pi}{2} - \beta\right)\left(\frac{1}{\cos\varrho} - 1\right)\right) \tag{7.7}$$

Schnittgerade der Ebenen (7.5) *und* (7.6), *auf der Ebene* $z = \alpha + \beta - \frac{\pi}{2}$ *gelegen.* □

Hilfssatz 4.1 kann erweitert werden zu

HILFSSATZ 7.3. *Zu jedem* ϱ *mit* $0 < \varrho < \frac{\pi}{2}$ *und zu jedem* α *mit* $\frac{\pi}{6} \leqq \alpha \leqq \frac{\pi}{2}$ *ist*

$$h_\varrho\left(\alpha, \frac{\pi}{3}\right) \leqq F(\alpha) = \alpha - \pi + 2 \arccos \frac{1}{4 \cos\frac{\alpha}{2}}. \tag{7.8}$$

Das Gleichheitszeichen gilt genau dann, wenn (α, ϱ) *ein Punkt der Kurve* (4.4) *ist, siehe Fig.* 6.

BEWEIS. Analog zum Beweis von Hilfssatz 4.1; man beachte nur, daß jetzt $h_\varrho\left(\alpha, \frac{\pi}{3}\right)$ im Bereich $D_1 \cup D_2 \cup D_3$ zu untersuchen ist, und daß $h_\varrho\left(\alpha, \frac{\pi}{3}\right)$ in D_3 streng monoton fallend in ϱ ist. □

8. Der konvexe Bereich $\overline{G}$ in der (α, β)-Ebene

In der (α, β)-Ebene sei $\overline{G} = \overline{G}(\varrho)$ die konvexe Hülle

$$\overline{G} := \operatorname{conv}\left\{\left(\frac{\pi}{2}, \frac{\pi}{2}\right), \left(0, \frac{\pi}{2}\right), \left(\frac{\pi}{6}, \frac{\pi}{3}\right), \left(\frac{\pi}{2}, \frac{\pi}{4}(1 - \cos\varrho)\right)\right\},$$

siehe Fig. 2. Dabei ist der letzte Punkt $(\frac{\pi}{2}, \frac{\pi}{4}(1 - \cos\varrho))$ ein Schnittpunkt der Ebene (7.6) mit der (α, β)-Ebene.

HILFSSATZ 8.1. *Für den konvexen Bereich* $\overline{G}$ *gilt*

$$\overline{G} \supset G_1 \cup G_2 \cup G_3 \quad \textit{und} \tag{8.1}$$

(8.2) *Der Punkt* (3.3) *ist innerer Punkt von* $\overline{G}$.

Beweis. Man betrachte die Gerade durch $(\frac{\pi}{6},\frac{\pi}{3})$ und $(\frac{\pi}{2},\frac{\pi}{4}(1-\cos\varrho))$ mit der Gleichung

$$\beta = -\frac{\alpha}{4}(1+3\cos\varrho)+\frac{\pi}{8}(3+\cos\varrho). \tag{8.3}$$

Die Geraden (8.3) und $\alpha = 2\beta$ schneiden sich im Punkt

$$\Big(2\Big(\frac{\pi}{12}+\frac{\pi}{6}\frac{1}{1+\cos\varrho}\Big),\frac{\pi}{12}+\frac{\pi}{6}\frac{1}{1+\cos\varrho}\Big).$$

Deshalb ist der Punkt (3.3) genau dann innerer Punkt von $\overline{G}$, wenn

$$\frac{\pi}{12}+\frac{\pi}{6}\frac{1}{1+\cos\varrho}-\arcsin\frac{1}{2\cos\frac{\varrho}{2}}<0, \qquad 0<\varrho<\frac{\pi}{2}. \tag{8.4}$$

Für $\varrho = 0$ und $\varrho = \frac{\pi}{2}$ tritt in (8.4) das Gleichheitszeichen auf. Für die Ableitung $A(\varrho)$ der linken Seite von (8.4) gilt $A(\frac{\pi}{2})>0$, und $A(\varrho)=0$ für $\varrho=0$ und für genau ein weiteres $\varrho<\frac{\pi}{2}$. Deshalb ist (8.4) richtig und damit (8.2).

Wegen der Konkavität von (3.1) liegt also das $G_1\cup G_2\cup G_3$ begrenzende Stück von (3.1) in $\overline{G}$. Das G_3 begrenzende Stück von (3.2) liegt ebenfalls in $\overline{G}$; denn die Steigung von (3.2) im Punkt (3.3) ist $-\cos\varrho$, die Steigung der Geraden (8.3) ist $-\frac{1}{4}(1+3\cos\varrho)<-\cos\varrho$, und (3.2) ist konvex. Damit gilt (8.1). □

9. Die konkave Hilfsfunktion $H_\varrho(\alpha,\beta)$

Außer $h_\varrho(\alpha,\beta)$ wird eine weitere konkave Hilfsfunktion $H_\varrho(\alpha,\beta)$ benötigt. Dabei genügt es, entsprechend der Voraussetzung (5.1) von Satz 2 vorauszusetzen, daß $0<\varrho<\varrho_0$.

Hilfssatz 9.1. *Zu jedem ϱ mit $0<\varrho<\varrho_0=\arccos\frac{1}{\sqrt{7}}$ gibt es eine für $0\leqq\alpha\leqq\frac{\pi}{2}$, $0\leqq\beta\leqq\frac{\pi}{2}$ definierte Funktion $H_\varrho(\alpha,\beta)$ mit*

(9.1) $\quad H_\varrho(\alpha,\beta)$ *ist konkav und in α und β nicht fallend,*

$$H_\varrho\Big(\alpha,\frac{\pi}{3}\Big)=F(\alpha) \quad \textit{für } \frac{\pi}{6}\leqq\alpha\leqq\frac{\pi}{2}, \tag{9.2}$$

(9.3) $\quad H_\varrho(\alpha,\beta)\geqq h_\varrho(\alpha,\beta)$, *und das Gleichheitszeichen gilt genau dann, wenn $\beta=\frac{\pi}{3}$ und (α,ϱ) ein Punkt der Kurve (4.4) ist.*

(9.4) $$H_\varrho(\alpha,\beta) > e_\varrho(\alpha,\beta) \quad in\ G_3.$$

BEMERKUNG. Aus Hilfssatz 9.1 folgt sofort, daß Satz 2 im Fall $\mu = \nu = 0$ richtig ist.

BEWEIS von Hilfssatz 9.1. Als erstes wird $H_\varrho(\alpha,\beta)$ definiert. Bei festem ϱ gilt, mit α_1 nach (4.1) und α_2 nach (4.2), siehe Fig. 1 und Fig. 6

$$G_2 \cap \left\{\left(\alpha, \frac{\pi}{3}\right)\right\} = \begin{cases} \{(\alpha, \frac{\pi}{3}) \mid \alpha_1 \leqq \alpha \leqq \alpha_2\} & \text{für } \varrho \leqq \frac{\pi}{3}, \\ \{(\alpha, \frac{\pi}{3}) \mid \alpha_1 \leqq \alpha \leqq \frac{\pi}{2} & \text{für } \frac{\pi}{3} \leqq \varrho < \varrho_0. \end{cases}$$

Dort ist h_ϱ streng konkav nach Hilfssatz 7.1 und $h_\varrho(\alpha, \frac{\pi}{3}) \leqq F(\alpha)$ nach Hilfssatz 7.3. Nach (4.5) ist $F(\alpha)$ streng konkav für $\frac{\pi}{6} \leqq \alpha \leqq \frac{\pi}{2}$. Nun rechnet man mit Hilfe von (7.4) nach, daß für die Ableitungen von $F(\alpha)$ und $h_\varrho(\alpha, \frac{\pi}{3})$ nach α gilt

$$F_\alpha\left(\frac{\pi}{6}\right) < (h_\varrho)_\alpha\left(\alpha_1, \frac{\pi}{3}\right)$$

und

$$F_\alpha\left(\frac{\pi}{2}\right) > \begin{cases} (h_\varrho)_\alpha(\alpha_2, \frac{\pi}{3}) & \text{für } \varrho \leqq \frac{\pi}{3} \\ (h_\varrho)_\alpha(\frac{\pi}{2}, \frac{\pi}{3}) & \text{für } \frac{\pi}{3} \leqq \varrho < \varrho_0; \end{cases}$$

dabei ist die Voraussetzung $\varrho < \varrho_0$ für die Gültigkeit der letzten Ungleichung wesentlich. Zu jedem α_0 mit $\frac{\pi}{6} \leqq \alpha_0 \leqq \frac{\pi}{2}$ gibt es also genau ein $\overline{\alpha}_0 = \overline{\alpha}_0(\varrho)$ mit $\alpha_1 < \overline{\alpha}_0 < \alpha_2$ bzw. $\frac{\pi}{2}$ so, daß $F_\alpha(\alpha_0) = (h_\varrho)_\alpha(\overline{\alpha}_0, \frac{\pi}{3})$, siehe Fig. 7. Die Tangentialebene der Fläche $z = h_\varrho(\alpha,\beta)$ im Punkt $(\overline{\alpha}_0, \frac{\pi}{3})$ werde nun parallel verschoben so, daß sie den Punkt $(\alpha,\beta,z) = (\alpha_0, \frac{\pi}{3}, F(\alpha_0))$ enthält. Jede solche verschobene Tangentialebene begrenzt einen Halbraum, der den Punkt $(\alpha,\beta,z) = (\frac{\pi}{2}, \frac{\pi}{3}, 0)$ im Innern enthält, und der Durchschnitt aller dieser Halbräume für $\frac{\pi}{6} \leqq \alpha_0 \leqq \frac{\pi}{2}$ ist eine konvexe Menge, die von einer konkaven Funktion $z = H_\varrho(\alpha,\beta)$ begrenzt wird.

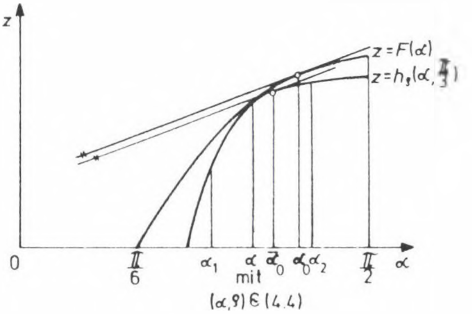


Fig. 7

$H_\varrho(\alpha,\beta)$ ist also konkav. Weil alle betrachteten Tangentialebenen in α und β nicht fallend sind, gilt (9.1). (9.2) gilt, weil $H_\varrho(\alpha, \frac{\pi}{3})$ die Hüllkurve der Tangenten der konkaven Funktion $F(\alpha)$ ist. Wegen der Konstruktion von H_ϱ mit Hilfe von Tangentialebenen an die konkave Funktion h_ϱ gilt $H_\varrho(\alpha,\beta) \geqq h_\varrho(\alpha,\beta)$. Die Tangentialebene an $z = h_\varrho(\alpha,\beta)$ im Punkt $(\overline{\alpha}_0, \frac{\pi}{3})$ wird genau dann nicht verschoben, wenn $(\overline{\alpha}, \frac{\pi}{3})$ ein Punkt der Kurve (4.4) ist, nach Hilfssatz 7.3. Weil $(\overline{\alpha}_0, \frac{\pi}{3}) \in G_2^0$ und $h_\varrho(\alpha,\beta)$ in G_2^0 streng konkav und sonst konkav ist, gilt also (9.3). Zu zeigen bleibt damit (9.4), was leider aufwendig ist.

Zum Beweis von (9.4) wird zunächst gezeigt

$$(9.5)\qquad H_\varrho(\alpha,\beta) \geqq 0 \quad \text{in } \overline{G}, \overline{G} \text{ gemäß 8.}$$

Wegen $H_\varrho(\alpha,\beta) \geqq h_\varrho(\alpha,\beta)$ in $G_1 \cup G_2$ und $H_\varrho(\frac{\pi}{6}, \frac{\pi}{3}) = 0$ nach (9.2) und nach Definition von $\overline{G}$ genügt es zu zeigen, daß $H_\varrho(\frac{\pi}{2}, \frac{\pi}{4}(1-\cos\varrho)) \geqq 0$. Dafür ist hinreichend, daß für die Schar der Tangentialebenen an $h_\varrho(\alpha,\beta)$ für $(\alpha,\beta) \in G_2 \cap \{(\alpha, \frac{\pi}{3})\}$ die entsprechende Ungleichung gilt, für jedes ϱ, d.h.

$$(9.6)\quad h_\varrho\left(\alpha, \frac{\pi}{3}\right) + \left(\frac{\pi}{2} - \alpha\right)(h_\varrho)_\alpha\left(\alpha, \frac{\pi}{3}\right) + \left(\frac{\pi}{4}(1-\cos\varrho) - \frac{\pi}{3}\right)(h_\varrho)_\beta\left(\alpha, \frac{\pi}{3}\right) \geqq 0$$

für $(\alpha,\varrho) \in D_2$, $\varrho < \varrho_0$.

Für jedes feste ϱ ist die Tangentialebene von h_ϱ im Punkt $(\alpha_1, \frac{\pi}{3})$ die Ebene (7.6). Sie hat im Punkt $(\frac{\pi}{2}, \frac{\pi}{4}(1-\cos\varrho))$ den Funktionswert $z = 0$. Deshalb gilt (9.6) für $\alpha = \alpha_1(\varrho)$. Also genügt es zu zeigen, daß (9.6) wachsend in α ist, d.h. daß

$$(9.7)\quad \begin{aligned} &-\left(\frac{\pi}{2} - \alpha\right)(4\sin^2\alpha - 1 + 4\sin^2\alpha\cos^2\alpha - 12\sin^4\alpha\cos^2\varrho) + \\ &+\left(\frac{\pi}{3} + \pi\cos\varrho\right)\sqrt{3}\sin^3\alpha\cos\alpha\sin^2\varrho \geqq 0 \quad \text{für } (\alpha,\varrho) \in D_2,\ \varrho < \varrho_0. \end{aligned}$$

D_2' sei der Bereich von D_2 mit $0 < \varrho \leqq \arccos\frac{1}{3}$. Dann ist $(\alpha,\varrho) = (\frac{\pi}{3}, \arccos\frac{1}{3})$ Randpunkt von D_2'. (9.7) wird in D_2' gezeigt.

Längs $\alpha = \alpha_1(\varrho)$, d.h. $\varrho = \varrho(\alpha)$ wird aus (9.7) die Ungleichung

$$-\left(\frac{\pi}{2} - \alpha\right) + \frac{\pi}{3\sqrt{3}}\sin\alpha\cos\alpha + \frac{\pi}{3}\cos^2\alpha \geqq 0 \quad \text{für } \frac{\pi}{6} < \alpha \leqq \frac{\pi}{3},$$

und diese ist richtig, weil sie für $\alpha = \frac{\pi}{6}$ und $\alpha = \frac{\pi}{3}$ richtig ist, und weil ihre linke Seite konkav ist. Längs $\varrho = \arccos\frac{1}{3}$ wird aus (9.7) die Ungleichung

$$\begin{aligned} &-\left(\frac{\pi}{2} - \alpha\right)(12\sin^2\alpha - 3 + 12\sin^2\alpha\cos^2\alpha - 4\sin^4\alpha) + \\ &+\pi\frac{2}{3\sqrt{3}}8\sin^3\alpha\cos\alpha \geqq 0 \quad \text{für } \frac{\pi}{3} \leqq \alpha \leqq \frac{\pi}{2}; \end{aligned}$$

ihre Gültigkeit sieht man mit Hilfe der Abschätzung

$$\frac{\pi}{2} - \alpha \leqq \pi \frac{2}{3\sqrt{3}} \sin\alpha\cos\alpha \quad \text{für } \frac{\pi}{3} \leqq \alpha \leqq \frac{\pi}{2}.$$

Damit ist (9.7) längs des oberen Rands von D_2' richtig, so daß (9.7) gilt, wenn ihre linke Seite in ϱ fallend ist, oder wenn

$$-\left[24\left(\frac{\pi}{2}-\alpha\right)\sin\alpha - \frac{2\pi}{\sqrt{3}}\cos\alpha\right]\cos\varrho + 3\sqrt{3}\pi\cos\alpha\cos^2\varrho - \sqrt{3}\pi\cos\alpha \leqq 0 \text{ in } D_2'.$$

Dies ist richtig wegen der Abschätzung

$$24\left(\frac{\pi}{2} - \alpha\right)\sin\alpha \geqq \frac{8\pi}{\sqrt{3}}\cos\alpha \quad \text{für } \frac{\pi}{6} \leqq \alpha \leqq \frac{\pi}{2}.$$

Damit ist (9.5) gezeigt.

Zum Beweis von (9.4) setze man $e_\varrho(\alpha,\beta)$ in G_3 gemäß (3.12) auf das Randstück (3.2) von G_3 fort. Wegen (9.3) gilt (9.4) längs (3.9), $e_\varrho(\alpha,\beta)$ in G_3 ist in β konvex nach Hilfssatz 6.1, und H_ϱ ist konkav. Deshalb genügt es, $H_\varrho(\alpha,\beta) > e_\varrho(\alpha,\beta)$ längs des Randes von G_3 auf (3.2) und (3.1) zu zeigen, d.h.

(9.8)
$$H_\varrho(\alpha,\beta) > e_\varrho(\alpha,\beta) \text{ längs } \sin\alpha = \frac{\cos\beta}{\cos\frac{\varrho}{2}} \text{ für } \frac{\varrho}{2} \leqq \beta \leqq \beta_0 := \arcsin\frac{1}{2\cos\frac{\varrho}{2}},$$

und

$$(9.9) \qquad H_\varrho(\alpha,\beta) > e_\varrho(\alpha,\beta) \text{ längs } \tan\alpha = \frac{1}{\cos\varrho\tan\beta} \text{ für } \beta_0 \leqq \beta \leqq \frac{\pi}{4}.$$

Der Rand von $\overline{G}$ enthält die Verbindungsstrecke von $(\frac{\pi}{6}, \frac{\pi}{3})$ und $(\frac{\pi}{2}, \frac{\pi}{4}(1-\cos\varrho))$

$$(9.10) \qquad \alpha = \frac{\pi}{2}\frac{3+\cos\varrho}{1+3\cos\varrho} - \frac{4}{1+3\cos\varrho}\beta \text{ mit } \frac{\pi}{4}(1-\cos\varrho) \leqq \beta \leqq \frac{\pi}{3}.$$

Die konvexe Hülle von (9.10) in der (α,β)-Ebene und dem Punkt $(\alpha,\beta,z) = (\frac{\pi}{2}, \frac{\pi}{3}, F(\frac{\pi}{2}))$ ist ein Ebenenstück. Wegen (9.5), (9.2) und (9.1) liegt $H_\varrho(\alpha,\beta)$ oberhalb dieses Ebenenstücks. (9.8) und (9.9) ist also richtig, wenn dieses Ebenenstück echt oberhalb von $z = e_\varrho(\alpha,\beta)$ liegt, längs der zu betrachtenden Kurvenstücke von (3.2) und (3.1). Dieses Ebenenstück hat in α-Richtung die Steigung

$$C := \frac{3}{\pi}F\left(\frac{\pi}{2}\right) = 0.8098\ldots,$$

so daß (9.8) und (9.9) richtig ist, wenn gilt

$$(9.11)\qquad \begin{gathered} C\left[\alpha-\left(\frac{\pi}{2}\frac{3+\cos\varrho}{1+3\cos\varrho}-\frac{4}{1+3\cos\varrho}\beta\right)\right]>e_\varrho(\alpha,\beta) \\ \text{längs }\ \sin\alpha=\frac{\cos\beta}{\cos\frac{\varrho}{2}},\quad \frac{\varrho}{2}\leqq\beta\leqq\beta_0 \end{gathered}$$

$$(9.12)\qquad \begin{gathered} C\left[\alpha-\left(\frac{\pi}{2}\frac{3+\cos\varrho}{1+3\cos\varrho}-\frac{4}{1+3\cos\varrho}\beta\right)\right]>e_\varrho(\alpha,\beta) \\ \text{längs }\ \tan\alpha=\frac{1}{\cos\varrho\tan\beta},\quad \beta_0\leqq\beta\leqq\frac{\pi}{4}. \end{gathered}$$

Nach (8.2) ist der Punkt (3.3) innerer Punkt von $\overline{G}$ und dort ist aus geometrischen Gründen $e_\varrho(\alpha,\beta)=0$, so daß (9.11) und (9.12) richtig ist für $\beta=\beta_0$.

Um nun (9.11) zu zeigen, beachte man, daß (9.11) richtig ist, wenn

$$C\left[\arcsin\frac{\cos\beta}{\cos\frac{\varrho}{2}}-\left(\frac{\pi}{2}\frac{3+\cos\varrho}{1+3\cos\varrho}-\frac{4}{1+3\cos\varrho}\beta\right)\right]-e_\varrho\left(\arcsin\frac{\cos\beta}{\cos\frac{\varrho}{2}},\beta\right)$$

in β fallend ist, d.h. wenn

$$(9.13)\qquad \begin{gathered} -\left(1-\frac{\sqrt{\cos^2\frac{\varrho}{2}-\cos^2\beta}}{\sin\beta}\frac{4}{1+3\cos\varrho}\right)C-(1+\cos\varrho)+ \\ +2\cos^2\varrho\frac{1}{1-4\cos^2\beta\sin^2\frac{\varrho}{2}}\leqq 0,\qquad \frac{\varrho}{2}\leqq\beta\leqq\beta_0. \end{gathered}$$

Mit der Hilfsfunktion

$$\begin{gathered} f(\varrho,\beta):=-\left(1-\frac{\sqrt{\cos^2\frac{\varrho}{2}-\cos^2\beta}}{\sin\beta}\frac{1}{\cos\varrho}\right)-(1+\cos\varrho)+ \\ +2\cos^2\varrho\frac{1}{1-4\cos^2\beta\sin^2\frac{\varrho}{2}} \end{gathered}$$

ist (9.13) äquivalent zu

(9.14)

$$f(\varrho,\beta)-C+1-\frac{\sqrt{\cos^2\frac{\varrho}{2}-\cos^2\beta}}{\sin\beta}\left(\frac{1}{\cos\varrho}-\frac{4C}{1+3\cos\varrho}\right)\leqq 0,\quad \text{für }\frac{\varrho}{2}\leqq\beta\leqq\beta_0.$$

Nun ist $f(\varrho,\beta)$ für $\frac{\varrho}{2}\leqq\beta\leqq\beta_0$ wachsend in β, denn dies ist äquivalent zu

$$(9.15)\qquad 1-\frac{16\cos^3\varrho\sin^3\beta\sqrt{\cos^2\frac{\varrho}{2}-\cos^2\beta}}{(1-4\cos^2\beta\sin^2\frac{\varrho}{2})^2}\geqq 0,$$

und die linke Seite davon ist fallend in β [2] und gleich 0 für $\beta = \beta_0$. Die Funktion $-\frac{\sqrt{\cos^2\frac{\varrho}{2} - \cos^2\beta}}{\sin\beta}$ ist für $\frac{\varrho}{2} \leqq \beta \leqq \beta_0$ fallend in β. Speziell für $\beta = \beta_0$ ist $f(\varrho, \beta) = 0$ und gilt in (9.14) das <-Zeichen.

Deshalb ist (9.14) richtig für $\beta_1 \leqq \beta \leqq \beta_0$ mit

$$-C + 1 - \frac{\sqrt{\cos^2\frac{\varrho}{2} - \cos^2\beta_1}}{\sin\beta_1}\left(\frac{1}{\cos\varrho} - \frac{4C}{1 + 3\cos\varrho}\right) = 0,$$

oder

$$\cos^2\beta_1 = \frac{\cos^2\frac{\varrho}{2} - \cos^2\varrho\, C_1^2}{1 - \cos^2\varrho\, C_1^2} \quad \text{mit} \quad C_1(\varrho) := \frac{1 - C}{1 - \frac{4\cos\varrho}{1+3\cos\varrho} C}.$$

Ebenso ist (9.14) richtig für $\beta_2 \leqq \beta \leqq \beta_1$ mit

$$f(\varrho, \beta_1) - C + 1 - \frac{\sqrt{\cos^2\frac{\varrho}{2} - \cos^2\beta_2}}{\sin\beta_2}\left(\frac{1}{\cos\varrho} - \frac{4C}{1 + 3\cos\varrho}\right) = 0,$$

oder

$$\cos^2\beta_2 = \frac{\cos^2\frac{\varrho}{2} - \cos^2\varrho\, C_2^2}{1 - \cos^2\varrho\, C_2^2} \quad \text{mit} \quad C_2(\varrho) := \frac{1 - C + f(\varrho, \beta_1)}{1 - \frac{4\cos\varrho}{1+3\cos\varrho} C}.$$

(9.14) und somit (9.13) ist also richtig für $\beta_2 \leqq \beta \leqq \beta_0$.

Zum Beweis von (9.13) bleibt also zu zeigen, daß die linke Seite von (9.13) wachsend ist in β für $\frac{\varrho}{2} \leqq \beta \leqq \beta_2$, d.h.

$$C - \frac{1 + 3\cos\varrho}{4\cos\varrho}\,\frac{16\cos^3\varrho\sin^3\beta\sqrt{\cos^2\frac{\varrho}{2} - \cos^2\beta}}{(1 - 4\cos^2\beta\sin^2\frac{\varrho}{2})^2} \geqq 0 \quad \text{für } \frac{\varrho}{2} \leqq \beta \leqq \beta_2.$$

Weil die linke Seite hiervon fallend in β ist (vgl. (9.15)), bleibt diese Ungleichung für $\beta = \beta_2$ zu zeigen, d.h.

$$C - \frac{1 + 3\cos\varrho}{4\cos\varrho}\,\frac{C_2}{\left(\frac{1 - C_2^2}{4\sin^2\frac{\varrho}{2}} + C_2^2\right)^2} \geqq 0 \quad \text{für } 0 < \varrho < \varrho_0.$$

Die Gültigkeit dieser Ungleichung mit der einzigen Variablen ϱ rechnet man

[2] Man unterscheidet dabei zweckmäßigerweise die Fälle $4\sin^2\frac{\varrho}{2} \leqq 1$ und $4\sin^2\frac{\varrho}{2} \geqq 1$.

schließlich numerisch nach[3], wobei für kleine ϱ wesentlich ist, daß

$$\frac{1-C_2^2}{4\sin^2\frac{\varrho}{2}} =$$

$$= \left[\frac{C}{2(1+3\cos\varrho-4\cos\varrho\, C)} - \frac{f(\varrho,\beta_1)(1+3\cos\varrho)}{2(1+3\cos\varrho-4\cos\varrho\, C)(1-\cos\varrho)}\right](1+C_2) \geqq$$

$$\geqq \frac{C}{2(1+3\cos\varrho-4\cos\varrho\, C)}(1+C_2).$$

Damit ist (9.11) gezeigt.

Zum Beweis von (9.12) beachte man: Wegen (2.1) und (3.6) ist $\tan\alpha = = \frac{1}{\cos\varrho\tan\beta}$ äquivalent zu $\sin r = \sin\beta\sin\varrho$ und zu $\sin d = \sin 2\beta\sin\varrho$, so daß (9.12) äquivalent ist zu

$$\text{(9.16)} \qquad -(1+\cos\varrho-C)\arccos\frac{\sin\beta\cos\varrho}{\sqrt{1-\sin^2\beta\sin^2\varrho}} - \frac{\pi}{2}\frac{3+\cos\varrho}{1+3\cos\varrho}C +$$

$$+\frac{4C}{1+3\cos\varrho}\beta + 2\beta + 2\cos\varrho\arccos\frac{\sin 2\beta\cos\varrho}{\sqrt{1-\sin^2 2\beta\sin^2\varrho}} > 0 \quad \text{für } \beta_0 \leqq \beta \leqq \frac{\pi}{4}.$$

Diese Ungleichung ist für $\beta = \beta_0$ richtig, wie oben gezeigt wurde. Sie ist auch für $\beta = \frac{\pi}{4}$ richtig; denn die dann entstehende Funktion ist gleich 0 für $\varrho = 0$ und streng monoton wachsend in ϱ.

Um (9.16) weiter zu untersuchen, betrachte man ihre Ableitung nach β

(9.17)

$$(1+\cos\varrho-C)\frac{\cos\varrho}{1-\sin^2\beta\sin^2\varrho} + \frac{4C}{1+3\cos\varrho} + 2 - 4\cos^2\varrho\,\frac{1}{1-\sin^2 2\beta\sin^2\varrho}.$$

Für $\beta = \beta_0$ ist (9.17) > 0, wie man nachrechnet. Für $\beta = \frac{\pi}{4}$ ist (9.17) < 0; denn weil $\frac{\cos\varrho}{1-\frac{1}{2}\sin^2\varrho}$ fallend ist, genügt es zu zeigen, daß

$$(1+\cos\varrho-C)C_3 + \frac{4C}{1+3\cos\varrho} - 2 < 0$$

für $0 < \varrho \leqq \frac{\pi}{3}$ und $C_3 := 1$, und für $\frac{\pi}{3} \leqq \varrho < \varrho_0$ und $C_3 := 0.8$, und Ableitung nach ϱ zeigt, daß diese Funktion ihr Maximum an den Intervallrändern annimmt.

---

[3] Damit ist gemeint: Gezeigt werden soll die Ungleichung $U(\varrho) \geqq 0$ in einer einzigen Variablen ϱ, wobei ϱ aus einem Intervall I ist. Man beachte, daß sich $U(\varrho)$ aus monotonen Funktionen $f_j(\varrho)$ zusammensetzt. Nun läßt sich das Intervall I so in Intervalle I_i zerlegen, daß die Gültigkeit von $U(\varrho) \geqq 0$ in I_i wegen der Monotonie von f_j aus den Werten von f_j an den Intervallrändern geschlossen werden kann, wenn diese numerisch bekannt sind.

Dieses Verfahren läßt sich allerdings im allgemeinen nur anwenden, wenn in I sogar gilt $U(\varrho) \geqq \varepsilon$ für ein festes $\varepsilon > 0$.

Um (9.17) weiter zu untersuchen, beachte man, daß ihre Ableitung nach β genau dann gleich 0 ist, wenn gilt

$$\frac{1+\cos\varrho - C}{16\cos\varrho} = \cos 2\beta \frac{(1-\sin^2\beta\sin^2\varrho)^2}{(1-\sin^2 2\beta\sin^2\varrho)^2}. \tag{9.18}$$

Schließlich ist die Ableitung nach β der rechten Seite von (9.18) für $\beta_0 \leqq \beta \leqq \frac{\pi}{4}$ genau dann < 0, wenn gilt

$$\begin{gathered}(1-\sin^2\beta\sin^2\varrho)(-\cos^2\varrho + 2\cos^2 2\beta\sin^2\varrho)+ \\ + \cos 2\beta\sin^2\beta\sin^2\varrho\,(-2+3\sin^2\varrho - 2\sin^2\beta\sin^2\varrho) < 0 \quad \text{für } \beta_0 \leqq \beta \leqq \frac{\pi}{4}.\end{gathered} \tag{9.19}$$

Alle in (9.19) auftretenden Klammern werden für $\beta = \beta_0$ maximal; somit ist in jedem der beiden Summanden genau ein Faktor > 0, so daß (9.19) gilt.

Nun betrachte man successive die Ungleichungen (9.19) bis (9.16). Wegen (9.19) ist die rechte Seite von (9.18) für $\beta_0 \leqq \beta \leqq \frac{\pi}{4}$ streng monoton fallend in β, so daß (9.18) dort für höchstens ein β gilt, für jedes feste ϱ. Deshalb hat die Funktion (9.17) im Intervall $\beta_0 \leqq \beta \leqq \frac{\pi}{4}$ höchstens einen Hoch- oder Tiefpunkt; wegen ihrer Werte in den Intervallenden hat sie also dort genau eine Nullstelle, die einem Hochpunkt der Funktion in (9.16) entspricht. Deshalb nimmt die Funktion in (9.16) ihr Minimum in den Intervallrändern an, und weil für diese (9.16) richtig ist, ist (9.16) richtig und also (9.12). □

10. Die Funktion $e_\varrho(\alpha,\beta)$ in G_4, und die konkave Funktion $E_\varrho(\alpha,\beta)$

In G_4 ist $e_\varrho(\alpha,\beta)$ durch (3.13) definiert. Es ist zweckmäßig, im jetzigen Abschnitt **10**. diese Funktion auf einen größeren Definitionsbereich fortzusetzen, bei gleicher Bezeichnung. Betrachtet wird also die Funktion

$$e_\varrho(\alpha,\beta) = \left(\alpha - \arccos\frac{\tan r}{\tan\varrho}\right)(1-\cos\varrho), \quad \text{für } 0 \leqq r \leqq \varrho, \text{ mit} \tag{10.1}$$

$$r = \arccos\frac{\cos\beta}{\sin\alpha}, \quad \text{für } \alpha+\beta \geqq \frac{\pi}{2},\ \alpha \neq 0. \tag{10.2}$$

Der Definitionsbereich von (10.1) ist also der Bereich von $\{(\alpha,\beta)\,|\,\alpha+\beta \geqq \frac{\pi}{2}\}$ mit $\alpha \neq 0$ und $r \leqq \varrho$, vgl. Fig. 1.

Für die folgenden Untersuchungen wesentlich ist die Streckenschar zum Scharparameter C in der (α,β)-Ebene

$$\alpha = \frac{\pi}{2} - C\beta \quad (0 \leqq C \leqq 1), \quad 0 \leqq \beta \leqq \frac{\pi}{2}. \tag{10.3}$$

Jede Scharstrecke geht durch den Punkt $\left(\frac{\pi}{2},0\right)$. Sie verbindet ihn mit dem Punkt $((1-C)\frac{\pi}{2},\frac{\pi}{2})$, so daß die Schar den Bereich $\{(\alpha,\beta)\,|\,\alpha+\beta \geqq \frac{\pi}{2}\}$ überdeckt.

Nun gilt

Hilfssatz 10.1. *Die Funktion* (10.2) *ist längs jeder Strecke* (10.3) *konvex.*

Beweis. Längs einer festen Strecke (10.3), d.h. bei festem C, wird aus (10.2)

$$r = \arccos \frac{\cos\beta}{\cos C\beta} \quad \text{für } 0 \leqq \beta \leqq \frac{\pi}{2}. \tag{10.4}$$

Die Konvexität von (10.4) ist äquivalent zu

$$\begin{aligned} &-\cos^2 C\beta \sin^2 C\beta \cos\beta + 2C\cos^3 C\beta \sin C\beta \sin\beta - C^2\cos^2 C\beta\cos\beta + \\ &+C^2\cos^3\beta - C^2\cos^2 C\beta \sin^2 C\beta \cos\beta \geqq 0 \quad \text{für } 0 \leqq \beta \leqq \frac{\pi}{2},\ 0 \leqq C \leqq 1. \end{aligned} \tag{10.5}$$

Nun gilt wegen der Konkavität von $\sin x$ und der Konvexität von $\tan x$ $\sin C\beta \geqq C\sin\beta$ und $C\cos C\beta \sin\beta - \sin C\beta\cos\beta \geqq 0$. Deshalb gilt

$$\begin{aligned} &2C\cos^3 C\beta \sin C\beta \sin\beta - \cos^2 C\beta \sin^2 C\beta \cos\beta \geqq \\ &\geqq C\cos^3 C\beta \sin C\beta \sin\beta + \cos^2 C\beta (C\sin\beta)[C\cos C\beta\sin\beta - \sin C\beta\cos\beta] \geqq \\ &\geqq C^2\cos^3 C\beta \sin^2\beta + C\cos^2 C\beta (C\sin\beta)\sin\beta[\cos C\beta - \cos\beta] = \\ &= 2C^2\cos^3 C\beta\sin^2\beta - C\cos^2 C\beta\cos\beta\sin^2\beta. \end{aligned}$$

Mit Hilfe dieser Abschätzung wird aus (10.5)

$$\begin{aligned} &(\cos C\beta - \cos\beta)(2\cos^2 C\beta - \cos C\beta\cos\beta - \cos^2\beta) + \\ &+\cos^2 C\beta\cos\beta(\cos^2\beta - 2\cos C\beta\cos\beta + \cos^2 C\beta) \geqq 0, \end{aligned}$$

was offensichtlich richtig ist. □

Hilfssatz 10.2. *Die Funktion* (10.1) *ist (innerhalb ihres Definitionsbereichs) längs jeder Strecke* (10.3) *konvex.*

Beweis. Längs einer festen Strecke (10.3), d.h. bei festem C wird aus (10.1)

$$e_\varrho\left(\frac{\pi}{2} - C\beta, \beta\right) = \left(\frac{\pi}{2} - C\beta - \arccos\frac{\tan r}{\tan\varrho}\right)(1-\cos\varrho), \quad \text{für } 0 \leqq r \leqq \varrho$$

und r wie in (10.4). Nun ist die Funktion $-\arccos\frac{\tan r}{\tan\varrho}$ in r wachsend und konvex für $0 \leqq r \leqq \varrho$, wie man durch Ableiten sieht. Zusammen mit Hilfssatz 10.1 folgt daraus die Behauptung. □

Es wird eine weitere konkave Hilfsfunktion eingeführt, und zwar gilt

HILFSSATZ 10.3. *Zu jedem ϱ mit $0 < \varrho < \frac{\pi}{2}$ gibt es eine in* $\operatorname{conv} G_4$ *definierte Funktion $E_\varrho(\alpha,\beta)$ mit*

$$(10.6) \qquad E_\varrho(\alpha,\beta) \text{ ist konkav,}$$

$$(10.7) \qquad E_\varrho(\alpha,\beta) \geqq e_\varrho(\alpha,\beta) \quad \text{in} \quad G_4,$$

$$(10.8) \qquad \begin{gathered} E_\varrho(\alpha,\beta) = e_\varrho(\alpha,\beta) = \left(\alpha - 2\arcsin\frac{1}{2\cos\frac{\varrho}{2}}\right)(1-\cos\varrho) \\ \text{für } (\alpha,\beta) \text{ auf der Kurve } (3.2), \\ E_\varrho(\alpha,\beta) = e_\varrho(\alpha,\beta) = 0 \quad \text{für } (\alpha,\beta) = \left(\frac{\pi}{2},0\right), \end{gathered}$$

(10.9) *Jeder Punkt von* $\operatorname{conv} G_4$ *liegt entweder auf der Verbindungsstrecke des Punktes $(\frac{\pi}{2},0)$ mit einem Punkt der Kurve (3.2), oder auf der Verbindungsstrecke zweier Punkte von (3.2), und längs jeder solcher Verbindungsstrecke ist $E_\varrho(\alpha,\beta)$ linear.*

BEWEIS. Die Kurve (3.2) ist die G_4 begrenzende konvexe Kurve

$$(3.2) \qquad \begin{cases} r = \frac{\varrho}{2} & \text{oder} \\ \beta = \arccos\left(\sin\alpha\cos\frac{\varrho}{2}\right), & 0 \leqq \alpha \leqq \frac{\pi}{2}. \end{cases}$$

Zusammen mit der Konkavität von (3.1) folgt also die Aussage von (10.9) bzgl. $\operatorname{conv} G_4$.

Nun gilt

$$(10.10) \quad z = \left(\alpha - \arccos\frac{\tan\frac{\varrho}{2}}{\tan\varrho}\right)(1-\cos\varrho) = \left(\alpha - 2\arcsin\frac{1}{2\cos\frac{\varrho}{2}}\right)(1-\cos\varrho)$$

ist Ebene durch den Punkt (3.3) in der (α,β)-Ebene; identisch mit $z = e_\varrho(\alpha,\beta)$ (vgl. (10.1)) für $(\alpha,\beta) \in (3.2)$.

Jeder Punkt (α,β,z) auf (10.10) über einem Punkt von (3.2), d.h. jeder Punkt $\left(\alpha, \arccos\left(\sin\alpha\cos\frac{\varrho}{2}\right), \left(\alpha - \arccos\frac{\tan\frac{\varrho}{2}}{\tan\varrho}\right)(1-\cos\varrho)\right)$ werde mit dem Punkt $(\alpha,\beta,z) = (\frac{\pi}{2},0,0)$ verbunden. Alle diese Verbindungsstrecken für $0 \leqq \alpha \leqq \frac{\pi}{2}$ bestimmen eine Strahlfläche, die wegen der Konvexität von (3.2) konkav ist. Es sei $z = E_\varrho(\alpha,\beta)$ diese Strahlfläche, wenn (α,β) in dem Bereich von $\operatorname{conv} G_4$ liegt mit $r \leqq \frac{\varrho}{2}$. In dem Bereich von $\operatorname{conv} G_4$ mit $r \geqq \frac{\varrho}{2}$ aber sei $z = E_\varrho(\alpha,\beta)$ durch (10.10) definiert.

(10.8), (10.9) und (10.6) folgen dann direkt aus der Definition von $E_\varrho(\alpha,\beta)$, und (10.7) folgt aus Hilfssatz 10.2. □

11. Eine Ungleichung für einen Punkt auf $h_\varrho(\alpha,\beta)$ und einen Punkt auf $E_\varrho(\alpha,\beta)$

Nach diesen Vorbereitungen kann nun gezeigt werden

HILFSSATZ 11.1. *Es seien Punkte* (α_1,β_1) *und* (α_2,β_2) *gegeben mit*

$$(\alpha_1,\beta_1)\in \operatorname{conv} G_4 \quad \textit{und} \quad \alpha_2+\beta_2 \geqq \frac{\pi}{2}. \tag{11.1}$$

Für ganze Zahlen $\mu,\nu>0$ *mit* $\nu\leqq 2\mu$ *gelte*

$$\frac{\mu\beta_1+\nu\beta_2}{\mu+\nu}\geqq\frac{\pi}{3}. \tag{11.2}$$

Dann gilt für ϱ *mit* $0<\varrho<\varrho_0$ *die Ungleichung*

$$\mu E_\varrho(\alpha_1,\beta_1)+\nu h_\varrho(\alpha_2,\beta_2)\leqq(\mu+\nu)H_\varrho\Big(\frac{\mu\alpha_1+\nu\alpha_2}{\mu+\nu},\frac{\mu\beta_1+\nu\beta_2}{\mu+\nu}\Big). \tag{11.3}$$

Wenn nicht gleichzeitig $(\alpha_1,\beta_1)=(\frac{\pi}{2},0)$ *und* $(\alpha_2,\beta_2)=(0,\frac{\pi}{2})$, *so gilt in* (11.3) *sogar das* $<$*-Zeichen.*

BEWEIS. Wegen (11.1) ist $\beta_1<\frac{\pi}{3}$, wegen (11.2) ist also $\beta_2>\frac{\pi}{3}$, so daß $\beta_2>\beta_1$, und so daß es ein λ_0 gibt mit $0<\lambda_0<1$ und mit

$$\lambda_0\beta_1+(1-\lambda_0)\beta_2=\frac{\pi}{3}. \tag{11.4}$$

Wegen $\beta_2>\beta_1$ ist $\frac{\beta_1+\frac{\nu}{\mu}\beta_2}{1+\frac{\nu}{\mu}}$ wachsend in $\frac{\nu}{\mu}$. Jedes Paar (β_1,β_2), das (11.1) und (11.2) erfüllt, erfüllt deshalb erst recht

$$\frac{\beta_1+2\beta_2}{3}\geqq\frac{\pi}{3}. \tag{11.5}$$

Nun betrachte man im Raum $\{(\alpha,\beta,z)\}$ die Punkte $(\alpha_1,\beta_1,E_\varrho(\alpha_1,\beta_1))$ und $(\alpha_2,\beta_2,h_\varrho(\alpha_2,\beta_2))$. Ihre Verbindungsstrecke ist die Punktmenge

$$\lambda(\alpha_1,\beta_1,E_\varrho(\alpha_1,\beta_1))+(1-\lambda)(\alpha_2,\beta_2,h_\varrho(\alpha_2,\beta_2)) \qquad (0\leqq\lambda\leqq 1). \tag{11.6}$$

Dann bedeutet (11.3), daß die Strecke (11.6) im Punkt mit $\lambda=\frac{\mu}{\mu+\nu}$ unterhalb der Fläche $z=H_\varrho(\alpha,\beta)$ liegt. (11.6) liegt im Punkt mit $\lambda=0$ unterhalb der Fläche $H_\varrho(\alpha,\beta)$ wegen (9.3), $H_\varrho(\alpha,\beta)$ ist konkav nach (9.1), und wegen (11.2) und (11.4) ist $\frac{\mu}{\mu+\nu}\leq\lambda_0$. Deshalb ist (11.3) also richtig, wenn (11.6) im Punkt mit $\lambda=\lambda_0$ unterhalb von $H_\varrho(\alpha,\beta)$ liegt. Der Punkt von (11.6) mit $\lambda=\lambda_0$ liegt wegen (11.4) in der Ebene $\beta=\frac{\pi}{3}$, und es ist $H_\varrho(\alpha,\frac{\pi}{3})=F(\alpha)$ nach (9.2); deshalb ist (11.3) richtig, wenn (11.6) die Ebene $\beta=\frac{\pi}{3}$ unterhalb von $F(\alpha)$ schneidet. Wenn (11.6) die Ebene $\beta = \frac{\pi}{3}$ echt unterhalb von $F(\alpha)$ schneidet, so steht in (11.3) sogar das $<$-Zeichen.

Wir definieren: Ein Punktepaar habe die (*echte*) *Schnitteigenschaft*, wenn seine Verbindungsstrecke die Ebene $\beta = \frac{\pi}{3}$ (echt) unterhalb von $F(\alpha)$ schneidet. Damit ist Hilfssatz 11.1 richtig, wenn gilt

Das Punktepaar $(\alpha_1, \beta_1, E_\varrho(\alpha_1, \beta_1))$ *und* $(\alpha_2, \beta_2, h_\varrho(\alpha_2, \beta_2))$ *mit* (11.1) *und* (11.5) *hat die Schnitteigenschaft, wobei* $0 < \varrho < \varrho_0$.

(11.7) *Das Punktepaar hat sogar die echte Schnitteigenschaft, wenn nicht gleichzeitig*

$$(\alpha_1, \beta_1) = (\tfrac{\pi}{2}, 0) \text{ und } (\alpha_2, \beta_2) = (0, \tfrac{\pi}{2}).$$

Zum Beweis von (11.7) wird die folgende Bemerkung hilfreich sein: Aus der Konkavität von $F(\alpha)$ folgt für die Ableitung von $F(\alpha)$, daß $1 \geqq F'(\alpha) \geqq 1 - \cos\varrho$. Deshalb gilt

Liegt ein Punkt $(\alpha_0, \frac{\pi}{3}, z_0)$ unterhalb von $F(\alpha)$, d.h. ist $z_0 \leqq F(\alpha_0)$, dann liegt die durch den Punkt $(\alpha_0, \frac{\pi}{3}, z_0)$ gehende Gerade

(11.8) $$\{(\alpha, \tfrac{\pi}{3}, z) \mid z = z_0 + m(\alpha - \alpha_0)\}$$

echt unterhalb von $F(\alpha)$ für $\frac{\pi}{6} \leqq \alpha < \alpha_0$, wenn $m \geqq 1$, und für $\alpha_0 < < \alpha \leqq \frac{\pi}{2}$, wenn $m \leqq 1 - \cos\varrho$.

Im folgenden wesentlich ist die Ebene (vgl. (10.8))

(11.9) $$z = \left(\alpha - 2\arcsin\frac{1}{2\cos\frac{\varrho}{2}}\right)(1 - \cos\varrho).$$

Im folgenden sei ein Punkt (,) ein Punkt der (α, β)-Ebene, und ein Punkt (, ,) ein Punkt des (α, β, z)-Raums. $e_\varrho(\alpha, \beta)$ *bezieht sich auf die Funktion* (10.1). Der Beweis von (11.7) geschieht nun in 5 Schritten.

1. Schritt. (α_1, β_1, z_1) sei ein Punkt der Ebene (11.9) mit

$$\frac{\varrho}{2} \leqq \beta_1 \leqq \arcsin\frac{1}{2\cos\frac{\varrho}{2}} \quad \text{und mit} \quad (\alpha_1, \beta_1) \in (3.2).$$

Dann haben (α_1, β_1, z_1) und $(0, \frac{\pi}{2}, 0)$ die echte Schnitteigenschaft.

Beweis. (3.2) enthält den Punkt (3.3), d.h. den Punkt

$$\left(2\arcsin\frac{1}{2\cos\frac{\varrho}{2}}, \arcsin\frac{1}{2\cos\frac{\varrho}{2}}\right).$$

Der Punkt $(0, \frac{\pi}{2}, 0)$ und ein Punkt $\left(2\arcsin\frac{1}{2\cos\frac{\varrho}{2}}, \beta, 0\right)$ für $\frac{\pi}{6} \leqq \beta \leqq \leqq \arcsin\frac{1}{2\cos\frac{\varrho}{2}}$ haben die Schnitteigenschaft trivialerweise. Die Ebene (11.9) enthält die Gerade $(2\arcsin\frac{1}{2\cos\frac{\varrho}{2}}, \beta, 0)$ und hat die Steigung $m = 1 - \cos\varrho$ in α-Richtung. Wegen Bemerkung (11.8) hat ein im 1. Schritt betrachtetes Punktepaar also die echte Schnitteigenschaft, wenn $\frac{\pi}{6} \leqq \beta_1 \leqq \arcsin\frac{1}{2\cos\frac{\varrho}{2}}$.

Der Fall $\frac{\varrho}{2} \leqq \beta_1 \leqq \frac{\pi}{6}$ tritt genau für $\varrho < \frac{\pi}{3}$ auf. Dann enthält die Kurve (3.5) Punkte (α_3, β_3) mit $\beta_3 < \frac{\pi}{3}$. Wegen (7.1) und (10.1) gilt $h_\varrho(\alpha_3, \beta_3) = \alpha_3(1-\cos\varrho) = e_\varrho(\alpha_3, \beta_3)$. $(0, \frac{\pi}{2}, 0)$ und $(\alpha_3, \beta_3, h_\varrho(\alpha_3, \beta_3))$ für $\beta_3 \leqq \frac{\pi}{3}$ haben die echte Schnitteigenschaft wegen der Konvexität von (3.5) und wegen (9.3) und (7.1). Außerdem haben $(0, \frac{\pi}{2}, 0)$ und $(\frac{\pi}{2}, 0, 0)$ die Schnitteigenschaft trivialerweise. $(0, \frac{\pi}{2}, 0)$ und jeder innere Punkt der Verbindungsstrecke von $(\alpha_3, \beta_3, e_\varrho(\alpha_3, \beta_3))$ mit $(\frac{\pi}{2}, 0, 0)$ haben deshalb ebenfalls die echte Schnitteigenschaft wegen der Konkavität von $F(\alpha)$.

Die Verbindungsstrecken des Punktes $(\frac{\pi}{2}, 0)$ mit den Punkten (α_3, β_3) für $\beta_3 \leqq \frac{\pi}{3}$ überdecken nun das Kurvenstück (3.2) für $\beta_1 \leqq \frac{\pi}{6}$ wegen Hilfssatz 10.1. Wegen (10.8) ist nun $z_1 = e_\varrho(\alpha_1, \beta_1)$, so daß aus Hilfssatz 10.2 folgt, daß $(0, \frac{\pi}{2}, 0)$ und (α_1, β_1, z_1) auch für $\beta_1 \leqq \frac{\pi}{6}$ die echte Schnitteigenschaft haben, w.z.z.w.

2. Schritt. Es werden Punktepaare (α_1, β_1, z_1) und (α_2, β_2, z_2) betrachtet mit $\frac{1}{3}\beta_1 + \frac{2}{3}\beta_2 = \frac{\pi}{3}$. Dabei liege (α_1, β_1, z_1) in der Ebene (11.9) und (α_2, β_2, z_2) auf der Geraden (7.7). Gezeigt wird, daß die folgenden drei derartigen Punktepaare die echte Schnitteigenschaft haben, für $0 < \varrho < \varrho_0$:

(11.10) $$\left(2 \arcsin \frac{1}{2\cos\frac{\varrho}{2}}, \arcsin \frac{1}{2\cos\frac{\varrho}{2}}, 0\right)$$

und

(11.11) $$\left(\left(\frac{\pi}{2} - \beta\right)\frac{1}{\cos\varrho}, \beta, \left(\frac{\pi}{2} - \beta\right)\left(\frac{1}{\cos\varrho} - 1\right)\right) \text{ mit } \beta = \frac{\pi}{2} - \frac{1}{2}\arcsin\frac{1}{2\cos\frac{\varrho}{2}},$$

(11.12) $$\left(2 \arcsin \frac{1}{2\cos\frac{\varrho}{2}}, \frac{\pi}{6}, 0\right)$$

und

(11.13) $$\left(\left(\frac{\pi}{2} - \beta\right)\frac{1}{\cos\varrho}, \beta, \left(\frac{\pi}{2} - \beta\right)\left(\frac{1}{\cos\varrho} - 1\right)\right) \text{ mit } \beta = \frac{5\pi}{12},$$

(11.14) $$\left(\frac{\pi}{2}\frac{3+\cos\varrho}{1+3\cos\varrho} - \frac{2\varrho}{1+3\cos\varrho}, \frac{\varrho}{2}, \left(\frac{\pi}{2}\frac{3+\cos\varrho}{1+3\cos\varrho} - \frac{2\varrho}{1+3\cos\varrho} - 2\arcsin\frac{1}{2\cos\frac{\varrho}{2}}\right)(1-\cos\varrho)\right)$$

und

(11.15) $$\left(\left(\frac{\pi}{2} - \beta\right)\frac{1}{\cos\varrho}, \beta, \left(\frac{\pi}{2} - \beta\right)\left(\frac{1}{\cos\varrho} - 1\right)\right) \quad \text{mit} \quad \beta = \frac{\pi}{2} - \frac{\varrho}{4}.$$

BEWEIS für das Punktepaar (11.10) und (11.11). Mit $t := \arcsin \frac{1}{2\cos\frac{\varrho}{2}}$ wird aus (11.10) und (11.11)

$$(2t, t, 0) \tag{11.16}$$

und

$$\left(t\frac{\sin^2 t}{1-2\sin^2 t}, \frac{\pi}{2} - \frac{1}{2}t, t\frac{\sin^2 t}{1-2\sin^2 t} - \frac{1}{2}t\right), \tag{11.17}$$

jeweils für $\frac{\pi}{6} < t < \arcsin \frac{1}{2\cos\frac{\varrho_0}{2}} =: t_0$.

(11.16) beschreibt für $\frac{\pi}{6} \leqq t \leqq t_0$ ein Geradenstück $G(t)$, das durch t linear parametrisiert ist. (11.17) beschreibt für $\frac{\pi}{6} \leqq t \leqq t_0$ ein Kurvenstück $\ell(t)$ in der Ebene $z = \alpha + \beta - \frac{\pi}{2}$. Sei $G\ell(t)$ für $\frac{\pi}{6} \leqq t \leqq t_0$ die Verbindungsstrecke von $\ell(\frac{\pi}{6})$ und $\ell(t_0)$, linear parametrisiert in t. Nun rechnet man nach, daß (11.16) und (11.17) die Schnitteigenschaft haben für $t = \frac{\pi}{6}$ und $t = t_0$, d.h. die Punktepaare $G(\frac{\pi}{6})$, $G\ell(\frac{\pi}{6})$ und $G(t_0)$, $G\ell(t_0)$ haben die Schnitteigenschaft. Wegen der Konkavität von $F(\alpha)$ und aus Linearitätsgründen hat dann für jedes t das Punktepaar $G(t)$ und $G\ell(t)$ die Schnitteigenschaft.

Die β-Koordinaten von $\ell(t)$ und $G\ell(t)$ hängen beide linear von t ab, sind also für jedes feste t gleich groß. Die Funktion $t\frac{\sin^2 t}{1-2\sin^2 t}$ ist für $\frac{\pi}{6} < t < t_0$ streng konvex in t, während die α-Koordinate von $G\ell(t)$ linear von t abhängt; deshalb ist die α-Koordinate von $\ell(t)$ kleiner als die von $G\ell(t)$ für jedes t mit $\frac{\pi}{6} < t < t_0$. Schließlich liegen $\ell(t)$ und $G\ell(t)$ beide in der Ebene $z = \alpha + \beta - \frac{\pi}{2}$ mit Steigung $m = 1$ in α-Richtung. Aus der Schnitteigenschaft von $G(t)$ und $G\ell(t)$ folgt deshalb nach (11.8) die echte Schnitteigenschaft von $G(t)$ und $\ell(t)$ für jedes t mit $\frac{\pi}{6} < t < t_0$.

BEWEIS für das Punktepaar (11.12) und (11.13). Völlig analog, wobei man wieder $t := \arcsin \frac{1}{2\cos\frac{\varrho}{2}}$ setzt; die benötigte Schnitteigenschaft von (11.12) und (11.13) für $\varrho = 0$ und $\varrho = \varrho_0$ und die strenge Konvexität der Funktion $\frac{\pi}{6}\frac{\sin^2 t}{1-2\sin^2 t}$ rechnet man nach.

BEWEIS für das Punktepaar (11.14) und (11.15). Die echte Schnitteigenschaft von (11.14) und (11.15) ist nach Definition äquivalent zu

$$F(\alpha(\varrho)) - \left(\alpha(\varrho) - \frac{2}{3}\arcsin\frac{1}{2\cos\frac{\varrho}{2}}\right)(1-\cos\varrho) > 0 \quad \text{für } 0 < \varrho < \varrho_0 \tag{11.18}$$

mit

$$\alpha(\varrho) := \frac{1}{3}\left(\frac{\pi}{2}\frac{3+\cos\varrho}{1+3\cos\varrho} - \frac{2\varrho}{1+3\cos\varrho}\right) + \frac{\varrho}{6}\frac{1}{\cos\varrho}.$$

Zum Beweis von (11.18) rechnet man nach, daß für die Ableitung von $\alpha(\varrho)$ nach ϱ gilt

$$\alpha'(\varrho) \geqq \frac{\pi}{12}\sin\varrho, \tag{11.19}$$

so daß $\alpha(\varrho)$ wachsend in ϱ ist. Deshalb setzt sich die linke Seite von (11.18) aus Differenz und Produkt von in ϱ wachsenden Funktionen zusammen, und numerisches Nachrechnen zeigt (vgl. Fußnote [3] in **9.**), daß (11.18) gilt für $0.1 \leqq \varrho < \varrho_0$. Zum Beweis von (11.18) für $0 < \varrho \leqq 0.1$ rechnet man numerisch nach, daß $\alpha(\varrho) - \frac{2}{3} \arcsin \frac{1}{2\cos\frac{\varrho}{2}} < 0.2$ für $0 < \varrho \leqq 0.1$, so daß zu zeigen bleibt

$$F(\alpha(\varrho)) - 0.2(1-\cos\varrho) > 0 \quad \text{für } 0 < \varrho \leqq 0.1. \tag{11.20}$$

Nun ist die linke Seite von (11.20) streng wachsend in ϱ, wenn gilt

$$F'(\alpha(\varrho))\alpha'(\varrho) - 0.2\sin\varrho > 0 \quad \text{für } 0 < \varrho \leqq 0.1. \tag{11.21}$$

Nach Hilfssatz 4.3 ist $F'(\alpha)$ fallend in α und also fallend in ϱ, so daß $F'(\alpha(\varrho)) \geqq 0.9$ für $0 < \varrho \leqq 0.1$. Zusammen mit (11.19) ist also (11.21) richtig, so daß die Funktion in (11.20) streng wachsend in ϱ ist. Weil außerdem in (11.20) für $\varrho = 0$ das Gleichheitszeichen gilt, ist (11.20) für $0 < \varrho \leqq 0.1$ richtig.

3. Schritt. Es werden Punktepaare (α_1, β_1, z_1) und (α_2, β_2, z_2) betrachtet mit $\frac{1}{3}\beta_1 + \frac{2}{3}\beta_2 = \frac{\pi}{3}$. Dabei liege (α_1, β_1, z_1) in der Ebene (11.9) und (α_2, β_2, z_2) auf der Geraden (7.7). Gezeigt wird, daß alle derartigen Punktepaare die echte Schnitteigenschaft haben, wenn $\frac{\varrho}{2} \leqq \beta_1 \leqq \arcsin \frac{1}{2\cos\frac{\varrho}{2}}$ und wenn $(\alpha_1, \beta_1) \in (3.2)$.

Beweis. Es sei $\mathcal{S}$ der Streckenzug, der die Punkte (11.10), (11.12) und (11.14) verbindet; $\mathcal{S}$ ist eine Punktmenge $\left\{ S(\beta_1) \mid \frac{\varrho}{2} \leqq \beta_1 \leqq \arcsin \frac{1}{2\cos\frac{\varrho}{2}} \right\}$. Aus dem 2. Schritt folgt aus Linearitätsgründen die echte Schnitteigenschaft für jedes Punktepaar $S(\beta_1)$ und (α_2, β_2, z_2). Weil $S(\beta_1)$ und (α_1, β_1, z_1) beide in der Ebene (11.9) liegen mit Steigung $m = 1 - \cos\varrho$ in α-Richtung, haben nach (11.8) also auch (α_1, β_1, z_1) und (α_2, β_2, z_2) die echte Schnitteigenschaft, sofern α_1 größer oder gleich der α-Koordinate $\alpha_S(\beta_1)$ von $S(\beta_1)$ ist, d.h. sofern gilt

$$\alpha_S(\beta_1) \leqq \alpha_1(\beta_1) \quad \text{für } \frac{\varrho}{2} \leqq \beta_1 \leqq \arcsin \frac{1}{2\cos\frac{\varrho}{2}}. \tag{11.22}$$

Nach Definition des Punktes (11.12) ist (11.22) offensichtlich richtig für $\frac{\pi}{6} \leqq \beta_1 \leqq \arcsin \frac{1}{2\cos\frac{\varrho}{2}}$, so daß (11.22) zu zeigen bleibt für $\frac{\varrho}{2} \leqq \beta_1 \leqq \frac{\pi}{6}$.

Man betrachte die Gerade (8.3). Für jeden Punkt $(\alpha(\beta_1), \beta_1)$ von (8.3) gilt nach (8.1) (vgl. Fig. 2), daß

$$\alpha(\beta_1) \leqq \alpha_1(\beta_1) \quad \text{für } \frac{\varrho}{2} \leqq \beta_1 \leqq \frac{\pi}{6}. \tag{11.23}$$

Nun wurde der Punkt (11.14) gerade als Punkt $(\alpha(\frac{\varrho}{2}), \frac{\varrho}{2},)$ definiert, so daß $\alpha_S(\frac{\varrho}{2}) = \alpha(\frac{\varrho}{2})$. Außerdem ist

$$\alpha_S\left(\frac{\pi}{6}\right) = 2\arcsin \frac{1}{2\cos\frac{\varrho}{2}} \leqq \frac{\pi}{6}\left(1 + \frac{4}{1+3\cos\varrho}\right) = \alpha\left(\frac{\pi}{6}\right),$$

wie man mit Hilfe der 1. Ableitung zeigt. Deshalb gilt für die linearen Funktionen $\alpha_S(\beta_1)$ und $\alpha(\beta_1)$

$$(11.24)\qquad \alpha_S(\beta_1) \leqq \alpha(\beta_1) \quad \text{für } \frac{\varrho}{2} \leqq \beta_1 \leqq \frac{\pi}{6}.$$

Aus (11.24) und (11.23) folgt (11.22) für $\frac{\varrho}{2} \leqq \beta_1 \leqq \frac{\pi}{6}$.

4. Schritt. $(\alpha_1, \beta_1, E_\varrho(\alpha_1, \beta_1))$ und $(\alpha_2, \beta_2, z_2) \in (7.7)$ haben die echte Schnitteigenschaft, wenn $(\alpha_1, \beta_1) \in \operatorname{conv} G_4$ und wenn $\frac{1}{3}\beta_1 + \frac{2}{3}\beta_2 \geqq \frac{\pi}{3}$ und $\beta_1 \neq 0$.

Beweis. Zunächst beachte man: Nach (10.9) ist $E_\varrho(\alpha_1, \beta_1)$ in $\operatorname{conv} G_4$ bestimmt durch seine Werte in $(\frac{\pi}{2}, 0)$ und in $(\alpha_1, \beta_1) \in (3.2)$. Dabei gilt nach (10.8), daß $(\alpha_1, \beta_1, E_\varrho(\alpha_1, \beta_1))$ in der Ebene (11.9) liegt für $(\alpha_1, \beta_1) \in (3.2)$, und $E_\varrho(\frac{\pi}{2}, 0) = 0$.

Nun haben $(\frac{\pi}{2}, 0, 0)$ und $(0, \frac{\pi}{2}, 0,)$ die Schnitteigenschaft trivialerweise, und für ihre β-Koordinaten gilt $\frac{1}{3} \cdot 0 + \frac{2}{3}\frac{\pi}{2} = \frac{\pi}{3}$. Zusammen mit dem 3. Schritt folgt daraus also aus Linearitätsgründen und wegen der Konkavität von $F(\alpha)$, daß $(\alpha_1, \beta_1, E_\varrho(\alpha_1, \beta_1))$ und $(\alpha_2, \beta_2, z_2) \in (7.7)$ die echte Schnitteigenschaft haben, wenn $(\alpha_1, \beta_1) \in \operatorname{conv} G_4$ und wenn speziell $\frac{1}{3}\beta_1 + \frac{2}{3}\beta_2 = \frac{\pi}{3}$ und $\beta_1 \neq 0$.

Weil $(\frac{\pi}{2}, 0, 0)$ und $(0, \frac{\pi}{2}, 0)$ die Schnitteigenschaft haben, folgt aus dem 1. Schritt ebenso aus Linearitätsgründen und wegen der Konkavität von $F(\alpha)$, daß $(\alpha_1, \beta_1, E_\varrho(\alpha_1, \beta_1))$ und $(0, \frac{\pi}{2}, 0)$ die echte Schnitteigenschaft haben, wenn $(\alpha_1, \beta_1) \in \operatorname{conv} G_4$ und $\beta_1 \neq 0$.

Aus diesen beiden Ergebnissen folgt schließlich wieder mit Linearitätsgründen und wegen der Konkavität von $F(\alpha)$ die echte Schnitteigenschaft von $(\alpha_1, \beta_1, E_\varrho(\alpha_1, \beta_1))$ und $(\alpha_2, \beta_2, z_2) \in (7.7)$, wenn $(\alpha_1, \beta_1) \in \operatorname{conv} G_4$ und wenn $\frac{1}{3}\beta_1 + \frac{2}{3}\beta_2 \geqq \frac{\pi}{3}$ und $\beta_1 \neq 0$.

5. Schritt. $(\alpha_1, \beta_1, E_\varrho(\alpha_1, \beta_1))$ und $(\alpha_2, \beta_2, h_\varrho(\alpha_2, \beta_2))$ haben die Schnitteigenschaft, wenn $(\alpha_1, \beta_1) \in \operatorname{conv} G_4$, $\alpha_2 + \beta_2 \geqq \frac{\pi}{2}$ und $\frac{1}{3}\beta_1 + \frac{2}{3}\beta_2 \geqq \frac{\pi}{3}$. Sie haben sogar die echte Schnitteigenschaft, wenn nicht gleichzeitig $(\alpha_1, \beta_1) = (\frac{\pi}{2}, 0)$ und $(\alpha_2, \beta_2) = (0, \frac{\pi}{2})$.

Beweis. Zunächst sei $\beta_1 \neq 0$. Zu dem gegebenen Punkt $(\alpha_2, \beta_2, h_\varrho(\alpha_2, \beta_2))$ betrachte man den Punkt der Geraden (7.7) mit derselben β-Koordinate β_2; dies ist ein Punkt $P(\beta_2)$. Wegen des 4. Schritts haben $(\alpha_1, \beta_1, E_\varrho(\alpha_1, \beta_1))$ und $P(\beta_2)$ die echte Schnitteigenschaft. Nach Hilfssatz 7.2 liegt aber der Punkt $(\alpha_2, \beta_2, h_\varrho(\alpha_2, \beta_2))$ unterhalb der Ebenen (7.5) und (7.6) mit Steigung $m = 1 - \cos\varrho$ bzw. $m = 1 + \cos\varrho \geqq 1$ in α-Richtung, deren Schnittgerade (7.7) ist. Daraus folgt nach (11.8), daß das gegebene Punktepaar die echte Schnitteigenschaft hat, wenn $\beta_1 \neq 0$.

Ist aber $\beta_1 = 0$, so folgt aus den Voraussetzungen, daß $\beta_2 = \frac{\pi}{2}$. Dann ist $h_\varrho(\alpha_2, \beta_2) = \alpha_2(1 - \cos\varrho)$ nach (7.1). Weil $(\frac{\pi}{2}, 0, 0)$ und $(0, \frac{\pi}{2}, 0)$ die Schnitteigenschaft haben, folgt also nach (11.8), daß das gegebene Punktepaar auch im Fall $\beta_1 = 0$ die echte Schnitteigenschaft hat, wenn $\alpha_2 > 0$.

Deshalb hat das gegebene Punktepaar die echte Schnitteigenschaft, wenn nicht gleichzeitig $\beta_1 = 0$, d.h. $(\alpha_1, \beta_1) = (\frac{\pi}{2}, 0)$, und $(\alpha_2, \beta_2) = (0, \frac{\pi}{2})$ gilt. Weil aber in diesem Fall trivialerweise die Schnitteigenschaft gilt, ist die Behauptung richtig.

Die Aussage des 5. Schritts ist identisch mit (11.7), womit Hilfssatz 11.1 gezeigt ist. □

12. Beweis von Satz 2

Nach allen diesen Vorbereitungen kann nun Satz 2 bewiesen werden.

Es seien also k Punkte (α_i, β_i) mit $\frac{\sum_{i=1}^{k} \beta_i}{k} \leqq \frac{\pi}{3}$ gegeben, und für ganze Zahlen $\mu, \nu \geqq 0$ mit $\nu \leqq 2\mu$ gelte (5.2), ..., (5.5).

Dann folgt aus (5.2) wegen (10.7) und (10.6)

$$\begin{gathered}\sum_{i=1}^{\mu} e_\varrho(\alpha_i, \beta_i) \leqq \sum_{i=1}^{\mu} E_\varrho(\alpha_i, \beta_i) \leqq \mu E_\varrho(\overline{\alpha}_1, \overline{\beta}_1) \\ \text{mit } \overline{\alpha}_1 := \frac{\sum_{i=1}^{\mu} \alpha_i}{\mu}, \quad \overline{\beta}_1 := \frac{\sum_{i=1}^{\mu} \beta_i}{\mu},\end{gathered} \tag{12.1}$$

so daß $(\overline{\alpha}_1, \overline{\beta}_1) \in \operatorname{conv} G_4$.

Aus (5.3) folgt wegen Hilfssatz 7.1

$$\begin{gathered}\sum_{i=\mu+1}^{\mu+\nu} e_\varrho(\alpha_i, \beta_i) = \sum_{i=\mu+1}^{\mu+\nu} h_\varrho(\alpha_i, \beta_i) \leqq \nu h_\varrho(\overline{\alpha}_2, \overline{\beta}_2) \\ \text{mit } \overline{\alpha}_2 := \frac{\sum_{i=\mu+1}^{\mu+\nu} \alpha_i}{\nu}, \quad \overline{\beta}_2 := \frac{\sum_{i=\mu+1}^{\mu+\nu} \beta_i}{\nu},\end{gathered} \tag{12.2}$$

so daß $\overline{\alpha}_2 + \overline{\beta}_2 \geqq \frac{\pi}{2}$.

(5.5) is äquivalent zu

$$\frac{\mu\overline{\beta}_1 + \nu\overline{\beta}_2}{\mu + \nu} \geqq \frac{\pi}{3}. \tag{12.3}$$

Wegen (5.2) und (12.3) ist entweder $\mu = \nu = 0$ oder $\mu, \nu \neq 0$. Für $\mu, \nu \neq 0$ folgt aus (12.1), (12.2) und (12.3) nach Hilfssatz 11.1

$$(12.4)\qquad \begin{aligned} &\sum_{i=1}^{\mu+\nu} e_\varrho(\alpha_i, \beta_i) \leqq \mu E_\varrho(\overline{\alpha}_1, \overline{\beta}_1) + \nu h_\varrho(\overline{\alpha}_2, \overline{\beta}_2) \leqq \\ &\leqq (\mu+\nu) H_\varrho\Big(\frac{\mu\overline{\alpha}_1 + \nu\overline{\alpha}_2}{\mu+\nu}, \frac{\mu\overline{\beta}_1 + \nu\overline{\beta}_2}{\mu+\nu}\Big) \qquad (\mu, \nu \neq 0), \end{aligned}$$

wobei das Gleichheitszeichen an der 2. Stelle nur dann gilt, wenn $(\overline{\alpha}_1, \overline{\beta}_1) = (\frac{\pi}{2}, 0)$ und $(\overline{\alpha}_2, \overline{\beta}_2) = (0, \frac{\pi}{2})$. Dann ist wegen (12.3) $\nu = 2\mu$ und $\frac{\mu\overline{\alpha}_1 + \nu\overline{\alpha}_2}{\mu+\nu} = \frac{\pi}{6}$, $\frac{\mu\overline{\beta}_1 + \nu\overline{\beta}_2}{\mu+\nu} = \frac{\pi}{3}$.

Aus (5.4) folgt wegen $e_\varrho(\alpha, \beta) = h_\varrho(\alpha, \beta)$ in $G_1 \cup G_2$ nach Hilfssatz 7.1 und wegen (9.3) und (9.4)

$$(12.5)\qquad \sum_{i=\mu+\nu+1}^{k} e_\varrho(\alpha_i, \beta_i) \leqq \sum_{i=\mu+\nu+1}^{k} H_\varrho(\alpha_i, \beta_i),$$

wobei das Gleichheitszeichen genau dann gilt, wenn $\beta_i = \frac{\pi}{3}$ und (α_i, ϱ) ein Punkt der Kurve (4.4) ist für jedes $i = \mu + \nu + 1, \dots, k$.

Nun folgt wegen der Konkavität von $H_\varrho(\alpha, \beta)$ nach (9.1)

$$(12.6)\qquad \begin{aligned} &(\mu+\nu) H_\varrho\Big(\frac{\mu\overline{\alpha}_1 + \nu\overline{\alpha}_2}{\mu+\nu}, \frac{\mu\overline{\beta}_1 + \nu\overline{\beta}_2}{\mu+\nu}\Big) + \sum_{i=\mu+\nu+1}^{k} H_\varrho(\alpha_i, \beta_i) \leqq \\ &\leqq k H_\varrho\left(\frac{\sum\limits_{i=1}^{k} \alpha_i}{k}, \frac{\sum\limits_{i=1}^{k} \beta_i}{k}\right). \end{aligned}$$

Weil $H_\varrho(\alpha, \beta)$ in β nicht fallend ist nach (9.1), wegen $\frac{\sum_{i=1}^{k} \beta_i}{k} \leqq \frac{\pi}{3}$ nach Voraussetzung und wegen (9.2) gilt weiter

$$H_\varrho\left(\frac{\sum\limits_{i=1}^{k} \alpha_i}{k}, \frac{\sum\limits_{i=1}^{k} \beta_i}{k}\right) \leqq H_\varrho\left(\frac{\sum\limits_{i=1}^{k} \alpha_i}{k}, \frac{\pi}{3}\right) = F\left(\frac{\sum\limits_{i=1}^{k} \alpha_i}{k}\right),$$

womit (5.6) gezeigt ist.

Gilt in (5.6) das Gleichheitszeichen, so auch in (12.4), (12.5) und (12.6). Dann wird aus (12.6) die Gleichung

$$(12.7)\quad (\mu+\nu)F\Big(\frac{\pi}{6}\Big)+\sum_{i=\mu+\nu+1}^{k} F(\alpha_i)=kF\left(\frac{\sum_{i=1}^{k}\alpha_i}{k}\right), \quad (\mu=\nu=0 \text{ oder } \mu,\nu\neq 0),$$

wobei (α_i,ϱ) ein Punkt der Kurve (4.4) ist für $i=\mu+\nu+1,\ldots,k$.

Es ist $\alpha_i > \frac{\pi}{6}$ für $i=\mu+\nu+1,\ldots,k$, so daß (12.7) nur für $\mu=\nu=0$ möglich ist. Im Fall $\mu=\nu=0$ aber ist die Gültigkeit des Gleichheitszeichens in (12.5) äquivalent zu (5.7). □

II. ZERLEGUNGEN DER SPHÄRE

13. Vorbereitungen

Satz 1 wird nun mit Hilfe von geeigneten Zerlegungen der Sphäre bewiesen. Vom Bisherigen wird benötigt

- die Definitionen bzgl. g-Dreiecken in **2.**,
- die Aufteilung von G in Bereiche G_1, G_2, G_3, G_4 gemäß **3.**, vgl. Fig. 1,
- die Kurve (4.4) samt der geometrischen Interpretation der entsprechenden Wertepaare (α,ϱ),
- Satz 2.

Man beachte, daß zu jedem festen ϱ die Punkte aus $G=G(\varrho)$ und die g-Dreiecke bzgl. ϱ sich bijektiv entsprechen.

Bei einem g-Dreieck wird der Basiswinkel immer mit $\alpha_{\ldots}$ bezeichnet, der halbe Winkel an der Spitze immer mit $\beta_{\ldots}$. Dementsprechend ist der *α-Winkel* eines g-Dreiecks sein Basiswinkel, der *β-Winkel* sein halber Winkel an der Spitze.

Entsprechend den Voraussetzungen von Satz 1 sei also auf der Sphäre ein Kreissystem aus n kongruenten, abgeschlossenen sphärischen Kreisen $K_1,\ldots,K_n$ mit Radius ϱ gegeben. Diese Kreise heißen im folgenden auch *Systemkreise*. Ihre Mittelpunkte seien $O_1,\ldots,O_n$. Nach (1.4), (1.5) und (1.6) kann angenommen werden, daß

$$(13.1)\qquad n \geqq 4,$$

$$(13.2)\qquad 0<\varrho<\varrho_0=\arccos\frac{1}{\sqrt{7}},$$

und

(13.3) die Systemkreismittelpunkte $O_1,\ldots,O_n$ liegen nicht alle in einer abgeschlossenen Halbsphäre.

$E(K_1,\ldots,K_n)$ sei wieder der von $K_1,\ldots,K_n$ einfach überdeckte Bereich der Sphäre S^2.

14. Die $\mathcal{L}$-Zerlegung

Man betrachte einen Systemkreis K_i. Zu jedem K_j $(j \neq i)$ sei G_{ij} der Großkreis nicht durch O_i und O_j, bzgl. dem K_i und K_j symmetrisch liegen. H_{ij} sei die abgeschlossene Halbsphäre, die von G_{ij} begrenzt wird und O_i enthält. Es ist $D_i := \bigcap_{j \neq i} H_{ij}$ die *Dirichletsche Zelle* von K_i. D_i ist sphärisch konvex, weil nach (13.3) nicht alle O_i auf einem Großkreis liegen. Es ist

$$D_i = \{X \mid |\overline{XO_i}| \leqq |\overline{XO_j}|,\ 1 \leqq j \leqq n\}.$$

Die Familie $\{D_i\}_{i=1}^{n}$ pflastert die Sphäre S^2.

Ein Kreis auf S^2 mit Radius $< \pi$ heißt wie üblich *Stützkreis*, wenn in seinem Inneren kein, auf seinem Rand aber mindestens drei Systemkreismittelpunkte liegen. Man sieht sofort, daß die Stützkreismittelpunkte genau die Ecken von Dirichletschen Zellen sind. Wegen (13.3) ist der Radius jedes Stützkreises $< \frac{\pi}{2}$. Die auf dem Rand eines Stützkreises liegenden Systemkreismittelpunkte spannen also ein sphärisch konvexes Polygon auf. Aus der Konstruktion folgt, daß die Familie $\mathcal{L} = \{L_j\}$ dieser Polygone die S^2 pflastert.

Verschiebt man jeden Kreismittelpunkt O_i um höchstens ε, so verkleinert sich $|E(K_1, \ldots, K_n)|$ um höchstens $n2\pi\varepsilon$. Deshalb kann o.B.d.A. angenommen werden, daß $\mathcal{L}$ eine *Dreiecksfamilie* ist. Außerdem kann angenommen werden

(14.1) *Der Umkreismittelpunkt eines $\mathcal{L}$-Dreiecks L liegt nicht auf dem Rand von L.*

Nun sei $\mathcal{Z}$ eine Zerlegung der Indexmenge von $\mathcal{L}$. Zu $z \in \mathcal{Z}$ sei $\mathcal{L}_z := \{L_j \mid j \in z\}$ und $S_z := \bigcup_{j \in z} L_j$; die Familie $\{S_z\}_{z \in \mathcal{Z}}$ pflastert also S^2. Es gilt

HILFSSATZ 14.1. *Gegeben sei eine Zerlegung $\mathcal{Z}$ der Indexmenge von $\mathcal{L}$. Zu $z \in \mathcal{Z}$ sei $\mathcal{L}_z$ und S_z wie oben definiert, und $\mathcal{L}_z$ bestehe aus k_z $\mathcal{L}$-Dreiecken. Ein $\mathcal{L}$-Dreieck L_j habe die Winkelsumme ω_{L_j}. Dann ist die Ungleichung* (1.2) *in Satz 1 richtig, wenn für jedes $z \in \mathcal{Z}$ gilt*

$$(14.2) \qquad |E(K_1, \ldots, K_n) \cap S_z| \leqq 6k_z F\left(\frac{\sum_{j \in z} \omega_{L_j}}{6k_z}\right).$$

Gilt (14.2), *so gilt das Gleichheitszeichen in* (1.2) *nur dann, wenn das Gleichheitszeichen in* (14.2) *gilt.*

BEWEIS. Weil n Systemkreise gegeben sind, ist die Summe der Winkelsummen aller $\mathcal{L}$-Dreiecke gleich $n2\pi$, d.h.

$$\sum_{z \in \mathcal{Z}} \sum_{j \in z} \omega_{L_j} = n2\pi.$$

$\mathcal{L}$ besteht aus $k := \sum_{z\in\mathcal{Z}} k_z$ $\mathcal{L}$-Dreiecken. Weil diese die S^2 pflastern, gilt nach dem Eulerschen Polyedersatz

$$k = 2(n-2).$$

Außerdem pflastert $\{S_z\}_{z\in\mathcal{Z}}$ die S^2, und die Funktion F ist konkav.

Aus der Gültigkeit von (14.2) folgt deshalb

$$|E(K_1,\dots,K_n)| = \sum_{z\in\mathcal{Z}} |E(K_1,\dots,K_n)\cap S_z| \leqq \sum_{z\in\mathcal{Z}} 6k_z F\left(\frac{\sum_{j\in z}\omega_{L_j}}{6k_z}\right) \leqq$$

(14.3)

$$\leqq 6kF\left(\frac{\sum_{z\in\mathcal{Z}}\sum_{j\in z}\omega_{L_j}}{6k}\right) = 12(n-2)F\left(\frac{n}{n-2}\frac{\pi}{6}\right),$$

d.h. die Gültigkeit von (1.2). Aus (14.3) folgt auch die Behauptung bzgl. des Gleichheitszeichens in (1.2). □

Zum Beweis von (1.2) ist also (14.2) zu zeigen für eine geeignete Zerlegung der Indexmenge von $\mathcal{L}$. Eine solche Zerlegung erhält man folgendermaßen:

15. Die Graphen Γ, die Familien $\mathcal{L}_\Gamma$ und die Bereiche S_Γ

Man betrachte die Dreiecksfamilie $\mathcal{L} = \{L_j\}$. Die Umkreismittelpunkte $\{l_j\}$ von $\{L_j\}$ sind die Stützkreismittelpunkte, so daß für den Umkreisradius $U(L_j)$ von L_j gilt, daß $U(L_j) < \frac{\pi}{2}$.

Zu einem $\mathcal{L}$-Dreieck $L_{\dots}$ ist $l_{\dots}$ sein Umkreismittelpunkt und umgekehrt, wobei die Indizierungen übereinstimmen.

Eine Seite s von $L\in\mathcal{L}$ heißt *trennende Seite* von L,[4] wenn gilt

(15.1) Der Großkreis durch s trennt echt l von L,

(15.2) $$|s| > \varrho,$$

und

(15.3) $$U(L) \geqq \varrho.$$

Um nun die Indexmenge von $\mathcal{L}$ geeignet zu zerlegen, wird ein gerichteter Graph Γ_S definiert: Seine Knotenmenge sei $\{l_j\}$. Zwei Knoten l_{j_1} und l_{j_2}

---

[4] Vgl. [8] und [9]. Hier wird aber zusätzlich (15.2) und (15.3) verlangt.

werden genau dann durch eine von l_{j_1} nach l_{j_2} gerichtete Kante (l_{j_1}, l_{j_2}) verbunden, wenn L_{j_1} und L_{j_2} eine gemeinsame Seite haben, die trennende Seite von L_{j_1} ist.

Γ_S hat folgende Eigenschaften:

(15.4) *Die Ausgangsvalenz eines Knotens von Γ_S ist höchstens* 1 (nach Definition).

(15.5) *Ist (l_{j_1}, l_{j_2}) gerichtete Kante von Γ_S, dann nicht (l_{j_2}, l_{j_1});*

denn sonst würde der Stützkreis um l_{j_2} eine Ecke von L_{j_1} im Innern enthalten, d.h. einen Systemkreismittelpunkt im Widerspruch zur Definition eines Stützkreises. □

Wenn es in Γ_S eine Bahn von l nach l' gibt, folgt l' auf l, liegt l vor l' und l' *nach* l. Mit diesen Bezeichnungen gilt weiter

(15.6) *Folgt l' auf l, so ist $U(L') > U(L)$* (folgt unmittelbar).

(15.7) *Γ_S ist zyklenfrei;*

denn wegen (15.4) wäre jeder Zyklus ein Kreis. Zwei seiner Knoten l und l' folgten also jeweils aufeinander. Dann wäre aber $U(L') > U(L) > U(L')$. □

Nun sei Γ eine Zusammenhangskomponente von Γ_S. Die den Knoten von Γ entsprechenden $\mathcal{L}$-Dreiecke bilden die Familie $\mathcal{L}_\Gamma$, und es sei $S_\Gamma := = \bigcup_{L \in \mathcal{L}_\Gamma} L = \bigcup_{l \in \Gamma} L$ (vgl. **14.**). Aus Hilfssatz 14.1 folgt sofort

HILFSSATZ 15.1. *Zu einer Zusammenhangskomponente Γ von Γ_S sei $\mathcal{L}_\Gamma$ und S_Γ wie oben definiert, und $\mathcal{L}_\Gamma$ bestehe aus k_Γ $\mathcal{L}$-Dreiecken. Ein $\mathcal{L}$-Dreieck L habe die Winkelsumme ω_L. Dann ist die Ungleichung* (1.2) *in Satz 1 richtig, wenn für jedes Γ gilt*

$$(15.8) \qquad |E(K_1, \dots, K_n) \cap S_\Gamma| \leqq 6k_\Gamma F\left(\frac{\sum\limits_{l\in\Gamma} \omega_L}{6k_\Gamma}\right).$$

Gilt (15.8), *so gilt das Gleichheitszeichen in* (1.2) *nur dann, wenn das Gleichheitszeichen in* (15.8) *gilt.*

16. Die Zusammenhangskomponenten Γ, bei denen $\mathcal{L}_\Gamma$ ein Dreieck L enthält mit $|s| \leqq \varrho$ für jede Seite s von L

HILFSSATZ 16.1. *Sei Γ eine Zusammenhangskomponente von Γ_S so, daß $\mathcal{L}_\Gamma$ ein Dreieck L enthält mit $|s| \leqq \varrho$ für jede Seite s von L. Dann ist* (15.8) *richtig mit dem $<$-Zeichen.*

BEWEIS. Es sei also $L \in \mathcal{L}_\Gamma$ so, daß für jede Seite s von L gilt $|s| \leqq \varrho$. Wegen (15.2) ist dann l Senke von Γ. Wenn es eine Kante (l', l) in Γ gibt, so hat L' eine trennende Seite s' mit $|s'| > \varrho$ und s' ist auch Seite von L im Widerspruch zur Voraussetzung. Deshalb ist $S_\Gamma = \{L\}$. Weil L von jedem Systemkreis um seine Ecken überdeckt wird, gilt $E(K_1, \ldots, K_n) \cap L = \emptyset$. □

17. Die Zusammenhangskomponenten Γ, bei denen $\mathcal{L}_\Gamma$ ein Dreieck L enthält mit $U(L) < \varrho$

HILFSSATZ 17.1. *Sei Γ eine Zusammenhangskomponente von Γ_S so, daß $\mathcal{L}_\Gamma$ ein Dreieck L enthält mit $U(L) < \varrho$. Dann ist* (15.8) *richtig mit dem $<$-Zeichen.*

BEWEIS. Es sei also $L \in \mathcal{L}_\Gamma$ so, daß $U(L) < \varrho$. Nach (15.3) ist dann l Senke von Γ. Wenn es eine Kante (l', l) in Γ gibt, so ist $U(L') \geqq \varrho$ und nach (15.6) ist $U(L) > U(L')$ im Widerspruch zu $U(L) < \varrho$. Deshalb ist $S_\Gamma = \{L\}$.

O.B.d.A. sei $L = O_1O_2O_3$. Weil $L \subset K_1 \cup K_2 \cup K_3$, ist $E(K_1, \ldots, K_n) \cap \cap L \subset E(K_1, K_2, K_3) \cap L$. $|E(K_1, K_2, K_3) \cap L|$ wird echt vergrößert, wenn ϱ verkleinert wird bis zu $\varrho = U(L)$. Deshalb wird jetzt $|E(K_1, K_2, K_3) \cap L|$ nach oben abgeschätzt für $\varrho = U(L)$. Man beachte, daß nach (14.1) gilt $l \notin \partial L$.

Gilt $l \in L^0$, so zerlegen die Strecken $\overline{lO_i}$ das Dreieck L in drei g-Dreiecke Δ_m $(m = 1, 2, 3)$ mit den α-Winkeln α_m und den β-Winkeln β_m (vgl. **13.**). Es ist $\beta_1 + \beta_2 + \beta_3 = \pi$. Sind alle g-Dreiecke Δ_m breit, so erfüllt also die Punktmenge $\{(\alpha_m, \beta_m)\}_{m=1}^3$ die Voraussetzungen von Satz 2, mit $\mu = 0$, $\nu = 0$. Ist aber ein g-Dreieck schmal, so ist sein β-Winkel $< \frac{\pi}{4}$, d.h. höchstens eines der g-Dreiecke ist schmal, z.B. Δ_1. Weil dann $\beta_2 \geqq \frac{\pi}{4}$, $\beta_3 \geqq \frac{\pi}{4}$, gilt $(\alpha_2, \beta_2), (\alpha_3, \beta_3) \in G_1 \cup G_2$. Die Punktmenge $\{(\alpha_m, \beta_m)\}_{m=1}^3$ erfüllt also ebenfalls die Voraussetzungen von Satz 2, mit $\mu = 1$, $\nu = 2$.

Δ_m wird von den beiden Systemkreisen um seine Basisecken mit Radius $\varrho = U(L)$ überdeckt. Nach (2.4) und (2.5) ist deshalb $|E(K_1, K_2, K_3) \cap L| = \sum_m |E(K_1, K_2, K_3) \cap \Delta_m| \leqq \sum_m 2e_\varrho(\alpha_m, \beta_m)$, und nach Satz 2 ist

$$2\sum_m e_\varrho(\alpha_m, \beta_m) \leqq 6F\left(\frac{\alpha_1 + \alpha_2 + \alpha_3}{3}\right).$$

(15.8) gilt also im Fall $l \in L^0$.

Ist aber $l \notin L$, so sei s die Kante von L so, daß der Großkreis durch s echt l von L trennt. O_3 sei die s gegenüberliegende Ecke von L. Das Lot von O_3 aus auf s habe den Fußpunkt F; es zerlegt L in zwei Dreiecke Δ_m $(m = 1, 2)$ mit $O_m \in \Delta_m$.

$|E(K_1, K_2, K_3) \cap \Delta_1|$ wird folgendermaßen abgeschätzt: Das Lot von l aus auf $\overline{O_1O_3}$ schneidet den Rand von Δ_1 außer in $\overline{O_1O_3}$ im Punkt S_1, wobei $S_1 \in s$ oder $S_1 \in \overline{FO_3}$ (siehe Fig. 8 und 9). Es sei $\varrho_1 := |\overline{S_1O_3}| < \varrho$. Dann ist $S_1O_1O_3$ ein g-Dreieck bzgl. ϱ_1 und $\Delta_1 \setminus S_1O_1O_3$ kann als die Hälfte eines g-Dreiecks bzgl. ϱ_1 mit Spitze S_1 aufgefaßt werden (das für $S_1 = F$ entartet ist).

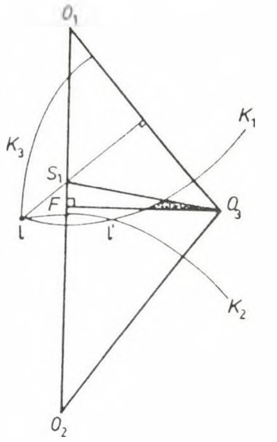


Fig. 8

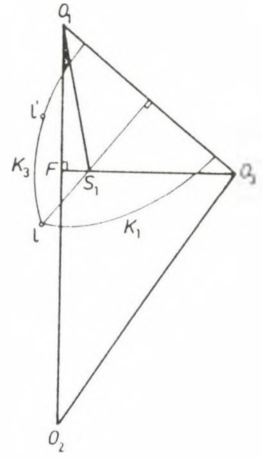


Fig. 9

Es ist

$$E(K_1, K_2, K_3) \cap S_1O_1O_3 \subset E(K_1, K_3) \cap S_1O_1O_3 \subset \\ \subset E(K_{\varrho_1}(O_1), K_{\varrho_1}(O_3)) \cap S_1O_1O_3.$$

Um $|E(K_1,K_2,K_3)\cap(\Delta_1\setminus S_1O_1O_3)|$ abzuschätzen, betrachte man zunächst den Fall $S_1 \in s$. Dann ist $\Delta_1 \setminus S_1O_1O_3 = S_1FO_3 \subset K_3$, und abzuschätzen bleibt $|S_1FO_3\setminus(K_1\cup K_2)|$. Seien O_3' und l' das Spiegelbild von O_3 und l bzgl. s. Es ist $l' \in \partial K_1 \cap \partial K_2$ und $l' \in \partial K_\varrho(O_3')$. Deshalb ist $S_1FO_3\setminus(K_1 \cup K_2) \subset$ $\subset S_1FO_3 \setminus K_\varrho(O_3') \subset S_1FO_3 \setminus K_{\varrho_1}(O_3') = E(K_{\varrho_1}(O_3), K_{\varrho_1}(O_3')) \cap S_1FO_3$.

Ist aber $S_1 \in \overline{FO}_3$, so ist $\Delta_1 \setminus S_1O_1O_3 = S_1FO_1 \subset K_1$, und abzuschätzen bleibt $|S_1FO_1 \setminus K_3|$. Seien O_1' und l' das Spiegelbild von O_1 und l bzgl. O_3F. Es ist $l' \in \partial K_3$ und $l' \in \partial K_\varrho(O_1')$. Deshalb ist $S_1FO_1 \setminus K_3 \subset S_1FO_1 \setminus$ $\setminus K_\varrho(O_1') \subset S_1FO_1 \setminus K_{\varrho_1}(O_1') = E(K_{\varrho_1}(O_1), K_{\varrho_1}(O_1')) \cap S_1FO_1$.

Nun seien α_1, β_1 der α- und β-Winkel von $S_1O_1O_3$, und α_2, β_2 seien der α- und β-Winkel von $S_1O_3O_3'$ bzw. von $S_1O_1O_1'$. Dann ist $\beta_1 + \beta_1 +$ $+ \beta_2 = \pi$, und zu ϱ_1 erfüllt die Punktmenge $\{(\alpha_1, \beta_1), (\alpha_1, \beta_1), (\alpha_2, \beta_2)\}$ die Voraussetzungen von Satz 2. Deshalb gilt nach (2.4), (2.5) und Satz 2, daß $|E(K_1, K_2, K_3) \cap \Delta_1| = |E(K_1, K_2, K_3) \cap S_1O_1O_3| + |E(K_1, K_2, K_3) \cap (\Delta_1 \setminus$ $\setminus S_1O_1O_3)| \leqq 2e_{\varrho_1}(\alpha_1, \beta_1) + e_{\varrho_1}(\alpha_2, \beta_2) \leqq 3F(\frac{\alpha_1+\alpha_1+\alpha_2}{3})$.

Weil für $|E(K_1, K_2, K_3)\cap\Delta_2|$ eine analoge Abschätzung gilt, folgt aus der Konkavität von F, daß $|E(K_1, K_2, K_3) \cap L| \leqq 6F(\frac{\omega_L}{6})$, wenn ω_L die Winkelsumme von L ist. (15.8) gilt also auch im Fall $l \notin L$.

Weil zu Beginn des Beweises das gegebene ϱ verkleinert wurde zu $\varrho =$ $= U(L)$ und dadurch $|E(K_1, K_2, K_3)\cap L|$ echt vergrößert wurde, gilt in (15.8) sogar das $<$-Zeichen. □

18. Die Zusammenhangskomponenten Γ, bei denen für jedes Dreieck $L \in \mathcal{L}_\Gamma$ gilt, daß L eine Seite s mit $|s| > \varrho$ besitzt, und daß $U(L) \geqq \varrho$

18.1. *Geknickte g-Dreiecke*

In Anlehnung an **2.** und **13.** wird folgende Begriffsbildung zweckmäßig sein: Auf der Sphäre sei ein Viereck $C_1A_1C_2A_2$ so gegeben, daß es von $\overline{C_1C_2}$ in zwei kongruente Dreiecke zerlegt wird, daß der Innenwinkel bei C_2 größer als π sei, und daß $|\overline{A_1C_2}| \geqq \varrho$ und $|\overline{A_1C_1}| \leqq \frac{\pi}{2}$. Ein solches Viereck heiße *geknicktes g-Dreieck* (*bzgl.* ϱ). A_1 und A_2 heißen *Basisecken*, C_1 *Spitze*, C_2 *Gegenspitze* und $\overline{C_1C_2}$ *Diagonale* des geknickten g-Dreiecks. Der Winkel bei A_1 oder A_2 heiße *Basiswinkel.*

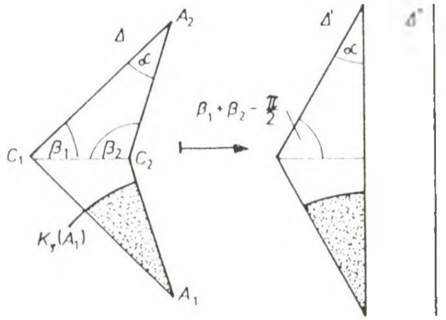


Fig. 10

Auch der Basiswinkel eines geknickten g-Dreiecks wird immer mit $\alpha_{\dots}$ bezeichnet, so daß der α-*Winkel* eines geknickten g-Dreiecks sein Basiswinkel sei. Der halbe Innenwinkel in der Spitze und der halbe Innenwinkel in der Gegenspitze eines geknickten g-Dreiecks wird immer mit $\beta_{\dots}$ bezeichnet; jeder solche Winkel heißt β-*Winkel*, so daß ein geknicktes g-Dreieck *zwei* β-Winkel besitzt.

Analog zu (2.4) sei für ein geknicktes g-Dreieck mit den Basisecken A_1, A_2

(18.1) $\qquad e_\varrho(\Delta) := |E(K_\varrho(A_1), K_\varrho(A_2)) \cap \Delta|$ für geknickte g-Dreiecke.

Jedem geknickten g-Dreieck lassen sich folgendermaßen zwei breite g-Dreiecke zuordnen.

HILFSSATZ 18.1. *Zu jedem geknicktem g-Dreieck Δ gibt es ein breites g-Dreieck Δ' und ein breites entartetes g-Dreieck Δ'' mit*

Die Summe der α-Winkel von Δ' und Δ'' ist gleich dem α-Winkel von Δ.

(18.2) $\qquad$ *Die Summe der β-Winkel von Δ' und Δ'' ist gleich* *der Summe der β-Winkel von Δ.*

(18.3) *Der β-Winkel von Δ' ist größer als der β-Winkel von Δ an der Spitze.*

(18.4) *Δ' und Δ'' entsprechen Punkten von G_1 (vgl. Fig. 1).*

(18.5) $$e_\varrho(\Delta) = e_\varrho(\Delta') = e_\varrho(\Delta') + e_\varrho(\Delta'').$$

BEWEIS (siehe Fig. 10). Sei Δ ein geknicktes g-Dreieck mit α-Winkel α und den beiden β-Winkeln β_1 und β_2. A_1 sei eine Basisecke und C_2 die Gegenspitze von Δ. Δ' sei das gleichschenklige Dreieck mit Basiswinkel α und mit $|\Delta'| = |\Delta|$. Weil bei einem geknickten g-Dreieck $|\overline{A_1 C_2}| \geqq \varrho$ gilt, ist Δ' ein breites g-Dreieck, das einem Punkt aus G_1 entspricht. Weil Δ' und Δ denselben α-Winkel haben, ist also $e_\varrho(\Delta') = e_\varrho(\Delta)$. Wegen $|\Delta'| = |\Delta|$ ist der β-Winkel von Δ' gleich $\beta_1 + \beta_2 - \frac{\pi}{2}$; daraus folgt (18.3).

Δ'' sei ein entartetes breites g-Dreieck, d.h. mit α-Winkel 0 und β-Winkel $\frac{\pi}{2}$. Es ist $e_\varrho(\Delta'') = 0$.

Δ' und Δ'' erfüllen dann zusammen die Behauptung von Hilfssatz 18.1. □

18.2. *Die Polygonfamilie $\mathcal{D}_\Gamma$*

Nun werden also die Zusammenhangskomponenten Γ von Γ_S betrachtet, bei denen für jedes Dreieck $L \in \mathcal{L}_\Gamma$ gilt, daß L eine Seite s mit $|s| > \varrho$ besitzt, und daß $U(L) \geqq \varrho$ ist. Zunächst bemerken wir

HILFSSATZ 18.2. *Jede Zusammenhangskomponente Γ von Γ_S ist ein Baum mit genau einer Senke, und jede Quelle von Γ wird mit der Senke durch eine Bahn verbunden.*

BEWEIS. Dies folgt aus (15.7), (15.4) und (15.5). □

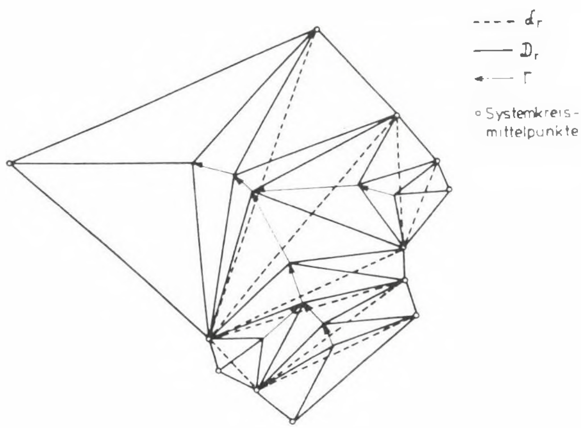


Fig. 11

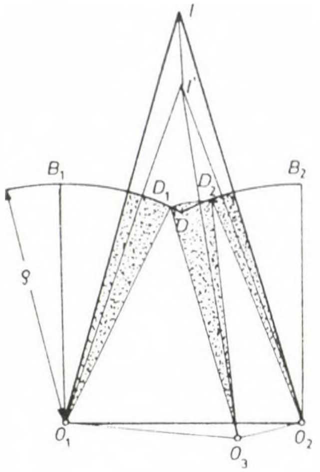


Fig. 12

Zu jeder Quelle von Γ betrachte man die sie mit der Senke verbindende Bahn und durchlaufe sie rückwärts, d.h. von der Senke ausgehend. Dabei verbinde man jeden Knoten von Γ mit den Ecken des zugehörigen $\mathcal{L}_\Gamma$-Dreiecks (wegen (14.1) liegt kein Knoten von Γ auf dem Rand des zugehörigen $\mathcal{L}_\Gamma$-Dreiecks). Dann entsteht zusammen mit den nicht trennenden Seiten von $\mathcal{L}_\Gamma$-Dreiecken eine Polygonfamilie $\mathcal{D}_\Gamma$, die S_Γ pflastert, siehe Fig. 11. Die Polygone von $\mathcal{D}_\Gamma$ sind entweder gleichschenklige Dreiecke — die Basis ist dann eine nicht trennende Seite von $\mathcal{L}_\Gamma$, die zwei Basisecken sind Systemkreismittelpunkte und die Spitze ist ein Knoten von Γ — oder Vierecke — zwei Ecken des Vierecks sind Systemkreismittelpunkte und Ecken einer trennenden Seite von $\mathcal{L}_\Gamma$, die anderen beiden Ecken sind Knoten von Γ und die ihnen zugeordnete Kante von Γ zerlegt das Viereck in zwei kongruente Dreiecke und weist von der Ecke mit Innenwinkel $> \pi$ zur Ecke mit Innenwinkel $< \pi$. Weil der Radius jedes Stützkreises $< \frac{\pi}{2}$ ist nach (13.3), und weil $U(L) \geqq \varrho$ nach Voraussetzung, sind die Dreiecke von $\mathcal{D}_\Gamma$ g-Dreiecke und die Vierecke geknickte g-Dreiecke. $\mathcal{D}_\Gamma$ ist also eine Familie aus g-Dreiecken. Aus allem diesem folgt unmittelbar

Hilfssatz 18.3. *Sei Γ eine Zusammenhangskomponente von Γ_S so, daß für jedes Dreieck $L \in \mathcal{L}_\Gamma$ gilt, daß L eine Seite s mit $|s| > \varrho$ besitzt, und daß $U(L) \geqq \varrho$. Dann läßt sich S_Γ in eine Familie $\mathcal{D}_\Gamma$ von g-Dreiecken zerlegen (Fig. 11), und es ist*

(18.6) $$|E(K_1, \ldots, K_n) \cap S_\Gamma| = \sum_{\Delta \in \mathcal{D}_\Gamma} |E(K_1, \ldots, K_n) \cap \Delta|. \qquad \square$$

Hilfssatz 18.4. *$\mathcal{L}_\Gamma$ bestehe aus k_Γ $\mathcal{L}$-Dreiecken, und ein $\mathcal{L}$-Dreieck L habe die Winkelsumme ω_L. Dann gilt*

(18.7) *Die Summe der α-Winkel aller $\Delta \in \mathcal{D}_\Gamma$ ist* $\frac{1}{2} \sum_{l \in \Gamma} \omega_L$.

(18.8) *Die Summe der β-Winkel aller $\Delta \in \mathcal{D}_\Gamma$ ist πk_Γ.* □

HILFSSATZ 18.5. *Jeder Knoten von Γ inzidiert mit genau 3 g-Dreiecken von $\mathcal{D}_\Gamma$. Jedes schmale oder breite g-Dreieck von $\mathcal{D}_\Gamma$ inzidiert mit genau 1 Knoten von Γ, jedes geknickte g-Dreieck inzidiert mit genau 2 Knoten. Die Kanten von Γ sind genau die Diagonalen der geknickten g-Dreiecke.* □

18.3. *Eine Abschätzung von $|E(K_1, \dots, K_n) \cap \Delta|$ für $\Delta \in \mathcal{D}_\Gamma$*

HILFSSATZ 18.6. *Für jedes g-Dreieck $\Delta \in \mathcal{D}_\Gamma$ sei $e_\varrho(\Delta)$ gemäß (2.4) bzw. (18.1) definiert. Dann gilt*

$$|E(K_1, \dots, K_n) \cap \Delta| \leqq e_\varrho(\Delta). \tag{18.9}$$

BEWEIS. Als erstes sei Δ ein geknicktes g-Dreieck von $\mathcal{D}_\Gamma$. Δ wird durch eine Γ-Kante (l', l) in zwei kongruente Dreiecke zerlegt mit den Ecken O_1 bzw. O_2. $\overline{ll'}$ ist die gemeinsame Seite der konvexen Dirichletschen Zellen zu O_1 und O_2. Deshalb ist $\Delta \setminus (K_1 \cup K_2)$ fremd zu jedem Systemkreis, so daß $E(K_1, \dots, K_n) \cap \Delta \subset E(K_1, K_2) \cap \Delta$, und nach Definition von $e_\varrho(\Delta)$ gilt (18.9) für geknickte g-Dreiecke.

Nun sei Δ ein nicht geknicktes g-Dreieck von $\mathcal{D}_\Gamma$. Dann gilt nach Definition von $\mathcal{D}_\Gamma$, daß $\Delta = lO_1O_2$, und $\overline{O_1O_2}$ ist keine trennende Seite von $\mathcal{L}_\Gamma$. Es sei $D := \Delta \cap \partial K_1 \cap \partial K_2$, falls dieser Punkt existiert, sonst sei D der Mittelpunkt von $\overline{O_1O_2}$. Für $D = l$ ist offensichtlich $E(K_1, \dots, K_n) \cap \Delta \subset$ $\subset E(K_1, K_2) \cap \Delta$ und (18.9) gilt nach Definition von $e_\varrho(\Delta)$. Deshalb sei im folgenden $D \neq l$.

Nach Definition von $\mathcal{D}_\Gamma$ ist l Umkreismittelpunkt eines $\mathcal{L}_\Gamma$-Dreiecks, das auf derselben Seite von O_1O_2 liegt wie l. Außerdem gibt es ein $\mathcal{L}$-Dreieck $L' =$ $= O_1O_2O_3$, das nicht auf derselben Seite von O_1O_2 liegt wie l; sein Umkreismittelpunkt sei l'. $\overline{ll'}$ ist die gemeinsame Seite der konvexen Dirichletschen Zellen zu O_1 und O_2. Ist also $\overline{lD} \subset \overline{ll'}$, so ist $\Delta \setminus (K_1 \cup K_2)$ fremd zu jedem Systemkreis, d.h. es ist $E(K_1, \dots, K_n) \cap \Delta \subset E(K_1, K_2) \cap \Delta$ und (18.9) gilt.

Es bleibt der Fall zu betrachten, daß l' echt zwischen l und D liegt.

Als erstes sei Δ breit, d.h. $|\overline{O_1O_2}| > \varrho$. Es gilt $U(L') = |\overline{l'O_1}| > \varrho$, so daß die Seite $\overline{O_1O_2}$ von L' nach (15.1), (15.2) und (15.3) trennende Seite von L' ist. Deshalb ist (l', l) Kante von Γ und $\overline{O_1O_2}$ ist trennende Seite von $\mathcal{L}_\Gamma$ im Widerspruch zu oben. Für breite g-Dreiecke liegt also l' nicht echt zwischen l und D.

Nun sei Δ schmal (siehe Fig. 12). Es gilt $\Delta = lO_1l'O_2 \cup l'O_1O_2$. Weil $\overline{ll'}$ die gemeinsame Seite der konvexen Dirichletschen Zellen von O_1 und O_2 ist, gilt $E(K_1, \dots, K_n) \cap lO_1l'O_2 = (K_1 \cup K_2) \cap lO_1l'O_2$.

Um $|E(K_1, \dots, K_n) \cap l'O_1O_2|$ abzuschätzen, wird $l'O_iO_3$ $(i = 1, 2)$ betrachtet. Es sei $D_i := l'O_iO_3 \cap \partial K_i \cap \partial K_3$. Nun gilt für die Kreisbogen $\overset{\frown}{DD_1}, \overset{\frown}{DD_2}, \overset{\frown}{D_1D_2}$, daß $|\overset{\frown}{D_1D_2}| \leqq |\overset{\frown}{DD_1}| + |\overset{\frown}{DD_2}|$, wie man folgendermaßen einsieht:

Diese Ungleichung ist offensichtlich richtig, wenn DD_1D_2 einen Umkreisradius $U(DD_1D_2) < \varrho$ hat. $U(DD_1D_2) < \varrho$ ist klar, wenn der Umkreismittelpunkt von DD_1D_2 auf derselben Seite von D_1D_2 liegt wie D. Sonst aber sei B_i $(i = 1, 2)$ der Punkt auf dem Lot von O_1O_2 durch O_i, auf derselben Seite von O_1O_2 wie l und im Abstand ϱ zu O_i. Man überlegt, daß $U(DB_1B_2) < \varrho$. Der Umkreis von DB_1B_2 enthält D_1, D_2 im Inneren, so daß auch in diesem Fall $U(DD_1D_2) < \varrho$ ist. Deshalb gilt also $|\overset{\frown}{D_1D_2}| \leqq |\overset{\frown}{DD_1}| + |\overset{\frown}{DD_2}|$.

Deshalb gilt für $l'O_iO_3$, daß $\sum_{i=1,2} |E(K_i, K_3) \cap l'O_iO_3| \leqq \sum_{i=1,2} |(K_i \cup K_3) \cap l'O_iO_3 \setminus D_iO_iO_3| \leqq |(K_1 \cup K_2) \cap l'O_1O_2 \setminus DO_1O_2|$.

Weil l' Mittelpunkt des Stützkreises durch O_1, O_2 und O_3 war, gibt es einen Punkt l_i'' auf $l'D_i$, der ebenfalls Mittelpunkt eines Stützkreises durch O_i und O_3 ist $(i = 1, 2)$. Gilt $\overline{l'D_i} \subset \overline{l'l_i''}$, so ist $l'O_iO_3 \setminus (K_i \cup K_3)$ fremd zu jedem Systemkreis, d.h. es ist $E(K_1, \ldots, K_n) \cap l'O_iO_3 \subset E(K_i, K_3) \cap l'O_iO_3$, so daß $|E(K_1, \ldots, K_n) \cap l'O_1O_2| \leqq \sum_{i=1,2} |E(K_1, \ldots, K_n) \cap l'O_iO_3| \leqq \sum_{i=1,2} |E(K_i, K_3) \cap l'O_iO_3| \leqq |(K_1 \cup K_2) \cap l'O_1O_2 \setminus DO_1O_2|$ und insgesamt $|E(K_1, \ldots, K_n) \cap \Delta| \leqq e_\varrho(\Delta)$. Ansonsten wiederholt man die oberen Überlegungen endlich oft und erhält ebenfalls diese Abschätzung. Deshalb gilt (18.9) auch für schmale Dreiecke. □

Beim Beweis von Hilfssatz 18.6 wurde mitbewiesen

HILFSSATZ 18.7. *Es sei $\Delta \in \mathcal{D}_\Gamma$ ein geknicktes oder breites g-Dreieck, und K_1, K_2 seien die beiden Systemkreise um die Basisecken von Δ. Dann ist $\Delta \setminus (K_1 \cup K_2)$ fremd zu jedem Systemkreis.* □

Wegen (18.6) und (18.9) ist also

$$|E(K_1, \ldots, K_n) \cap S_\Gamma| = \sum_{\Delta \in \mathcal{D}_\Gamma} |E(K_1, \ldots, K_n) \cap \Delta| \leqq \sum_{\Delta \in \mathcal{D}_\Gamma} e_\varrho(\Delta).$$

Um mit Hilfe von (2.5) und Satz 2 die gewünschte Abschätzung zu erhalten, müssen die geknickten g-Dreiecke aus $\mathcal{D}_\Gamma$ eliminiert werden — was prinzipiell mit Hilfssatz 18.1 möglich ist —, vor allem aber müssen die Voraussetzungen von Satz 2 erfüllt sein — was im allgemeinen nicht gegeben ist. Deshalb wird in 18.6 die Dreiecksmenge $\mathcal{D}_\Gamma$ in eine geeignete Dreiecksmenge ε_Γ transformiert werden, mit Hilfe der Verschiebungsoperation V von 18.5 und der Quellenoperation Q von 18.4:

18.4. *Die Quellenoperation Q*

Auf $\mathcal{D}_\Gamma$ wird eine Quellenoperation Q eingeführt. Sie soll bewirken, daß eine Quelle von Γ mit höchstens einem schmalen g-Dreieck inzidiert. Es gilt

Hilfssatz 18.8. *l sei eine Quelle von Γ, die mit zwei schmalen g-Dreiecken $\Delta_1 = lO_1O_3$ und $\Delta_2 = lO_2O_3$ von $\mathcal{D}_\Gamma$ inzidiere. Dann inzidiert l außerdem mit einem geknickten g-Dreieck $\Delta_3 = l'O_1lO_2$ von $\mathcal{D}_\Gamma$.*

Dann gibt es Punkte $Q(l)$ und $Q(O_3)$ so, daß gilt

$$Q(\Delta_1) := Q(l)O_1Q(O_3),\ Q(\Delta_2) := Q(l)O_2Q(O_3) \text{ und } Q(\Delta_3) := l'O_1Q(l)O_2 \text{ sind g-Dreiecke, und } Q(\Delta_1) \text{ ist breit.} \tag{18.10}$$

Q sei die Operation, die in $\mathcal{D}_\Gamma$ die Dreiecke Δ_i durch $Q(\Delta_i)$ ersetzt ($i = 1, 2, 3$).

$Q(l)$ und $Q(O_3)$ können außerdem so gewählt werden, daß weiter gilt

(18.11) *Die Operation Q läßt die Summe der β-Winkel konstant, die Summe der α-Winkel wird höchstens verkleinert.*

$$\sum_{\Delta \in \mathcal{D}_\Gamma} e_\varrho(\Delta) \leqq \sum_{\Delta \in Q(\mathcal{D}_\Gamma)} e_\varrho(\Delta). \tag{18.12}$$

Beweis. Die Quelle l von Γ inzidiere also mit den drei g-Dreiecken $\Delta_1, \Delta_2, \Delta_3 \in \mathcal{D}_\Gamma$, wobei $\Delta_1 = lO_1O_3$ und $\Delta_2 = lO_2O_3$ schmal seien. Deshalb ist der β-Winkel an der Spitze von Δ_1 und Δ_2 jeweils $< \frac{\pi}{4}$ (vgl. Fig. 1), so daß Δ_3 geknickt ist, d.h. $\Delta_3 =: l'O_1lO_2$.

Das l zugeordnete $\mathcal{L}$-Dreieck ist $O_1O_2O_3$, so daß nach (15.2) und (15.3) gilt, daß $|\overline{O_1O_2}| > \varrho$ und $U(O_1O_2O_3) \geqq \varrho$.

H sei die von O_1O_2 begrenzte abgeschlossene Halbsphäre so, daß $l \in H$. Es sei $Q(l) := H \cap \partial K_1 \cap \partial K_2$ und $Q(O_3) := O_2$.

Daraus folgt sofort (18.10), und daß Q die Summe der β-Winkel konstant läßt. Daß Q die Summe der α-Winkel höchstens verkleinert, folgt deshalb aus $|Q(l)O_1O_3 \cup Q(l)O_2O_3| \leqq |Q(l)O_1Q(O_3) \cup Q(l)O_2Q(O_3)|$.

Sei $D_1 := H \cap \partial K_1 \cap \partial K_3$, $D_2 := H \cap \partial K_2 \cap \partial K_3$. Dann gilt für die Kreisbogen $\overset{\frown}{D_1D_2}$, $\overset{\frown}{D_1Q(l)}$, $\overset{\frown}{D_2Q(l)}$ wie beim Beweis von Hilfssatz 18.6 (siehe Fig. 12), daß $|\overset{\frown}{D_1D_2}| \leqq |\overset{\frown}{D_1Q(l)}| + |\overset{\frown}{D_2Q(l)}|$. Deshalb ist $e_\varrho(\Delta_3) + e_\varrho(\Delta_1) + e_\varrho(\Delta_2) \leqq |(K_1 \cup K_2) \cap l'O_1Q(l)O_2| = e_\varrho(Q(\Delta_3))$, d.h. (18.12) gilt. □

Die Operation Q wurde bzgl. einer festen Quelle l definiert, d.h. $Q = Q_l$. Q_l beeinflußt nur die mit l inzidierenden g-Dreiecke, und diese inzidieren mit keiner anderen Quelle. Deshalb läßt sich Q nacheinander bzgl. jeder Quelle anwenden, die mit zwei schmalen g-Dreiecken inzidiert.

18.5. *Die Verschiebungsoperation V*

Auf $\mathcal{D}_\Gamma$ wird eine Verschiebungsoperation V definiert. Sie soll bewirken, daß einem schmalen g-Dreieck von $\mathcal{D}_\Gamma$ ein geknicktes g-Dreieck so zugeordnet wird, daß die Summe ihrer β-Winkel $\geqq \pi$ ist.

l_S sei die Senke von Γ. Es sei $l \neq l_S$ ein Knoten von Γ, der mit einem schmalen g-Dreieck $\Delta = lO_1O_2$ inzidiert, siehe Fig. 13. Wegen $l \neq l_S$ gibt

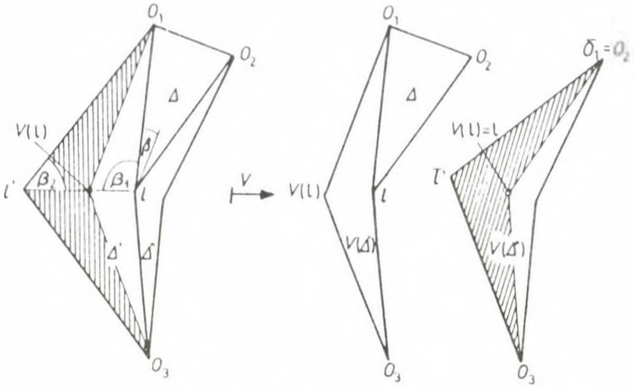


Fig. 13

es eine Kante (l, l') von Γ, die Diagonale eines geknickten g-Dreiecks Δ' ist, wobei o.B.d.A. $\Delta' = l'O_1lO_3$. l inzidiert außerdem mit genau einem weiteren g-Dreieck Δ''. Δ'' hat l, O_2, O_3 als Ecken, und Δ'' ist genau dann geknickt, wenn l keine Quelle ist. Die Operation V soll in $\mathcal{D}_\Gamma$ die g-Dreiecke Δ' und Δ'' durch gewisse g-Dreiecke $V(\Delta')$ und $V(\Delta'')$ ersetzen. Die β-Winkel von Δ und Δ' seien mit β, β_1, β_2 bezeichnet wie in Fig. 13.

Ist dann $\beta + \beta_1 + \beta_2 \geqq \pi$, so sei $V(\Delta') := \Delta'$, $V(\Delta'') := \Delta''$.

Ist aber $\beta + \beta_1 + \beta_2 < \pi$, so ist $\pi - (\beta + \beta_1) > \beta_2$. Deshalb gibt es einen Punkt $V(l)$ echt zwischen l und l' so, daß $\sphericalangle O_1V(l)l = \pi - (\beta + \beta_1)$. Es sei $V(\Delta') := V(l)O_1lO_3$. Die Operation V liefert also zu dem schmalen g-Dreieck Δ ein geknicktes g-Dreieck $V(\Delta')$ so, daß die Summe ihrer β-Winkel $\geqq \pi$ ist.

Wegen $|\overline{V(l)O_1}| > |\overline{lO_1}|$ kann man das geknickte g-Dreieck $l'O_1V(l)O_3$ schrumpfen lassen zu einem geknickten g-Dreieck $\overline{l'}\,\overline{O}_1V(l)\overline{O}_3$ so, daß $\overline{O}_1 \in \overline{V(l)O_1}$, $|\overline{V(l)\overline{O}_1}| = |\overline{lO_1}|$, und daß der α-Winkel gleich bleibt; der β-Winkel an der Spitze wird daher kleiner. Verschiebt man nun $\overline{l'}\,\overline{O}_1V(l)\overline{O}_3$ so, daß $V(l)$ mit l und $\overline{O}_1$ mit O_2 zusammenfällt und vereinigt dann mit Δ'', so erhält man ein g-Dreieck $V(\Delta'')$, das wie Δ'' schmal, breit, oder geknickt ist.

Man erreicht daher

(18.13) *Die Operation V läßt die Summe der α-Winkel konstant, die Summe der β-Winkel wird höchstens verkleinert.*

(18.14) *Die Operation V läßt die Anzahl der schmalen, breiten und geknickten g-Dreiecke jeweils konstant.*

Aus $\mathcal{D}_\Gamma$ entstehe durch V die Menge $V(\mathcal{D}_\Gamma)$. Dann gilt weiter

(18.15)
$$\sum_{\Delta \in \mathcal{D}_\Gamma} e_\varrho(\Delta) = \sum_{\Delta \in V(\mathcal{D}_\Gamma)} e_\varrho(\Delta),$$

wenn Δ'' geknickt ist oder wenn Δ'' einem Punkt aus $G_1 \cup G_2$ entspricht. Die Gleichung (18.15) gilt aber — leider — nicht, wenn Δ'' einem Punkt aus $G_3 \cup G_4$ entspricht.

Die Operation V wurde bzgl. eines festen Knotens l definiert, d.h. $V = V_l$, $\Delta = \Delta_l$, $\Delta' = \Delta'_l$ und $\Delta'' = \Delta''_l$. V_l beeinflußt nur Δ'_l und Δ''_l. Inzidieren Knoten $l_1 \neq l_S$ und $l_2 \neq l_S$ mit einem schmalen g-Dreieck, so läßt sich also $V_{l_2}(V_{l_1}(\mathcal{D}_\Gamma))$ bilden, wenn l_1 und l_2 durch keine Kante von Γ verbunden sind. $V_{l_2}(V_{l_1}(\mathcal{D}_\Gamma))$ läßt sich aber auch dann bilden, wenn (l_2, l_1) Kante von Γ ist; man setze dann $\Delta'_{l_2} := V_{l_1}(\Delta''_{l_1})$.

Einen gewissen Ersatz für (18.15) im Fall, daß Δ'' einem Punkt aus G_3 entspricht, liefert

HILFSSATZ 18.9. *Sei l_1 eine Quelle von Γ, die mit einem schmalen und einem breiten g-Dreieck inzidiere. $\mathcal{B}$ sei eine von l_1 ausgehende Bahn von Γ, mit den Knoten $l_1, l_2, \ldots, l_m$ $(m \geqq 1)$ in natürlicher Reihenfolge; jeder Knoten von $\mathcal{B}$ inzidiere mit einem schmalen g-Dreieck, und es sei $l_m \neq l_S$, der Senke von Γ. $\mathcal{D}_\mathcal{B}$ sei die Menge der g-Dreiecke von $\mathcal{D}_\Gamma$, die mit einem Knoten von $\mathcal{B}$ inzidieren, $S_\mathcal{B}$ sei ihre Vereinigung.*

Wendet man dann — von l_m ausgehend — successive in jedem Knoten von $\mathcal{B}$ die Operation V an, so daß man eine g-Dreiecksmenge

$$V_{l_1}(V_{l_2}(\ldots V_{l_m}(\mathcal{D}_\mathcal{B})\ldots))$$

erhält, so gilt

$$|E(K_1, \ldots, K_n) \cap S_\mathcal{B}| \leqq \sum_{\Delta \in V_{l_1}(V_{l_2}(\ldots V_{l_m}(\mathcal{D}_\mathcal{B})\ldots))} e_\varrho(\Delta). \tag{18.16}$$

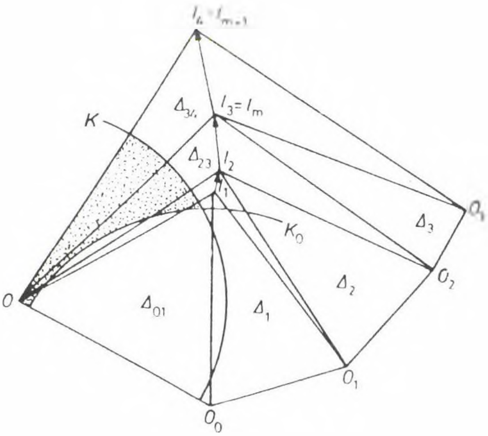


Fig. 14

Beweis. l_1 sei also Quelle von Γ, $\mathcal{B}$ sei eine von l_1 ausgehende Bahn mit den Knoten $l_1, \dots, l_m$ ($m \geqq 1$) in natürlicher Reihenfolge, und es sei $l_m \neq l_S$, der Senke von Γ. Nach Voraussetzung inzidiert jeder Knoten l_j mit einem schmalen g-Dreieck Δ_j ($j = 1, \dots, m$). Wegen $l_m \neq l_S$ gibt es eine Kante (l_m, l_{m+1}) von Γ. Die beiden Knoten l_j, l_{j+1} ($j = 1, \dots, m$) inzidieren mit einem geknickten g-Dreieck $\Delta_{j\ j+1}$. l_1 inzidiert nach Voraussetzung außer mit Δ_1 und Δ_{12} mit einem breiten g-Dreieck Δ_{01}. Δ_1 habe die Basisecken O_0 und O_1, Δ_{01} habe die Basisecken O und O_0. Dann ist O auch Basisecke von Δ_{12}.

Zunächst nehme man an, daß O Basisecke aller $\Delta_{j\ j+1}$ sei ($j = 2, \dots, m$), siehe Fig. 14. Die andere Basisecke von $\Delta_{j\ j+1}$ sei O_j. Sei $j \geqq 1$.

Wegen Hilfssatz 18.7 ist $Ol_jl_{j+1} \setminus K$ fremd zu jedem Systemkreis. Deshalb ist $|E(K_1, \dots, K_n) \cap O\, l_j\, l_{j+1}| \leqq |O\, l_j\, l_{j+1} \cap K \setminus K_0|$.

Ebenso ist $|E(K_1, \dots, K_n) \cap O_j\, l_j\, l_{j+1}| \leqq |O_j\, l_j\, l_{j+1} \cap K_j \setminus K_{j-1}|$. Der Kreis um l_j durch O_{j-1} ist ein Stützkreis und enthält also keinen Systemkreismittelpunkt im Inneren, so daß $|\overline{l_j\, O_{j-1}}| \leqq |\overline{l_j O_0}|$. Weil Δ_j schmal ist und Δ_{01} breit, gilt außerdem $|\overline{O_{j-1}\, O_j}| < |\overline{O_0\, O}|$. Deshalb gilt weiter $|O_j\, l_j\, l_{j+1} \cap K_j \setminus K_{j-1}| \leqq |O\, l_j\, l_{j+1} \cap K \setminus K_0|$.

Aus beidem zusammen folgt

$$|E(K_1, \dots, K_n) \cap \Delta_{j\ j+1}| \leqq 2|O\, l_j\, l_{j+1} \cap K \setminus K_0|.$$

Wendet man nun die Operation V successive bzgl. $l_m, l_{m-1}, \dots, l_1$ an, so sieht man, daß (18.16) auch dann gilt, wenn Δ_{01} einem Punkt aus G_3 entspricht.

Schließlich überlegt man, daß im allgemeinen Fall die Dreiecke von $\mathcal{D}_\mathcal{B}$ so umgelegt werden können, daß alle geknickten g-Dreiecke durch O gehen, und daß das einfach Bedeckte in $\Delta_{j\ j+1}$ dann erst recht mit Hilfe von $K \setminus K_0$ abgeschätzt werden kann. □

18.6. *Die Polygonfamilie ε_Γ*

Ziel von **18.** ist, $|E(K_1, \dots, K_n) \cap S_\Gamma|$ abzuschätzen für die in **18.** betrachteten Zusammenhangskomponenten Γ. Wegen (18.6) und (18.9) ist $|E(K_1, \dots, K_n) \cap S_\Gamma| \leqq \sum_{\Delta \in \mathcal{D}_\Gamma} e_\varrho(\Delta)$. Dabei besteht $\mathcal{D}_\Gamma$ aus breiten, schmalen oder geknickten g-Dreiecken. Jedem schmalen oder breiten g-Dreieck entspricht durch α- und β-Winkel ein Punkt $(\alpha, \beta) \in G = G_1 \cup G_2 \cup G_3 \cup G_4$ (siehe Fig. 1).

$\mathcal{D}_\Gamma$ wird jetzt in eine Familie ε_Γ nur aus breiten oder schmalen g-Dreiecken transformiert. Dabei sollen die g-Dreiecke von ε_Γ Punkten (α, β) entsprechen, die die Voraussetzungen von Satz 2 erfüllen.

Folgende Definition wird benötigt: Seien $\Delta_1, \dots, \Delta_k$ nicht geknickte g-Dreiecke mit β-Winkel $\beta_1, \dots, \beta_k$, und seien $\Delta'_1, \dots, \Delta'_r$ geknickte g-Dreiecke mit β-Winkel $\beta'_1, \dots, \beta'_{2r}$. Dann ist $\frac{\sum_{i=1}^{k} \beta_i + \sum_{i=1}^{2r} \beta'_i}{k+2r}$ der *mittlere β-Winkel* der Dreiecke $\Delta_1, \dots, \Delta_k, \Delta'_1, \dots, \Delta'_r$.

Es gilt

HILFSSATZ 18.10. *Γ sei eine Zusammenhangskomponente, bei der für jedes Dreieck $L \in \mathcal{L}_\Gamma$ gilt, daß L eine Seite s mit $|s| > \varrho$ besitzt, und daß $U(L) \geqq \varrho$. Es sei k_Γ die Anzahl der $\mathcal{L}$-Dreiecke von $\mathcal{L}_\Gamma$, und $\mathcal{D}_\Gamma$ sei wie in* **18.2** *definiert.*

Dann gibt es eine Menge ε_Γ aus $3k_\Gamma$ schmalen oder breiten g-Dreiecken mit

(18.17)
Die Summe der α-Winkel von ε_Γ ist höchstens kleiner als die von $\mathcal{D}_\Gamma$,
die Summe der β-Winkel von ε_Γ ist höchstens kleiner als die von $\mathcal{D}_\Gamma$.

$$|E(K_1, \dots, K_n) \cap S_\Gamma| \leqq \sum_{\Delta \in \varepsilon_\Gamma} e_\varrho(\Delta). \tag{18.18}$$

Außerdem läßt sich ε_Γ zerlegen in eine Menge ε^1_Γ nur aus breiten g-Dreiecken und in eine Menge ε^2_Γ, die höchstens doppelt so viele breite wie schmale g-Dreiecke hat. Der mittlere β-Winkel von ε^2_Γ ist $\geqq \frac{\pi}{3}$, und die breiten g-Dreiecke von ε^2_Γ entsprechen Punkten $(\alpha, \beta) \in G_1 \cup G_2$.

BEWEIS. Der Beweis verwendet
— Verschiebungsoperationen V nach **18.5**,
— Quellenoperationen Q nach **18.4** und
— Zerlegungen von geknickten g-Dreiecken in zwei breite nach Hilfssatz 18.1.

Sei l_Q eine Quelle von Γ, und auf $\mathcal{D}_\Gamma$ seien in Knoten $l_{\dots} \neq l_Q$ Verschiebungsoperationen $V_{l_{\dots}}$ angewandt. In der Menge $V_{l_{\dots}}(\dots V_{l_{\dots}}(\mathcal{D}_\Gamma)\dots)$ kann dann das geknickte g-Dreieck mit l_Q als Gegenspitze Ergebnis von Verschiebungsoperationen sein. Auch dann kann eine Quellenoperation bzgl. l_Q angewandt werden, weil Hilfssatz 18.8 offensichtlich auch nach der Ersetzung von $\mathcal{D}_\Gamma$ durch $V_{l_{\dots}}(\dots V_{l_{\dots}}(\mathcal{D}_\Gamma)\dots)$ gilt.

Die Zusammenhangskomponente Γ wird in Hilfssatz 18.2 beschrieben.

Die Anzahl k_Γ der $\mathcal{L}$-Dreiecke von $\mathcal{L}_\Gamma$ ist gleich der Anzahl der Knoten von Γ. Der Baum Γ hat somit k_Γ Knoten und $k_\Gamma - 1$ Kanten. Weil nach Hilfssatz 18.5 die Kanten von Γ genau die Diagonalen der geknickten g-Dreiecke sind, hat $\mathcal{D}_\Gamma$ genau $k_\Gamma - 1$ geknickte g-Dreiecke. Weil ebenfalls nach Hilfssatz 18.5 jeder Knoten von Γ mit genau drei g-Dreiecken inzidiert, und jedes geknickte g-Dreieck mit genau zwei Knoten, jedes nicht geknickte g-Dreieck mit genau einem Knoten, hat $\mathcal{D}_\Gamma$ außerdem genau $k_\Gamma + 2$ nicht geknickte g-Dreiecke.

Deshalb folgt aus (18.8)

(18.19) Der mittlere β-Winkel von $\mathcal{D}_\Gamma$ ist $\frac{\pi}{3}$.

Verschiebungsoperationen und Quellenoperationen lassen die Anzahl der geknickten und der nicht geknickten g-Dreiecke jeweils konstant. Entsteht also ε_Γ aus $\mathcal{D}_\Gamma$ durch Verschiebungsoperationen, Quellenoperationen und Zerlegung geknickter g-Dreiecke in zwei breite gemäß Hilfssatz 18.1, so besteht ε_Γ aus $3k_\Gamma$ schmalen oder breiten g-Dreiecken, wie behauptet. Wegen (18.13), (18.11) und (18.2) gilt dann außerdem die Behauptung (18.17).

Besitzt nun $\mathcal{D}_\Gamma$ höchstens zwei breite g-Dreiecke, so bestehe ε_Γ^1 aus den breiten g-Dreiecken von $\mathcal{D}_\Gamma$ mit β-Winkel $< \frac{\pi}{3}$, und ε_Γ^2 bestehe aus allen schmalen und zerlegten geknickten g-Dreiecken von $\mathcal{D}_\Gamma$ und aus den breiten g-Dreiecken mit β-Winkel $\geqq \frac{\pi}{3}$; die breiten g-Dreiecke von ε_Γ^2 entsprechen dann wegen (18.4) bzw. wegen ihres β-Winkels $\geqq \frac{\pi}{3}$ Punkten aus $G_1 \cup G_2$. Dann ist wegen (18.19) für ε_Γ^2 der mittlere β-Winkel $\geqq \frac{\pi}{3}$, und ε_Γ^2 hat höchstens $2(k_\Gamma - 1) + 2$ breite g-Dreiecke und mindestens k_Γ schmale g-Dreiecke. Für $\varepsilon_\Gamma := \varepsilon_\Gamma^1 \cup \varepsilon_\Gamma^2$ folgt (18.18) aus (18.6), (18.9) und (18.5). Deshalb erfüllt ε_Γ die Behauptung von Hilfssatz 18.10.

Es wird also im folgenden angenommen, daß $\mathcal{D}_\Gamma$ mindestens drei breite g-Dreiecke besitzt.

Γ ist ein Baum mit genau einer Senke l_S. Die Eingangsvalenz von l_S sei s. Ist $s = 0$, so besteht $\mathcal{D}_\Gamma$ genau aus den drei mit l_S inzidierenden breiten g-Dreiecken, und mit $\varepsilon_\Gamma^1 := \mathcal{D}_\Gamma$, $\varepsilon_\Gamma^2 := \emptyset$ gilt Hilfssatz 18.10.

Es sei also im folgenden $1 \leqq s \leqq 3$.

Ein von l_S verschiedener Knoten von Γ mit Eingangsvalenz 2 heiße *Verzweigungsknoten*, k_V sei die Anzahl der Verzweigungsknoten. Ein von l_S, von einem Verzweigungsknoten und einer Quelle verschiedener Knoten heiße *gewöhnlicher Knoten*, er hat Eingangs- und Ausgangsvalenz 1.

Die Anzahl der Quellen von Γ ist $s + k_V$. Eine Quelle kann mit 0, 1 oder 2 breiten g-Dreiecken inzidieren und heißt dann *vom* 0., 1. *bzw.* 2. *Typ*; die Anzahl der entsprechenden Quellen von Γ sei k_{Q0}, k_{Q1} bzw. k_{Q2}. Damit ist $s + k_V = k_{Q0} + k_{Q1} + k_{Q2}$ oder

(18.20) $$s - k_{Q1} - 2k_{Q2} + k_V + k_{Q2} = k_{Q0}.$$

Gesucht wird die Anzahl der geknickten g-Dreiecke, deren Gegenspitze nicht mit genau einem schmalen g-Dreieck inzidiert: Weil jede Quelle vom 0. Typ mit einem solchen geknickten g-Dreieck inzidiert, gibt es also k_{Q0} davon. Weitere k_{Q0} erhält man mit Hilfe von (18.20) folgendermaßen: Weil $\mathcal{D}_\Gamma$ nach Voraussetzung mindestens drei breite g-Dreiecke hat, hat $\mathcal{D}_\Gamma$ mindestens s breite g-Dreiecke, die nicht mit l_S inzidieren. Deshalb hat $\mathcal{D}_\Gamma$ mindestens

$s - k_{Q1} - 2k_{Q2}$ breite g-Dreiecke, die mit gewöhnlichen Knoten inzidieren. Weil jeder gewöhnliche Knoten, jeder Verzweigungsknoten und jede Quelle Gegenspitze eines geknickten g-Dreiecks ist, folgt also aus (18.20)) die Existenz weiterer k_{Q0} geknickter g-Dreiecke, deren Gegenspitze nicht mit genau einem schmalen g-Dreieck inzidiert. Also gilt

(18.21) $\mathcal{D}_\Gamma$ hat mindestens $2k_{Q0}$ geknickte g-Dreiecke, deren Gegenspitze nicht mit genau einem schmalen g-Dreieck inzidiert.

Als Senke ist l_S nicht Gegenspitze eines geknickten g-Dreiecks. Weil der β-Winkel eines schmalen g-Dreiecks $< \frac{\pi}{4}$ ist (vgl. Fig. 1), ist also l_S Spitze von höchstens einem schmalen g-Dreieck. Dementsprechend werden zwei Fälle betrachtet:

1. Fall. Die Senke l_S von Γ inzidiert mit keinem schmalen g-Dreieck.

Es sei l_{Q1} eine Quelle von Γ, die mit genau einem breiten g-Dreieck inzidiere. Es gibt genau eine Bahn von l_{Q1} nach l_S; $\mathcal{B}$ sei die längste Bahn mit Anfangspunkt l_{Q1}, bei der alle Knoten mit einem schmalen g-Dreieck inzidieren. Wie in Hilfssatz 18.9 sei $\mathcal{D}_\mathcal{B}$ die Menge der g-Dreiecke, die mit einem Knoten von $\mathcal{B}$ inzidieren, und es sei $S_\mathcal{B}$ ihre Vereinigung. Man durchlaufe nun $\mathcal{B}$ rückwärts, d.h. von seinem Endpunkt ausgehend, und wende in jedem Knoten von $\mathcal{B}$ die Verschiebungsoperation an. Jedem schmalen g-Dreieck von $\mathcal{D}_\mathcal{B}$ wird dadurch ein geknicktes g-Dreieck zugeordnet so, daß ihr mittlerer β-Winkel $\geqq \frac{\pi}{3}$ ist, und die Menge dieser schmalen und geknickten g-Dreiecke nach ihrer Zerlegung sei $\varepsilon^2_\mathcal{B}$. Aus dem mit l_{Q1} inzidierenden breiten g-Dreieck von $\mathcal{D}_\mathcal{B}$ entsteht durch die Verschiebungsoperationen ein breites g-Dreieck, das $\varepsilon^1_\mathcal{B}$ bilde. Sei $\varepsilon_\mathcal{B} := \varepsilon^1_\mathcal{B} \cup \varepsilon^2_\mathcal{B}$. Wegen Hilfssatz 18.9 gilt $|E(K_1, \ldots, K_n) \cap S_\mathcal{B}| \leqq \sum_{\Delta \in \varepsilon_\mathcal{B}} e_\varrho(\Delta)$, und $\varepsilon^2_\mathcal{B}$ ist eine Menge aus höchstens doppelt so viel breiten wie schmalen g-Dreiecken mit mittlerem β-Winkel $\geqq \frac{\pi}{3}$, und seine breiten g-Dreiecke entsprechen nach (18.4) Punkten aus $G_1 \cup G_2$.

Für alle solchen Bahnen $\mathcal{B}$ behandle man $\mathcal{D}_\mathcal{B}$ wie beschrieben. Man streiche diese Bahnen $\mathcal{B}$ in Γ, der entstehende Graph sei Γ_-. Aus $\mathcal{D}_\Gamma$ entferne man die entsprechenden Familien $\mathcal{D}_\mathcal{B}$, die verbleibende Menge aus g-Dreiecken sei $\mathcal{D}_-$; ihre Vereinigung sei S_-. Dann ist $|E(K_1, \ldots, K_n) \cap S_\Gamma| = |E(K_1, \ldots, K_n) \cap S_-| + \sum_\mathcal{B} |E(K_1, \ldots, K_n) \cap S_\mathcal{B}|$.

Knoten von Γ_-, die mit dem Endpunkt einer Bahn $\mathcal{B}$ durch eine Kante verbunden sind, inzidieren nicht mit einem schmalen g-Dreieck. Ein Knoten von Γ_-, der mit einem schmalen g-Dreieck inzidiert und keine Quelle von Γ ist, inzidiert also mit derselben Dreiecksmenge aus $\mathcal{D}_-$ wie aus $\mathcal{D}_\Gamma$, so daß sich bzgl. ihm eine Verschiebungsoperation auf $\mathcal{D}_-$ anwenden läßt. Durchläuft man die Bahnen von Γ_- mit l_S als Endpunkt rückwärts, d.h. von l_S ausgehend und wendet die Verschiebungsoperationen in der dadurch entstehenden Reihenfolge an, so läßt sich also in jedem Punkt von Γ_-, der mit einem schmalen g-Dreieck inzidiert und keine Quelle von Γ ist, eine Verschiebungsoperation auf $\mathcal{D}_-$ anwenden. Den entsprechenden schmalen g-Dreiecken von $\mathcal{D}_-$ wird dadurch ein geknicktes g-Dreieck zugeordnet so, daß ihr mittlerer β-Winkel $\geqq \frac{\pi}{3}$ ist. Diese schmalen und geknickten g-Dreiecke nach ihrer Zerlegung sollen ε^2 bilden.

Die übrigen schmalen g-Dreiecke von $\mathcal{D}_-$ inzidieren mit Quellen von Γ, und zwar mit den k_{Q0} Quellen, die mit genau zwei schmalen g-Dreiecken inzidieren. In jeder solchen Quelle läßt sich nach den Eingangsüberlegungen nun eine Quellenoperation ausführen, so daß genau k_{Q0} schmale g-Dreiecke übrig bleiben. Wegen (18.21) hat $\mathcal{D}_\Gamma$ mindestens $2k_{Q0}$ geknickte g-Dreiecke, deren Gegenspitze nicht mit genau einem schmalen g-Dreieck inzidiert, so daß sie zu $\mathcal{D}_-$ gehören und auch nicht für ε^2 benützt wurden. Zerlegt man sie gemäß Hilfssatz 18.1, so erhält man $2k_{Q0}$ entartete g-Dreiecke mit β-Winkel $\frac{\pi}{2}$. Die betrachteten k_{Q0} schmalen g-Dreiecke zusammen mit den $2k_{Q0}$ entarteten breiten g-Dreiecken haben einen mittleren β-Winkel $\geqq \frac{\pi}{3}$, und sie sollen zusammen mit ε^2 die Menge ε^2_- bilden. ε^1_- sei die Menge der nicht in ε^2_- auftretenden breiten und zerlegten geknickten g-Dreiecke von $\mathcal{D}_-$ nach den beschriebenen Verschiebungs- und Quellenoperationen. Sei $\varepsilon_- := \varepsilon^1_- \cup \varepsilon^2_-$. Wegen (18.12) und (18.15) gilt $|E(K_1, \ldots, K_n) \cap S_-| \leqq \sum_{\Delta \in \varepsilon_-} e_\varrho(\Delta)$, und ε^2_- ist nach Konstruktion eine Menge aus höchstens doppelt so viel breiten wie schmalen g-Dreiecken, mit mittlerem β-Winkel $\geqq \frac{\pi}{3}$, und seine breiten g-Dreiecke entsprechen Punkten aus $G_1 \cup G_2$.

Damit ist Hilfssatz 18.10 richtig im betrachteten 1. Fall.

2. Fall. Die Senke l_S von Γ inzidiert mit genau einem schmalen g-Dreieck.

Für die Eingangsvalenz s von l_S gilt nach Voraussetzung $1 \leqq s \leqq 3$, so daß jetzt $1 \leqq s \leqq 2$. Γ läßt sich in genau s (nichttriviale) Bäume zerlegen, die l_S als Knoten haben, dies seien die *Äste* von Γ.

Weil $\mathcal{D}_\Gamma$ mindestens drei breite g-Dreiecke hat, inzidieren mindestens zwei davon mit Knoten desselben Astes $\mathcal{A}$, und nicht mit l_S. Man durchlaufe $\mathcal{A}$ von l_S aus rückwärts, l_1 sei der dabei als erstes angetroffene Knoten, der entweder gewöhnlich ist und mit einem breiten g-Dreieck inzidiert, oder Verzweigungsknoten ist, oder Quelle; im letzten Fall enthält also $\mathcal{A}$ keinen Verzweigungsknoten und kein gewöhnlicher Knoten von $\mathcal{A}$ inzidiert mit einem breiten g-Dreieck, so daß die Quelle mit genau zwei breiten g-Dreiecken inzidiert.

$\mathcal{A}_1$ sei der Graph mit Knoten l_1 und allen Knoten von $\mathcal{A}$ vor l_1. Δ_1 sei das geknickte g-Dreieck mit l_1 als Gegenspitze, es werde gemäß Hilfssatz 18.1 zerlegt in ein breites g-Dreieck Δ'_1 und ein breites entartetes g-Dreieck Δ''_1. In der Menge der g-Dreiecke von $\mathcal{D}_\Gamma$, die mit Knoten von $\mathcal{A}_1$ inzidieren, ersetze man Δ_1 durch Δ''_1; die entstehende Menge sei $\mathcal{D}_{\mathcal{A}_1}$. $S_{\mathcal{A}_1}$ sei die Vereinigung der g-Dreiecke von $\mathcal{D}_{\mathcal{A}_1}$.

Hat Γ außer $\mathcal{A}$ einen weiteren Ast so, daß mindestens zwei breite g-Dreiecke mit Knoten dieses Astes inzidieren, und trivialerweise nicht mit l_S, so behandle man auch diesen Ast wie für $\mathcal{A}$ beschrieben, d.h. man definiere

l_2, $\mathcal{A}_2$, Δ_2, Δ_2', Δ_2'', $\mathcal{D}_{\mathcal{A}_2}$, $S_{\mathcal{A}_2}$ analog zu oben.

Nun sei Γ_- der Graph, der durch Streichen von $\mathcal{A}_1$ und gegebenenfalls von $\mathcal{A}_2$ aus Γ entsteht. In der Menge der g-Dreiecke von $\mathcal{D}_\Gamma$, die mit Knoten von Γ_- inzidieren, ersetze man Δ_1 durch Δ_1' und gegebenenfalls Δ_2 durch Δ_2'; die entstehende Menge sei $\mathcal{D}_{\Gamma_-}$. Dann gilt

$$|E(K_1,\ldots,K_n)\cap S_\Gamma| = \sum_{\Delta\in\mathcal{D}_{\Gamma_-}} |E(K_1,\ldots,K_n)\cap\Delta| +$$
$$+|E(K_1,\ldots,K_n)\cap S_{\mathcal{A}_1}| + |E(K_1,\ldots,K_n)\cap S_{\mathcal{A}_2}|,$$

wobei der Summand mit $S_{\mathcal{A}_2}$ nur gegebenenfalls auftritt.

Auf jeden Fall ist wegen (18.3) für $\mathcal{D}_{\Gamma_-}$ der mittlere β-Winkel $\geqq \frac{\pi}{3}$. $\mathcal{D}_{\Gamma_-}$ besitzt die breiten g-Dreiecke Δ_1' und gegebenenfalls Δ_2'', auf jeden Fall nach Konstruktion höchstens zwei breite g-Dreiecke. Nach Zerlegung aller geknickten g-Dreiecke gemäß Hilfssatz 18.1 hat $\mathcal{D}_{\Gamma_-}$ also höchstens doppelt so viele breite wie schmale g-Dreiecke. Die breiten g-Dreiecke von $\mathcal{D}_{\Gamma_-}$ mit β-Winkel $\leqq \frac{\pi}{3}$ sollen die Menge $\varepsilon^1_{\Gamma_-}$ bilden, die übrigen breiten g-Dreiecke, die schmalen und zerlegten geknickten g-Dreiecke von $\mathcal{D}_{\Gamma_-}$ sollen $\varepsilon^2_{\Gamma_-}$ bilden. Dann ist $\varepsilon^2_{\Gamma_-}$ eine Menge aus höchstens doppelt so viel breiten wie schmalen g-Dreiecken, mit mittlerem β-Winkel $\geqq \frac{\pi}{3}$, und die breiten g-Dreiecke entsprechen Punkten aus $G_1 \cup G_2$. Für $\varepsilon_{\Gamma_-} := \varepsilon^1_{\Gamma_-} \cup \varepsilon^2_{\Gamma_-}$ gilt nach (18.5), daß $\sum_{\Delta\in\mathcal{D}_{\Gamma_-}} |E(K_1,\ldots,K_n)\cap\Delta| = \sum_{\Delta\in\varepsilon_{\Gamma_-}} |E(K_1,\ldots,K_n)\cap\Delta|$.

Es bleibt, $\mathcal{D}_{\mathcal{A}_1}$ (und gegebenenfalls $\mathcal{D}_{\mathcal{A}_2}$ völlig analog) weiter zu behandeln. Ist l_1 Quelle, so inzidiert sie nach oben mit zwei breiten g-Dreiecken, und $\mathcal{D}_{\mathcal{A}_1}$ besteht nur aus breiten g-Dreiecken; dann sei $\varepsilon^1_{\mathcal{A}_1} := \mathcal{D}_{\mathcal{A}_1}$, $\varepsilon^2_{\mathcal{A}_2} := \emptyset$. Sonst aber sei $s(\mathcal{A}_1)$ die Eingangsvalenz von l_1, d.h. $1 \leqq s(\mathcal{A}_1) \leqq 2$. Dann wird $\mathcal{D}_{\mathcal{A}_1}$ analog zu $\mathcal{D}_\Gamma$ im 1. Fall behandelt, wobei l_1 die Rolle des dortigen l_S spiele. Dabei beachte man: Nach Definition von $\mathcal{A}_1$ gibt es mindestens zwei breite g-Dreiecke von $\mathcal{D}_\Gamma$, die mit Knoten von $\mathcal{A}_1$ inzidieren, also zu $\mathcal{D}_{\mathcal{A}_1}$ gehören. Deshalb gibt es in $\mathcal{D}_{\mathcal{A}_1}$ mindestens $s(\mathcal{A}_1)$ breite g-Dreiecke, die nicht mit l_1 inzidieren. Für $\mathcal{D}_{\mathcal{A}_1}$ gilt also die (18.21) entsprechende Aussage.

Aus allem zusammen folgt, daß Hilfssatz 18.10 auch im betrachteten 2. Fall richtig ist. □

Für die Beziehung zwischen ε_Γ und $\mathcal{D}_\Gamma$ gilt außerdem

HILFSSATZ 18.11. *Es gilt* $\varepsilon^2_\Gamma = \emptyset$ *genau dann, wenn* $\mathcal{D}_\Gamma$ *keine schmalen g-Dreiecke enthält. Dann entsteht* ε_Γ *aus* $\mathcal{D}_\Gamma$ *ausschließlich durch Zerlegen der geknickten g-Dreiecke von* $\mathcal{D}_\Gamma$.

BEWEIS. Nach dem Beweis von Hilfssatz 18.10 entsteht ε_Γ aus $\mathcal{D}_\Gamma$ durch Anwendung von Verschiebungsoperationen und Zerlegen von geknickten g-Dreiecken — diese Operationen beeinflussen schmale g-Dreiecke nicht — und durch Quellenoperationen — bei diesen existieren vorher und nachher schmale g-Dreiecke. Deshalb gilt $\varepsilon_\Gamma^2 = \emptyset$ genau dann, wenn $\mathcal{D}_\Gamma$ keine schmalen g-Dreiecke enthält.

Klar ist, daß dann ε_Γ aus $\mathcal{D}_\Gamma$ ausschließlich durch Zerlegen der geknickten g-Dreiecke von $\mathcal{D}_\Gamma$ entsteht. □

18.7. *Die Gültigkeit der Ungleichung* (15.8)

Es gilt nun endlich der zu den Hilfssätzen 16.1 und 17.1 analoge

HILFSSATZ 18.12. *Es sei Γ eine Zusammenhangskomponente, bei der für jedes Dreieck $L \in \mathcal{L}_\Gamma$ gilt, daß L eine Seite s mit $|s| > \varrho$ besitzt, und daß $U(L) \geqq \varrho$. Es sei k_Γ die Anzahl der $\mathcal{L}$-Dreiecke von $\mathcal{L}_\Gamma$, und ein $\mathcal{L}$-Dreieck L habe die Winkelsumme ω_L. $\mathcal{D}_\Gamma$ sei wie in* **18.2** *definiert. Dann gilt*

$$|E(K_1, \dots, K_n) \cap S_\Gamma| \leqq 6k_\Gamma F\left(\frac{\sum\limits_{L\in\Gamma} \omega_L}{6k_\Gamma}\right), \tag{18.22}$$

d.h. die Ungleichung (15.8) *ist richtig.*

Aus der Gültigkeit des Gleichheitszeichens in (18.22) *folgt, daß $\mathcal{L}_\Gamma$ aus einem einzigen, gleichseitigen Dreieck L besteht, und daß für die Systemkreise K_1, K_2, K_3 in den Ecken von L gilt: $\partial K_i \cap L$ wird von $\{K_1, K_2, K_3\}$ genau zur Hälfte einfach überdeckt* $(i = 1, 2, 3)$.

BEWEIS. Wegen Hilfssatz 18.10 ist nach (18.18)

$$|E(K_1, \dots, K_n) \cap S_\Gamma| \leqq \sum_{\Delta \in \varepsilon_\Gamma} e_\varrho(\Delta).$$

Dabei besteht ε_Γ aus $3k_\Gamma$ schmalen oder breiten g-Dreiecken, die also Punkten $(\alpha_i, \beta_i) \in G = G_1 \cup G_2 \cup G_3 \cup G_4$ entsprechen (siehe Fig. 1), $i = 1, \dots, 3k_\Gamma$. Nach Definition (2.5) ist

$$\sum_{\Delta \in \varepsilon_\Gamma} e_\varrho(\Delta) = 2 \sum_{i=1}^{3k_\Gamma} e_\varrho(\alpha_i, \beta_i).$$

Dies wird mit Hilfe von Satz 2 weiter abgeschätzt.

Wegen (18.8) und (18.17) ist $\Sigma\beta_i \leqq \pi k_\Gamma$, so daß $\frac{\Sigma\beta_i}{3k_\Gamma} \leqq \frac{\pi}{3}$. Die weiteren Voraussetzungen (5.2), ..., (5.5) in Satz 2 sind äquivalent dazu, daß ε_Γ sich zerlegen läßt in eine Menge ε_Γ^2 aus schmalen und höchstens doppelt so vielen breiten g-Dreiecken und in eine Menge ε_Γ^1 aus breiten g-Dreiecken, wobei die breiten g-Dreiecke von ε_Γ^2 Punkten aus $G_1 \cup G_2$ entsprechen, und für ε_Γ^2 der mittlere β-Winkel $\geqq \frac{\pi}{3}$ ist. Dies ist nach Hilfssatz 18.10 richtig.

Aus Satz 2 folgt also

$$2 \sum_{i=1}^{3k_\Gamma} e_\varrho(\alpha_i, \beta_i) \leqq 2 \cdot 3k_\Gamma F\left(\frac{\sum\limits_{i=1}^{3k_\Gamma} \alpha_i}{3k_\Gamma}\right). \tag{18.23}$$

Aus (18.7) und (18.17) folgt $\sum_{i=1}^{3k_\Gamma} \alpha_i \leqq \frac{1}{2} \sum_{L\in\Gamma} \omega_L$, und weil F monoton wachsend ist, folgt (18.22).

Nun gelte das Gleichheitszeichen in (18.22). Dann gilt das Gleichheitszeichen in (18.23), und nach Satz 2 ist $\varepsilon_\Gamma^2 = \emptyset$, und alle g-Dreiecke von ε_Γ^1 sind kongruent mit β-Winkel $\frac{\pi}{3}$ und α-Winkel α so, daß (α, ϱ) ein Punkt der Kurve (4.4) ist.

Aus $\varepsilon_\Gamma^2 = \emptyset$ folgt nach Hilfssatz 18.11, daß ε_Γ aus $\mathcal{D}_\Gamma$ ausschließlich durch Zerlegen geknickter g-Dreiecke entsteht. Beim Zerlegen eines geknickten g-Dreiecks gemäß Hilfssatz 18.1 entsteht ein entartetes breites g-Dreieck mit β-Winkel $\frac{\pi}{2}$. Weil aber alle g-Dreiecke von $\varepsilon_\Gamma^1 = \varepsilon_\Gamma$ einen β-Winkel $\frac{\pi}{3}$ haben, kann $\mathcal{D}_\Gamma$ keine geknickten g-Dreiecke enthalten. Deshalb ist $\varepsilon_\Gamma = \mathcal{D}_\Gamma$, und $k_\Gamma = 1$. $\mathcal{D}_\Gamma$ ist also die Zerlegung eines einzigen g-Dreiecks L, das wegen der Kongruenz der g-Dreiecke von $\mathcal{D}_\Gamma$ gleichseitig ist. Wegen der Interpretation der Kurve (4.4) in Hilfssatz 4.1 gilt für die Systemkreise K_1, K_2, K_3 in den Ecken von L: Es wird $\partial K_i \cap L$ von $\{K_1, K_2, K_3\}$ genau zur Hälfte einfach überdeckt. □

19. Beweis von Satz 1

Der Beweis von Satz 1 unter den Voraussetzungen (1.4), (1.5) und (1.6) geschieht nun vollends mit Hilfssatz 15.1: Aus den Hilfssätzen 16.1, 17.1 und 18.12 folgt, daß (15.8) für jede Zusammenhangskomponente Γ von Γ_S gilt, so daß die Ungleichung (1.2) in Satz 1 richtig ist.

Daß die Abschätzung (1.2) für $n = 4, 6, 12$ scharf ist, wenn die Kreismittelpunkte und Kreisradien sich wie in Satz 1 beschrieben verhalten, wurde noch in **1.** gezeigt.

Umgekehrt gelte nun in (1.2) das Gleichheitszeichen. Nach Hilfssatz 15.1 folgt daraus die Gültigkeit des Gleichheitszeichens in (15.8). Deshalb treten nach den Hilfssätzen 16.1 und 17.1 nur die in **18.** betrachteten Zusammenhangskomponenten Γ auf, und für jedes solche Γ gilt in (18.22) das Gleichheitszeichen. Wegen Hilfssatz 18.12 besteht dann jedes $\mathcal{L}_\Gamma$ aus einem einzigen, gleichseitigen Dreieck L, und für die jeweiligen Systemkreise K_1, K_2, K_3 in den Ecken von L gilt: $\partial K_i \cap L$ wird von $\{K_1, K_2, K_3\}$ genau zur Hälfte einfach überdeckt. Die $\mathcal{L}$-Zerlegung ist somit eine reguläre Zerlegung der Sphäre in Dreiecke, so daß $n = 4, 6, 12$, und die Systemkreismittelpunkte sind die Ecken eines regulären Tetraeders, Oktaeders, bzw. Ikosaeders. Außerdem sind die Kreisradien von der in Satz 1 behaupteten Größe. □

Liste der wichtigsten Bezeichnungen

S^2	Sphäre im R^2
$K_1, \dots, K_n$	Systemkreise
$O_1, \dots, O_n$	Mittelpunkte von Systemkreisen

$E(K_1, \dots, K_n)$	von den Kreisen $K_1, \dots, K_n$ auf S^2 einfach überdeckter Bereich
	Zu einer Menge M ist
$\lvert M \rvert$	Flächeninhalt von M
∂M	Rand von M
M^0	Inneres von M
conv M	Konvexe Hülle von M
	Zu Punkten A, B, C, D ist
AB	Großkreis durch A und B
$\overline{AB}$	(kürzerer der beiden) Großkreisbogen durch A und B
ABC	sphärisches Dreieck mit den Ecken A, B, C
$ABCD$	sphärisches Viereck mit den Ecken A, B, C, D in natürlicher Reihenfolge
$U(ABC)$	Umkreisradius von ABC
$K_\varrho(A)$	Kreis mit Radius ϱ um A
	Bereiche
$G = G(\varrho)$	siehe Fig. 1 und (2.3)
$G_i = G_i(\varrho)$	siehe Fig. 1 und **3**.
$\overline{G} = \overline{G}(\varrho)$	siehe Fig. 2 und **8**.
D_1, D_2, D_3	siehe Fig. 6 und **4**.
$\varrho_0 = \arccos \frac{1}{\sqrt{7}}$	siehe (1.5), Satz 2 und (13.2)
ω_n	siehe (1.3)
	Funktionen
$e_\varrho(\Delta)$	siehe (2.4) und (18.1)
$e_\varrho(\alpha, \beta)$	siehe (2.5)
$\alpha_1(\varrho)$	siehe (4.1) und Fig. 6
$\alpha_2(\varrho)$	siehe (4.2) und Fig. 6
$h_\varrho(\alpha, \beta)$	siehe **7**.
$H_\varrho(\alpha, \beta)$	siehe **9**.
$E_\varrho(\alpha, \beta)$	siehe **10**.
$F(\alpha)$	siehe Satz 1, **4**. und Satz 2
	Bezeichnungen in II
$L_{\dots}$	$\mathcal{L}$-Dreieck, dessen Umkreismittelpunkt $l_{\dots}$ ist
$l_{\dots}$	Umkreismittelpunkt des $\mathcal{L}$-Dreiecks $L_{\dots}$
Γ_S	siehe **15**.
Γ	siehe **15**.
$\mathcal{L}_\Gamma$	siehe **15**.
S_Γ	siehe **15**.
k_Γ	siehe **15**.
$\mathcal{D}_\Gamma$	siehe **18.2**.
$V(\mathcal{D}_\Gamma)$	Verschiebungsoperation in **18.5**.
$Q(\mathcal{D}_\Gamma)$	Quellenoperation in **18.4**.
α-Winkel	siehe **13**.
β-Winkel	siehe **13**.
mittlerer β-Winkel	siehe **18.6**.

LITERATURVERZEICHNIS

[1] BALÁZS, J., Über ein Kreisüberdeckungsproblem, *Acta Math. Acad. Sci. Hungar.* **24** (1973), 377–382. *MR* **48** #9561

[2] BLIND, G. und BLIND, R., Ein Kreisüberdeckungsproblem, *Studia Sci. Math. Hungar.* **21** (1986), 35–57. *MR* **88m**:52023

[3] BLIND, G. und BLIND, R., Über ein Kreisüberdeckungsproblem auf der Sphäre, *Studia Sci. Math. Hungar.* (to appear).

[4] FEJES TÓTH, L., *Lagerungen in der Ebene, auf der Kugel und im Raum,* 2. Auflage, Die Grundlehren der math. Wissenschaften, Band 65, Springer Verlag, Berlin – Heidelberg – New York, 1972. *MR* **50** #5603

[5] GÁSPÁR, Zs. and TARNAI, T., Multisymmetric close packings of equal spheres on the spherical surface, *Acta Cryst. Sect. A* **43** (1987), 612–616. *MR* **89d**:52031

[6] GÁSPÁR, Zs. and TARNAI, T., Covering the sphere with 11 equal circles, *Elem. Math.* **41** (1986), 35–38. *MR* **88c**:52014

[7] MAKAI, E., Research problem 20, *Period. Math. Hungar.* **7** (1976), 319–320.

[8] MOLNÁR, J., On the ϱ-system of unit circles, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **20** (1977), 195–203. *MR* **58** #12733

[9] MOLNÁR, J., Packing of congruent spheres in a strip, *Acta Math. Acad. Sci. Hungar.* **31** (1978), 173–183. *MR* **58** #7406

(Eingegangen am 15. November 1990)

MATHEMATISCHES INSTITUT B
UNIVERSITÄT STUTTGART
PFAFFENWALDRING 57
D–70550 STUTTGART

WALDBURGSTRASSE 88
D–70563 STUTTGART
FEDERAL REPUBLIC OF GERMANY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 165–168

SOME ADDITIONAL PROPERTIES OF SUBGROUPS OF p-GROUPS HAVING SOFT SUBGROUPS

L. HÉTHELYI

Dedicated to S. G.

A subgroup A of a group G of prime power is said to be a soft subgroup of G if A is its own centralizer and if it is maximal in its normalizer. The basic properties were investigated in [2].

In an other paper the subgroups of these p-groups were studied. In this short note we investigate some further properties of subgroups of p-groups having soft subgroups. We shall show that every A-invariant maximal abelian subgroup of M has order at least $|A|$, where M is the unique maximal subgroup of G containing A. We shall also investigate $C_G(G')\cap N_i$ where N_i is the i-th term of the normalizer chain of A.

PROPOSITION 1. *Suppose that G is a p-group and that A is a soft subgroup of G. Let M be the unique maximal subgroup of G containing A. Then every A-invariant maximal abelian subgroup of M has order at least $|A|$.*

PROOF. Let B be an A-invariant maximal abelian subgroup of M. If $\mathrm{cl}(M)=1$ then $B=A$ and the proposition follows. Suppose that $\mathrm{cl}(M)\geqq 2$. By Lemma 2 in [2] there is a natural number k for which $A\cdot B=N_k$, where N_k is the k-th term of the normalizer chain of A.

If $N_k\neq M$ then the proposition follows by induction. So we may assume that $N_k=M$. Then $A\cap B=Z(M)$. Let a bar denote homomorphic images in $G/Z(M)$. Then $\overline{M}=\overline{A}\cdot\overline{B}$ holds. Moreover, $\overline{A}\cap\overline{B}=\overline{1}$, $\overline{N}_1\overline{B}=\overline{M}$ and $\overline{N}_1\cap\overline{B}\leqq Z(\overline{M})$, where $N_1=N_G(A)$. Let $\overline{B}_1\in SC(\overline{M})$ such that $\overline{B}\leqq\overline{B}_1$. Then $|\overline{B}_1|\geqq|\overline{N}_1|$ holds by induction. Thus $|B_1|\geqq|N_1|$ where B_1 is the inverse image in G of $\overline{B}_1$. However, $|N_1|=p|A|$. So it is enough to prove that $|B_1:B|\leqq p$.

However, as $B_1\cap A\leqq B_1\cap N_1\leqq Z_2(M)$ and as $|Z_2(M)\cap A:Z(M)|\leqq p$,

$$|B_1A|=|M|=\frac{|B_1|\,|A|}{|B_1\cap A|}\geqq\frac{|B_1|\,|A|}{p|Z(M)|}.$$

However,

$$\frac{|B|\,|A|}{|Z(M)|}=|M|=\frac{|B_1|\,|A|}{|B_1\cap A|}\geqq\frac{|B_1|\,|A|}{p|Z(M)|}.$$

1991 *Mathematics Subject Classification*. Primary 20D15.

Key words and phrases. Abelian subgroup, normalizer, p-group, chain, characteristic series.

0081–6906/94/$ 4.00

Thus $|B| \geqq \frac{|B_1|}{p}$. □

COROLLARY 1. *Suppose that G is a p-group and that A is a soft subgroup of G. Then there exists an abelian normal subgroup B of G such that $|B| \geqq |A|$.*

COROLLARY 2. *Suppose that G is a p-group and that A is a soft subgroup of G. Let M be the unique maximal subgroup of G containing A. Then for every $k = 1, \ldots, \mathrm{cl}(M)$ there exists a subgroup B of M such that $\mathrm{cl}(B) \leqq k$, $B \triangleleft M$ and $|B| \geqq |N_k|$, where N_k is the k-th term of the normalizer chain of A.*

PROPOSITION 2. *Suppose that G is a p-group and that A is a soft subgroup of G of index at least p^2 in G. Let M be the unique maximal subgroup of G containing A. Let $Z_{n-1}(M)$ be the last nontrivial term of the upper central series of M. Then there is exactly one normal subgroup B of G contained in M such that $Z_{n-1}(M) < B < M$. Moreover, B contains every normal abelian subgroup of G contained in M.*

PROOF. Proposition 2 follows from Theorem 2 of [3] and Proposition 1 and Lemma 2 of [2]. □

PROPOSITION 3. *Suppose that G is a p-group and that A is a soft subgroup of G. Let M be the unique maximal subgroup of G containing A. Then any characteristic series of M beginning with $Z(M)$ can be refined to one with factors of order at most p^2.*

PROOF. It is immediate from Proposition 1 of [2]. □

COROLLARY 3. *Suppose that G is a p-group and that A is a soft subgroup of G. Let M be the unique maximal subgroup of G containing A. Then every p'-subgroup of odd order of $\mathrm{Aut}(M/Z(M))$ is abelian.*

PROPOSITION 4. *Suppose that G is a p-group and that A is a soft subgroup of G. Let M be the unique maximal subgroup of G containing A and let N_i be the i-th term of the normalizer chain of A. Let $L_i = C_M(G') \cap N_i$. Then $L_i \triangleleft G$. Moreover, $|L_i : L_{i-1}| = p$ for $i = 1, \ldots, s$ where $N_s = C_M(G') \cdot A$ is the s-th term of the normalizer chain of A.*

PROOF. It is easy to see that $L_i \triangleleft N_i$ and as $[G', L_i] = 1$, $L_i \triangleleft M$. By Lemma 2 in [2] there are natural numbers j and s such that $L_i A = N_j$ and $C_G(G') \cdot A = N_s$. Moreover, $|N_s : A| = |C_M(G') : Z(M)|$ and there exists a series $Z(M) = U_1 < U_2 < \cdots < U_t = C_G(G')$ such that every U_i is normal in G and that $|U_i : U_{i-1}| = p$ for $i = 2, \ldots, t$. However, $U_i A \neq U_{i-1} A$. Otherwise, if $U_i A = N_m$, then $|U_i : Z(M)| = |N_m : A| = |U_{i-1} : Z(M)|$ would follow, which is a contradiction. Moreover, as $|L_j A : A| = |L_j : Z(M)|$, $L_j = U_l$ for some l. As $U_i A \neq U_j A$ $(i \neq j)$ $L_i = U_i$ follows for all i. Then $L_i \triangleleft G$ and $|L_i : L_{i-1}| = p$. □

COROLLARY 4. *Suppose that G is a p-group and that A is a soft subgroup of G. Let M be the unique maximal subgroup of G containing A. Let $N_1, \ldots, N_n$ be the normalizer chain of A. If $N_i' \leq Z(G')$ then $N_i' Z(M) \triangleleft G$.*

PROOF. Let L_i have the same meaning as in the proof of the previous proposition. Then as $N_i' \leq Z(G')$, $Z(M)\, N_i' \leq L_{i-1}$. However,

$$|L_{i-1} : Z(M)| = |N_{i-1} : A| = |N_i'\, Z(M) : Z(M)|$$

and so

$$L_{i-1}/Z(M) \cong N_i'\, Z(M)/Z(M).$$

Thus $L_{i-1} = N_i'\, Z(M)$. □

PROPOSITION 5. *Suppose that G is a p-group and that A is a soft subgroup of G. Let M be the unique maximal subgroup of G containing M. Suppose that G' is abelian. Let a bar denote homomorphic images in $G/Z(M)$. Then there exists an $\overline{a} \in \overline{G}$ such that $|\overline{G} : C_{\overline{G}}(\overline{a}) = |\overline{G}'|$.*

PROOF. As $C_M(G') \cdot A = M$, $|\overline{G}'| = |\overline{M} : \overline{A}| = |\overline{G} : \overline{N}_1|$, where $N_1 = N_G(A)$. However, there is an a for which $C_{\overline{G}}(\overline{a}) = \overline{N}_1$ and so $|\overline{G} : C_{\overline{G}}(\overline{a})| = |\overline{G}'|$. □

PROPOSITION 6. *Suppose that G is a p-group containing a soft subgroup A. Let M be the unique maximal subgroup of G containing A. Let a be an element of A such that $C_G(a) = A$. Let $v \in G \backslash M$. Let $L = \langle a, v \rangle$. Let $T_i = L \cdot Z_i(M)$ where $Z_i(M)$ is the i-th term of the upper central series of M. Let a bar denote homomorphic images in $G/Z_i(M)$. Then $\overline{T}_i' = \overline{T}_{i+1}'$ for $i = 1, \ldots, \mathrm{cl}(M) - 1$.*

PROOF. It is easy to see that $|T_{i+1} : T_i| \leq p$ and that $|Z(\overline{T}_{i+1})| = p$. Let $\widehat{\ }$ denote homomorphic images in $\overline{T}_{i+1}/Z(\overline{T}_{i+1})$. Then $\widehat{T}_{i+1} = \widehat{T}_i \times \widehat{Z_{i+1}(M)}$ and $\widehat{Z_{i+1}(M)} \leq Z(\widehat{T}_{i+1})$. Thus $\widehat{T}_{i+1}' = \widehat{T}_i'$. However, $Z(\overline{T}_{i+1}) = Z(\overline{T}_i)$. Thus the last nontrivial term of the lower central series of $\overline{T}_{i+1}$ and $\overline{T}_i$ coincide. Thus $\overline{T}_i' = \overline{T}_{i+1}'$. □

PROPOSITION 7. *Suppose that G is a p-group and that A is a soft subgroup of G. Let M denote the unique maximal subgroup of G containing A. Let C be a critical subgroup of M. Then $|C : Z(C)| \leq p^2$.*

PROOF. Let $N_k = C \cdot A$. Then as $[C, A] = N_k' \leq Z(C)$, $Z(C) \cdot A$ is of index p in N_k. Moreover, $Z(N_k) = Z(C) \cap A$. Let a bar denote homomorphic images in $N_k/Z(N_k)$. Then $|Z(\overline{N}_k)| \leq p^2$. Moreover, $\overline{A} \cap \overline{Z(C)} = 1$.

As $|\overline{N}_k : \overline{A} \cdot \overline{Z(C)}| = p$, $|\overline{A} \cap \overline{C}| \leq p$. Moreover, as $[C, A, C] = 1 = [A, C, C]$, $[C, C, A] = 1$ by the Three Subgroup Lemma. Then $C' \leq A \cap Z(C) = Z(N_k)$. Thus $\overline{A} \cap \overline{C} \leq Z(\overline{N}_k)$. Thus as $|Z(\overline{N}_k) \cap \overline{A}| \leq p$, $|\overline{A} \cap \overline{C}| \leq p$. Moreover, as

$$p|\overline{Z(C)} \cdot \overline{A}| = p|\overline{Z(C)}|\, |\overline{A}| = |\overline{C} \cdot \overline{A}| \geq \frac{|\overline{C}|\, |\overline{A}|}{p},$$

$|\overline{C}:\overline{Z(C)}| \leqq p^2$ follows. Thus $|C:Z(C)| \leqq p^2$. □

Finally, we make an observation about the embedding of M in an other p-group.

DEFINITION. Let G be a p-group, A a soft subgroup of G. Let M be the unique maximal subgroup of G containing A. Let G_1 a p-group. M is well-embedded in G_1 if M contains a soft subgroup of G_1.

PROPOSITION 8. *Suppose that G is a p-group and A is a soft subgroup of G of index at least p^4. Let M be the unique maximal subgroup of G containing A. Let M be well-embedded in G_1. Let B be a soft subgroup of G_1 contained in M. Then B is a soft subgroup of G.*

PROOF. Let $R(G)=G'\cdot Z(N_G(A))$. Then $R(G)$ is characteristic in M by Theorem 4 of [3]. Thus $B \nleqq R(G)$. Thus B is soft in G by Theorem 2 of [3]. □

REFERENCES

[1] HUPPERT, B., *Endliche Gruppen* I, Die Grundlehren der math. Wissenschaften, Bd. 134, Springer–Verlag, Berlin–New York, 1967. *MR* **37** #302

[2] HÉTHELYI, L., Soft subgroups of p-groups, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **27** (1984), 81–85. *MR* **87c**:20044

[3] HÉTHELYI, L., On subgroups of p-groups having soft subgroups, *J. London Math. Soc.* (2) **41** (1990), 425–437. *MR* **91i**:20021

(Received December 4, 1990)

BUDAPESTI MŰSZAKI EGYETEM
VEGYÉSZMÉRNÖKI KAR
MATEMATIKA TANSZÉK
EGRI JÓZSEF U. 1.
H–1521 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 169–175

ON THE COORDINATES OF MINIMUM VECTORS IN n-LATTICES

Á. G. HORVÁTH

Abstract

In this paper the following problem will be discussed: How to find such a basis of an n-lattice in E^n that with respect to this basis the absolute values of the coordinates belonging to the minima of this lattice are "small enough". We prove that in every lattice possessing n linearly independent minima one can find such a basis for which the maximum of the absolute values of the coordinates belonging to a minimum vector is not greater than the maximum of the indices of the admissible centerings of the n-dimensional lattices. This result is not sharp, we prove that in the lower dimensional cases, where $n \leqq 5$ in every lattice with n linearly independent minima there exists a basis for which all the coordinates of the minima are equal to -1, 0 or 1; and in the cases $n = 4$ and 5 the maximal admissible index is equal to 2.

1. Definitions

A lattice in E^n defined by the basis $A = \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ of E^n is the set $L = \sum_{i=1}^{n} x_i \mathbf{a}_i$ of all integral linear combinations of the basis A. A minimum vector (or minimum) of L is one of the shortest non-zero vectors in L. A minimum basis A of L is then such a basis in L for which all the basis vectors $\mathbf{a}_i$ are minimum vectors of L. The common length of the minima is denoted by $\min L$. Let us define now the numbers $L(A)$ and L_n in the following way:

$$L(A) = L(\{\mathbf{a}_1, \dots, \mathbf{a}_n\}) := \max\Big\{|x_i|; \mathbf{m} = \sum_{i=1}^{n} x_i \mathbf{a}_i \in L, |\mathbf{m}| = \min L\Big\} \tag{1}$$

$$L_n := \sup\Big\{\min\{L(A) \mid A \text{ is a basis of } L\} \mid L \subset E^n \text{ is a lattice possessing } n \text{ independent minima}\Big\}. \tag{2}$$

1991 *Mathematics Subject Classification.* Primary 52C07; Secondary 11H50.
Key words and phrases. Lattice, minimum vector basis.

Supported by the Hungarian National Found for Scientific Research Grant No. 1615 (1991).

0081–6906/94/$ 4.00 ©1994 *Akadémiai Kiadó, Budapest*

We will use some results from the theory on admissible centerings of n-lattices (see [5], [2]) so we have to recall some further definitions.

The lattice L' is a centering of the lattice L if $L' \supset L$. This centering is admissible iff $\min L' = \min L$. The index of the admissible centering is defined by the number $\operatorname{ind}(L'/L) = v(L)/v(L')$, where $v(L)$ is the volume of a basic parallelepiped in the lattice L. Let V_n be the maximum of the n-dimensional indices for all admissible centerings of all lattices L possessing n linearly independent minima. An important task in the geometry of numbers is to give estimations for the value of V_n. The first results in this field are due to A. Korkine–G. Zolotareff (see [4/a], [4/b]) and H. Minkowski [6]. Later a good estimation was given by Davenport and Watson [2], namely they proved that $V_n \leqq c_n^{n/2}$ where c_n denotes the Hermite constant defined as $\max\{(\min L)^2 \mid L \subset E^n$ is a lattice of determinant $1\}$. There holds $c_n \leqq$ $\leqq 2^{-0.198\ldots}\frac{n}{\pi e}(1+o(1))$ cf. the upper estimate of the packing density of balls in E^n in [3]. In the lower dimensional cases also the exact values of V_n are known due to the results of S. S. Ryškov [7] and N. V. Zaharova-Novikova [8] who discussed the cases $n \leqq 7$ and $n = 8$, respectively.

2. The theorem

THEOREM. *For an arbitrary dimension n $L_n \leqq V_n$ holds, i.e. in every n-lattice possessing n linearly independent minima there exists a basis such that the absolute values of the coordinates belonging to any minimum vector are not greater than V_n.*

PROOF. Let $\pm\mathbf{m}_1, \ldots, \pm\mathbf{m}_\sigma$ be all different minima of the lattice. It is well known that $\sigma \leqq 2^n - 1$ is valid (see [9]). Suppose that the positive volume of the parallelepiped $\pi(\mathbf{m}_1, \ldots, \mathbf{m}_n)$ is not less than the volume of any other n-dimensional parallelepiped spanned by the minimum vectors of L. Then

$$m_l = \sum_{j=1}^{n} \alpha_{lj}\mathbf{m}_j, \qquad \text{where } 1 \leqq l \leqq \sigma \tag{3}$$

and for the rational numbers α_{lj} the following hold:

$$|\alpha_{lj}| \leqq 1, \qquad j = 1, \ldots, n \text{ and } 1 \leqq l \leqq \sigma. \tag{4}$$

(Assuming the contrary we get such a minimum parallelepiped which has a volume greater than that of the parallelepiped $\pi(\mathbf{m}_1, \ldots, \mathbf{m}_n)$.) Consider now such a basis $A = \{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ of L in which $\mathbf{m}_1, \ldots, \mathbf{m}_n$ can be expressed as $\mathbf{m}_j = \sum_{i=1}^{n} v_{ji}\mathbf{a}_i$, where the coordinates v_{ji} satisfy the following inequalities:

$$\begin{array}{ll} \text{(i)} & v_{jj} > 0, \quad j = 1, \ldots, n \qquad v_{ji} = 0 \quad \text{for } 1 \leqq j < i \leqq n \\ \text{(ii)} & 0 \leqq v_{ji} < v_{jj} \quad \text{for } 1 \leqq i < j \leqq n. \end{array} \tag{5}$$

The existence of such a basis A is assured in every n-lattice (see [1]). With respect to this basis A the vectors $\mathbf{m}_1, \ldots, \mathbf{m}_\sigma$ can be expressed in the following form:

$$\mathbf{m}_l = \sum_{j=1}^{n} \alpha_{lj}\mathbf{m}_j = \sum_{i=1}^{n}\Big(\sum_{j=i}^{n} \alpha_{lj}v_{ji}\Big)\mathbf{a}_i, \qquad l = 1, \ldots, \sigma. \tag{6}$$

So the absolute values of the coordinates are:

$$p_{li} = \Big|\sum_{j=i}^{n} \alpha_{lj}v_{ji}\Big|. \tag{7}$$

If for any j $(j > i)$ v_{jj} is equal to one then on the base of (5)(ii), we get $v_{ji} = 0$, so for this reason we have:

$$p_{li} \leqq \sum_{\substack{j=i+1\\ v_{jj}>1}}^{n} |\alpha_{lj}|v_{ji} + |\alpha_{li}|v_{ii} \leqq \sum_{\substack{j=i+1\\ v_{jj}>1}}^{n} v_{jj} + v_{ii} \leqq \prod_{j=i}^{n} v_{jj} + 1. \tag{8}$$

In the last step we used that for a finite set of integers each greater than 1 their sum is at most their product, where equality holds only if the set consists of one element. If $v_{ii} = 1$ and there are at least two v_{jj}'s with $i+1 \leqq j \leqq n$, $v_{jj} > 1$ then the sum of these v_{jj}'s is strictly less than their product, so in fact

$$\sum_{\substack{j=i+1\\ v_{jj}>1}}^{n} v_{jj} + v_{ii} \leqq \prod_{j=i}^{n} v_{jj}.$$

The same holds if there is no j such that $i+1 \leqq j \leqq n$, $v_{jj} > 1$. If there is just one such v_{jj}, then in (8) we can use for this j the sharper estimate $|\alpha_{lj}|v_{ji} \leqq v_{ji} < v_{jj}$, hence $|\alpha_{lj}|v_{ji} \leqq v_{jj} - 1$, which gives

$$p_{li} \leqq \prod_{j=i}^{n} v_{jj}. \tag{9}$$

But $\prod_{j=i}^{n} v_{jj}$ is the index of an admissible centering, which completes the proof of the Theorem.

3. The case of the lower dimensions

In this paragraph we prove that the Theorem is not “sharp”. It is well known that $V_1 = V_2 = V_3 = 1$ and $V_4 = V_5 = 2$ (see [7]). We verify two statements:

STATEMENT 1. $L_4 = 1$.

PROOF. We will distinguish two cases.

1. If in the lattice L every linearly independent minimum system is a basis then the absolute values of the examined coordinates are less than two. (This is clear from the proof of the Theorem.)

2. If the lattice L has a minimum system with index 2, then the minima of the lattice can be expressed with respect to the basis constructed above (in the proof of the Theorem), in the following way (where $\mathbf{m}_1$, $\mathbf{m}_2$, $\mathbf{m}_3$, $\mathbf{m}_4$ are the edge vectors of the basic cube; at this time the lattice is the well-known space-centred cubic lattice, see e.g. in [7])

$$(10)\quad \begin{array}{c} \mathbf{m}_1 = \begin{bmatrix}1\\0\\0\\0\end{bmatrix} \quad \mathbf{m}_2 = \begin{bmatrix}0\\1\\0\\0\end{bmatrix} \quad \mathbf{m}_3 = \begin{bmatrix}0\\0\\1\\0\end{bmatrix} \quad \mathbf{m}_4 = \begin{bmatrix}1\\1\\1\\2\end{bmatrix} \quad \mathbf{m}_5 = \begin{bmatrix}0\\0\\0\\1\end{bmatrix} \quad \mathbf{m}_6 = \begin{bmatrix}1\\0\\0\\1\end{bmatrix} \\ \mathbf{m}_7 = \begin{bmatrix}0\\1\\0\\1\end{bmatrix} \quad \mathbf{m}_8 = \begin{bmatrix}0\\0\\1\\1\end{bmatrix} \quad \mathbf{m}_9 = \begin{bmatrix}0\\1\\1\\1\end{bmatrix} \quad \mathbf{m}_{10} = \begin{bmatrix}1\\0\\1\\1\end{bmatrix} \quad \mathbf{m}_{11} = \begin{bmatrix}1\\1\\0\\1\end{bmatrix} \quad \mathbf{m}_{12} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \end{array}$$

and so the characteristic matrix (see [7]) of the minima can be written in the following form:

$$(11)\quad \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

But a row-subtraction operation on this matrix is equivalent to a basis-change of the lattice so we get that with respect to a suitable basis the minima of L can be written in the following way:

$$(12)\quad \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ -1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix},$$

so L_4 is equal to one. □

STATEMENT 2. *L_5 is equal to one, too.*

PROOF. Suppose $v(L) = 1$. If the volumes of the minimum parallelepipeds are 1 or -1, then the elements of the characteristic matrix are also 0, 1 or -1. So we can assume that $v(\pi(\mathbf{m}_1, \dots, \mathbf{m}_5)) = 2$. Then the elements of the characteristic matrix have absolute value at most 2. We distinguish two cases:

1. If the lattice does not have a four-dimensional space-centred cubic sublattice (in this we include that the edge vectors of the cube are minima of the lattice and this sublattice is the intersection of the lattice and a 4-plane), then the setup of the characteristic matrix can be started in the following way:

$$(13)\qquad \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & \dots \\ 0 & 1 & 0 & 0 & 1 & \dots \\ 0 & 0 & 1 & 0 & 1 & \dots \\ 0 & 0 & 0 & 1 & 1 & \dots \\ 0 & 0 & 0 & 0 & 2 & \dots \end{bmatrix}.$$

Let $\mathbf{m}_l$ be an arbitrary minimum vector, where $5 < l \leqq \sigma$. Then the coordinates of this minima can be seen from (6), we have:

$$(14)\qquad m_{il} = \sum_{j=i}^{n} \alpha_{lj} v_{ji} = \sum_{j=i}^{5} \alpha_{lj} v_{ji},$$

where $v_{11} = \dots = v_{44} = v_{5i} = 1$ if $1 \leqq i < 5$, $v_{55} = 2$ and the other v_{ji} are equal to zero. So we get the following simple equalities:

$$(15)\qquad m_{il} = \alpha_{li} + \alpha_{l5}, \qquad m_{5l} = 2\alpha_{l5}.$$

Clearly we may assume $m_{5l} \geqq 0$. First assume that $m_{5l} = 2$, hence $\alpha_{l5} = 1$. At this time we have the possibility for the choice of the values of the other coordinates 0, 1, 2, respectively. It is clear that we do not have a minimal vector in the lattice mod (2) equivalent to one of the vectors $\mathbf{m}_1, \dots, \mathbf{m}_5$. So the number of the 1's among the first four coordinates is two or three. Now we examine two cases:

(a) If there is a zero among the first 4 coordinates;
then there is a sublattice L_1 of L which is a space-centred cubic 4-sublattice in L (e.g. if $m_{1l} = 0$, then the vectors $\mathbf{m}_2$, $\mathbf{m}_3$, $\mathbf{m}_4$, $\mathbf{m}_l$ form a minimum system with index 2, and the sublattice L_1 is spanned by the vectors $\mathbf{a}_2, \dots, \mathbf{a}_5$).

(b) If there are two or three 1's among the first 4 coordinates and the others are equal to 2;
then also there is a space-centred cubic 4-sublattice of the lattice L, for example if $\mathbf{m}_l = [2,1,1,1,2]^T$ or $\mathbf{m}_l = [2,2,1,1,2]^T$ then the sublattice spanned by the vectors $\mathbf{a}_2$, $\mathbf{a}_3$, $\mathbf{a}_4$, $\mathbf{a}_1 + \mathbf{a}_5$ contains the minimum-parallelepiped $\pi(\mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4, \mathbf{m}_l)$ with volume two. So this sublattice is a space-centred cubic lattice, too.

From this reason if $m_{5l} = 2$ then $\mathbf{m}_l = \mathbf{m}_5$.

Secondly we assume that $m_{5l} = 1$. At this time $\alpha_{l5} = \frac{1}{2}$ and the coordinates m_{il}, $1 \leqq i < 5$, are equal to zero or one.

Lastly if $m_{5l} = 0$ it may be seen that the other coordinates of this minimum are -1, 0 or 1 (e.g. if $m_{1l} = 2$, then the vectors $\mathbf{m}_2$, $\mathbf{m}_3$, $\mathbf{m}_4$, $\mathbf{m}_l$ form a minimum system of index 2 in the sublattice spanned by the vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$, $\mathbf{e}_4$).

So in the case of 1 the characteristic matrix can be written in the following form:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} \\ \\ A \\ \\ \\ \end{bmatrix} \begin{bmatrix} \\ \\ A' \\ \\ \\ \end{bmatrix}, \tag{16}$$

where A is a $(0, \pm 1)$ matrix in which the elements of the last row are equal to zero, and the matrix A' is a $(0,1)$ one. It can easily be seen by the subtraction of the first row from the last one that this matrix is equivalent to a $(0, \pm 1)$ one.

2. Consider now the case that the lattice L has a space-centred cubic 4-sublattice. Then the characteristic matrix is the following:

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} & A & \\ & & \\ & & \\ & B & \\ 1 & \cdots & 1 \end{bmatrix}, \tag{17}$$

where A has three rows and B is a one-row vector. If an element of B is equal to 2 resp. -2, then B does not have any negative resp. positive coordinate. In fact, otherwise the columns of the characteristic matrix containing the two mentioned elements of B together with $\mathbf{m}_1$, $\mathbf{m}_2$, $\mathbf{m}_3$ form a submatrix with determinant of absolute value greater than 2. So, if B is not a $(0, \pm 1)$ vector, we may assume that the coordinates of B are either positive or zero resp. either negative or zero. Subtracting the last row from, resp. adding the last row to the row containing B we get that B becomes a $(0, \pm 1)$ vector. Therefore we may assume B is a $(0, \pm 1)$ vector. Assume that A has such an element (for example in the first row) whose absolute value is greater than one. But this element must be a coordinate of a minimum vector $\mathbf{m} = [x_1\, x_2\, x_3\, x_4\, 1]^T$ so the parallelepiped $\pi = \pi[\mathbf{m}, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4, \mathbf{m}_5]$ is a minimum one. But after doing the suitable column-subtractions we have

$$v(\pi) = \left| \det \begin{bmatrix} x_1 & 0 & 0 & 1 & 0 \\ x_2 & 1 & 0 & 1 & 0 \\ x_3 & 0 & 1 & 1 & 0 \\ x_4 & 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix} \right| = |2x_1 - x_4| > 2, \tag{18}$$

where x_4 is 0 or ± 1 and this is a contradiction. For this reason A is a $(0, \pm 1)$ matrix. Thus the only element of the characteristic matrix, which is not 0 or ± 1 is $m_{44} = 2$. Subtract now the first row from the fourth row. Then we get a $(0, \pm 1)$ matrix, unless some minimum vector $\mathrm{m} = [x_1\, x_2\, x_3\, x_4\, 1]^T$ has $x_4 = -x_1 = \pm 1$. However, this case leads to a contradiction by (18). So we have verified this case and the statement, too. □

REMARK. The statement of the theorem is interesting in case of the well-known reduced bases (Minkowski, Hermite, Korkine–Zolotareff) only weaker statements are expected.

REFERENCES

[1] CASSELS, J. W. S., *An introduction to the geometry of numbers*, Die Grundlehren der math. Wissenschaften, Bd. 99, Springer-Verlag, Berlin, 1959. *MR* **28** #1175

[2] DAVENPORT, H. and WATSON, G. L., The minimal points of a positive definite quadratic form, *Mathematika* **1** (1954), 14–17. *MR* **16** - 18

[3] FEJES TÓTH, G., New results in the theory of packing and covering, *Convexity and its applications*, Birkhäuser, Basel–Boston, 1983, 318–359. *MR* **85i**: 52007

[4/A] KORKINE, A. and ZOLOTAREFF, G., Sur les formes quadratiques, *Math. Ann.* **6** (1873), 366–390. *Jb. Fortschritte Math.* **5**, 109

[4/B] KORKINE, A. and ZOLOTAREFF, G., Sur les formes quadratiques positives, *Math. Ann.* **11** (1877), 242–292. *Jb. Fortschritte Math.* **9**, 139

[5] GRUBER, P. M. and LEKKERKERKER, C. G., *Geometry of numbers*, 2nd ed., North-Holland Mathematical Library, Vol. 37, North-Holland, Amsterdam, 1987. *Zbl* **611** #10017. See also *Zbl* **198**, 380

[6] MINKOWSKI, H., Diskontinuitätsbereich für arithmetische Äquivalenz, *J. Reine Angew. Math.* **129** (1905), 220–274. *Jb. Fortschritte Math.* **37**, 251

[7] RYŠKOV, S. S., On the problem of determining perfect quadratic forms of several variables, Number theory, mathematical analysis and their applications, *Trudy Mat. Inst. Steklov* **142** (1976), 215–239, 270–271 (in Russian). *MR* **58** #27807

[8] ZAKHAROVA, N. V., Centerings of eight-dimensional lattices that preserve a frame of successive minima, Geometry of positive quadratic forms, *Trudy Mat. Inst. Steklov* **152** (1980), 97–123, 237 (in Russian). *MR* **82k**:10033 Correction: Novikova, N. V., Three admissible centerings of eight-dimensional lattices, Deposited in VINITI, No. 4842-81 Dep., 1981, I. 8 (in Russian).

[9] VORONOÏ, G., Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Premier Mémoire: Sur quelques propriétés des formes quadratiques positives parfaites, *J. Reine Angew. Math.* **133** (1908), 97–156. *Jb. Fortschritte Math.* **38**, 261

(Received August 15, 1990)

BUDAPESTI MŰSZAKI EGYETEM
GÉPÉSZMÉRNÖKI KAR
GEOMETRIA TANSZÉK
EGRI JÓZSEF U. 1. H. ÉP. II. 22
H-1521 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 177–188

ASYMPTOTIC BEHAVIOUR OF THE DISTRIBUTION FUNCTION OF MULTIVARIATE NONHOMOGENEOUS RENEWAL PROCESSES

S. P. NICULESCU and E. OMEY

Abstract

The Multivariate Central Limit Theorem for independent random vectors is used to obtain information about some limit distributions in multivariate nonhomogeneous renewal theory.

1. Introduction

Let $\mathbf{X}_n = (X_{n1}, \dots, X_{nk})$, $n \in \mathbb{N} = \{1, 2, \dots\}$, be a sequence of independent random vectors with values in $\mathbf{R}_+^k$, where $\mathbf{R}_+ = (0, \infty)$. For every $\mathbf{t} = (t_1, \dots, t_k) \in \mathbf{R}_+^k$ define $\mathbf{N}(\mathbf{t}) = (N_1(t_1), \dots, N_k(t_k))$, where $N_i(t_i) = \sup\{n : X_{1i} + \dots + X_{ni} \leqq t_i\}$, $i \in \{1, \dots, k\}$, with the convention that the supremum over an empty set is 0. The process $\{\mathbf{N}(\mathbf{t}), \mathbf{t} \in \mathbf{R}_+^k\}$ is the multivariate renewal process associated with $\{\mathbf{X}_n, n \in \mathbb{N}\}$.

The multivariate Central Limit Problem for independent random vectors is investigated in Section 2. The results are then used in Sections 3 and 4 to examine the asymptotic distribution of univariate and multivariate renewal processes. In the i.i.d. case similar problems have been previously considered by Ahmad [1], Csenki [3], Hunter [6], Mihoc and Niculescu [7, Chapter 9], Niculescu [8], Niculescu and Omey [10].

2. Multivariate Central Limit Theorem type results for independent random vectors

The celebrated *Lindeberg condition*, as it is presented for instance in Chow and Teicher [2, p. 290], is the necessary and sufficient condition for a sequence of one-dimensional random variables to belong to the domain of attraction of the standard normal law. In more dimensions we obtain

PROPOSITION 1. *Let* $\{A_{ni}, n \in \mathbb{N}\} \subset \mathbf{R}_+$, $i \in \{1, \dots, k\}$, *be* k *increasing sequences, and let* $\{\mathbf{X}_n, n \in \mathbb{N}\}$ *be a sequence of independent* $\mathbf{R}^k$*-valued random vectors. In order that there exist sequences* $\{B_{ni}, n \in \mathbb{N}\} \subset \mathbf{R}$,

1980 *Mathematics Subject Classification.* Primary 60F05; Secondary 60K05.
Key words and phrases. Weak convergence, nonhomogeneous renewal process.

$i \in \{1, \ldots, k\}$, such that

$$(2.1)\qquad \left(\frac{\sum_{s=1}^{n} X_{s1} - B_{n1}}{A_{n1}}, \ldots, \frac{\sum_{s=1}^{n} X_{sk} - B_{nk}}{A_{nk}}\right) \Longrightarrow \mathbf{V}$$

where $\mathbf{V}$ has a multivariate normal distribution $\Phi_{O,\Gamma}$ with standard normal distributed marginals, it is necessary and sufficient that the following conditions are satisfied:

$$(2.2)\qquad \text{for all } x \in \mathbb{R}_+ \text{ and } i \in \{1, \ldots, k\} \quad \lim_{n\to\infty} \sum_{s=1}^{n} P\{|X_{si}| > A_{ni}x\} = 0;$$

$$(2.3)\qquad \text{for all } x \in \mathbb{R}_+ \text{ and } i \in \{1, \ldots, k\} \quad \lim_{n\to\infty} \sum_{s=1}^{n} \left(E\left\{\frac{X_{si}^2}{A_{ni}^2} 1_{(|X_{si}| \leqq A_{ni}x)}\right\} - \left(E\left\{\frac{X_{si}}{A_{ni}} 1_{(|X_{si}| \leqq A_{ni}x)}\right\}\right)^2 \right) = 1;$$

$$(2.4)\qquad \text{for all } x, y \in \mathbb{R}_+ \text{ and } i, j \in \{1, \ldots, k\},\ i \neq j,$$
$$\lim_{n\to\infty} \sum_{s=1}^{n} \left(E\left\{\frac{X_{si}}{A_{ni}} \frac{X_{sj}}{A_{nj}} 1_{(|X_{si}| \leqq A_{ni}x, |X_{sj}| \leqq A_{nj}y)}\right\} - \right.$$
$$\left. - E\left\{\frac{X_{si}}{A_{ni}} 1_{(|X_{si}| \leqq A_{ni}x)}\right\} E\left\{\frac{X_{sj}}{A_{nj}} 1_{(|X_{sj}| \leqq A_{nj}y)}\right\} \right) = \Gamma_{ij}.$$

PROOF. The result follows easily by using the same methods as in de Haan, Omey and Resnick [5]. □

Specializing Proposition 1 to the zero mean case we obtain

PROPOSITION 2. *Let $\{\mathbf{Z}_n = (Z_{n1}, \ldots, Z_{nk}), n \in \mathbb{N}\}$ be a sequence of independent random vectors in $\mathbb{R}^k$ such that $E(Z_{ni}) = 0$, $E(Z_{ni}^2) < \infty$, $i \in \in \{1, \ldots, k\}$, $n \in \mathbb{N}$. For every $i \in \{1, \ldots, k\}$ and $n \in \mathbb{N}$ let $A_{ni}^2 = \sum_{s=1}^{n} E(Z_{si}^2)$. Then (2.1) holds with $X_{ni} = Z_{ni}$ and $B_{ni} = 0$, $i \in \{1, \ldots, k\}$, $n \in \mathbb{N}$, if and only if, as $n \to \infty$,*

$$(2.5)\qquad \frac{1}{A_{ni}^2} \sum_{s=1}^{n} E\left\{Z_{si}^2 1_{(|Z_{si}| \leqq A_{ni}x)}\right\} \to 1$$

for all $i \in \{1, \ldots, k\}$ and $x \in \mathbb{R}_+$; and

$$\frac{1}{A_{ni}A_{nj}}\sum_{s=1}^{n} E\left\{Z_{si}Z_{sj}1_{(|Z_{si}|\leqq A_{ni}x,|Z_{sj}|\leqq A_{nj}y)}\right\}\to\Gamma_{ij} \tag{2.6}$$

for all $i,j\in\{1,\dots,k\}$, $i\neq j$ *and* $x,y\in\mathbb{R}_+$.

PROOF. Via Proposition 1 it remains to show that (2.4) holds true if and only if the corresponding (2.6) holds true. Now, since $E(Z_{si})=0$, from (2.5), as $n\to\infty$, we deduce

$$0\leqq\left|\frac{1}{A_{ni}A_{nj}}\sum_{s=1}^{n}E\{Z_{si}1_{(|Z_{si}|\leqq A_{ni}x)}\}E\{Z_{sj}1_{(|Z_{sj}|\leqq A_{nj}y)}\}\right|=$$
$$=\left|\frac{1}{A_{ni}A_{nj}}\sum_{s=1}^{n}E\{Z_{si}1_{(|Z_{si}|>A_{ni}x)}\}E\{Z_{sj}1_{(|Z_{sj}|\leqq A_{nj}y)}\}\right|\leqq$$
$$\leqq\frac{y}{A_{ni}}\sum_{s=1}^{n}E\{|Z_{si}|1_{(|Z_{si}|>A_{ni}x)}\}\leqq\frac{y}{xA_{ni}^2}\sum_{s=1}^{n}E\{Z_{si}^2 1_{(|Z_{si}|>A_{ni}x)}\}\to 0$$

whatever $x,y\in\mathbb{R}_+$, where we used the Lindeberg condition, which is equivalent with the marginal convergence in (2.1). □

In general we obtain

THEOREM 1. *Let* $\{\mathbf{X}_n, n\in\mathbb{N}\}$ *be a sequence of independent random vectors with values in* $\mathbb{R}^k$ *such that* $\mu_{ni}=E\{X_{ni}\}\in\mathbb{R}$ *and*

$$A_{ni}^2=\sum_{s=1}^{n}E\{(X_{si}-\mu_{si})^2\}\in\mathbb{R}_+,\quad i\in\{1,\dots,k\},\quad n\in\mathbb{N}.$$

If for all $x\in\mathbb{R}_+$ *and* $i\in\{1,\dots,k\}$

$$\lim_{n\to\infty}\frac{1}{A_{ni}^2}\sum_{s=1}^{n}E\left\{X_{si}^2 1_{(|X_{si}|>A_{ni}x)}\right\}=0; \tag{2.7}$$

for all $x,y\in\mathbb{R}_+$ *and* $i,j\in\{1,\dots,k\}$, $i\neq j$,

$$\lim_{n\to\infty}\frac{1}{A_{ni}A_{nj}}\sum_{s=1}^{n}E\left\{|X_{si}X_{sj}|1_{(|X_{si}|>A_{ni}x;|X_{sj}|>A_{nj}y)}\right\}=0; \tag{2.8}$$

for all $i,j\in\{1,\dots,k\}$, $i\neq j$

$$\lim_{n\to\infty}\frac{1}{A_{ni}A_{nj}}\sum_{s=1}^{n}E\{(X_{si}-\mu_{si})(X_{sj}-\mu_{sj})\}=\Gamma_{ij}; \tag{2.9}$$

then (2.1) *holds true with* $B_{ni} = \sum_{s=1}^{n} \mu_{si}$ *and* A_{ni} *as above.*

PROOF. Let $Z_{ni} = X_{ni} - \mu_{ni}$, $i \in \{1, \dots, k\}$, $n \in \mathbb{N}$. We show that the assumptions in Proposition 2 are satisfied.

First observe that

$$\frac{1}{A_{ni}^2} \sum_{s=1}^{n} E\left\{(X_{si} - \mu_{si})^2 1_{(|X_{si}-\mu_{si}|>A_{ni}x)}\right\} =$$

$$= \frac{1}{A_{ni}^2} \sum_{s=1}^{n} E\{X_{si}^2 1_{(|X_{si}-\mu_{si}|>A_{ni}x)}\} + \frac{1}{A_{ni}^2} \sum_{s=1}^{n} \mu_{si}^2 P(|X_{si} - \mu_{si}| > A_{ni}x) -$$

$$- \frac{2}{A_{ni}^2} \sum_{s=1}^{n} \mu_{si} E\left\{X_{si} 1_{(|X_{si}-\mu_{si}|>A_{ni}x)}\right\} = I_1 + I_2 - 2I_3,$$

say. Note that for all $s \in \{1, \dots, n\}$, $n \in \mathbb{N}$ and $\varepsilon \in \mathbb{R}_+$

$$\mu_{si}^2 \leqq E(X_{si}^2) \leqq \varepsilon^2 A_{ni}^2 + E\{X_{si}^2 1_{(|X_{si}|>\varepsilon A_{ni})}\} \leqq \varepsilon^2 A_{ni}^2 + \sum_{s=1}^{n} E\{X_{si}^2 1_{(|X_{si}|>\varepsilon A_{ni})}\}.$$

Hence using (2.7) we can find n_0 large enough such that for every $n \geqq n_0$

$$\mu_{si}^2 \leqq 4\varepsilon^2 A_{ni}, \qquad s \in \{1, \dots, n\}, \quad i \in \{1, \dots, k\}. \tag{2.10}$$

For $n \geqq n_0$, using the well-known Tchebyshev inequality we deduce

$$0 \leqq I_2 \leqq 4\varepsilon^2 \sum_{s=1}^{n} \frac{E\{(X_{si} - \mu_{si})^2\}}{x^2 A_{ni}^2} = \frac{4\varepsilon^2}{x^2},$$

and consequently

$$0 \leqq \limsup_{n\to\infty} I_2 \leqq \frac{4\varepsilon^2}{x^2}.$$

As to I_1, note that $|X_{si} - \mu_{si}| > A_{ni}x$ implies $|X_{si}| > A_{ni}x - |\mu_{si}|$. Using (2.10) it follows

$$|X_{si}| > A_{ni}(x - \varepsilon), \qquad s \in \{1, \dots, n\}, \quad i \in \{1, \dots, k\}. \tag{2.11}$$

Choosing ε small enough to ensure $x - \varepsilon > 0$, from (2.7) we obtain

$$0 \leqq \limsup_{n\to\infty} I_1 \leqq \lim_{n\to\infty} \frac{1}{A_{ni}^2} \sum_{s=1}^{n} E\{X_{si}^2 1_{(|X_{si}|>A_{ni}(x-\varepsilon))}\} = 0.$$

As to I_3, using (2.7), (2.10) and (2.11), as $n \to \infty$, we deduce

$$0 \leqq |I_3| \leqq \frac{2\varepsilon}{A_{ni}} \sum_{s=1}^{n} E\{|X_{si}| 1_{(|X_{si}|>A_{ni}(x-\varepsilon))}\} \leqq$$

$$\leqq \frac{2\varepsilon}{(x-\varepsilon)A_{ni}^2} \sum_{s=1}^{n} E\{X_{si}^2 1_{(|X_{si}|>A_{ni}(x-\varepsilon))}\} \to 0.$$

Combining the results for I_1, I_2 and I_3 and then taking $\varepsilon \searrow 0$ (2.5) follows. Using similar methods we verify (2.6). The proof ends by using Proposition 2. □

In our next result we obtain sufficient conditions for (2.7) and (2.8). In order to formulate the result, let

$\mathcal{G}=\{g : \mathbf{R}_+ \to \mathbf{R}_+ :$ (i) $\lim_{x\to\infty} g(x) = \infty$; (ii) $g(x)$ is nondecreasing; (iii) $x/g(x)$ is well defined and nondecreasing.$\}$

PROPOSITION 3. *Suppose there exist functions $g_i, h_i \in \mathcal{G}$, $i \in \{1, \dots, k\}$, such that for all $i, j \in \{1, \dots, k\}$, $i \neq j$,*

$$\lim_{n\to\infty} \frac{1}{A_{ni}^2 g_i(A_{ni})} \sum_{s=1}^{n} E\{X_{si}^2 g_i(|X_{si}|)\} = 0, \tag{2.12}$$

and

$$\lim_{n\to\infty} \frac{\sum_{s=1}^{n} E\{|X_{si}X_{sj}| h_i(|X_{si}|) h_j(|X_{sj}|)\}}{A_{ni} h_i(A_{ni}) A_{nj} h_j(A_{nj})} = 0. \tag{2.13}$$

Then (2.7) and (2.8) of Theorem 1 hold true.

PROOF. First consider (2.7). For every $x \in \mathbf{R}_+$, from the monotonicity of g_i, it follows

$$\frac{1}{A_{ni}^2} \sum_{s=1}^{n} E\{X_{si}^2 1_{(|X_{si}|>A_{ni}x)}\} \leqq$$

$$\leqq \frac{1}{g(xA_{ni})A_{ni}^2} \sum_{s=1}^{n} E\{X_{si}^2 g_i(|X_{si}|) 1_{(|X_{si}|>A_{ni}x)}\} \leqq$$

$$\leqq \frac{g(A_{ni})}{g(xA_{ni})} \frac{1}{g(A_{ni})A_{ni}^2} \sum_{s=1}^{n} E\{X_{si}^2 g_i(|X_{si}|)\}.$$

Since $g_i(x)$ and $x/g_i(x)$ is nondecreasing we have $g(A_{ni})/g(xA_{ni}) \leqq 1$ if $x > 1$, and $\leqq 1/x$ if $x \leqq 1$. Hence (2.7) follows from (2.12). Now consider (2.8). As

$n \to \infty$ we deduce

$$\left|\frac{1}{A_{ni}A_{nj}}\sum_{s=1}^{n} E\{|X_{si}X_{sj}|1_{(|X_{si}|>A_{ni}x;|X_{sj}|>A_{nj}y)}\}\right| \leq$$

$$\leq \frac{h_i(A_{ni})h_j(A_{nj})}{h_i(A_{ni}x)h_j(A_{nj}y)} \frac{\sum_{s=1}^{n} E\{|X_{si}X_{sj}|h_i(|X_{si}|)h_j(|X_{sj}|)\}}{A_{ni}h_i(A_{ni})A_{nj}h_j(A_{nj})} \to 0,$$

as desired. □

In the i.i.d. case we obtain the following

COROLLARY 1. *Let* $\{\mathbf{X}_n, n\in\mathbf{N}\}$ *be a sequence of i.i.d. random vectors in* $\mathbf{R}^k$ *with* $\mu_i = E\{X_{1i}\}\in\mathbf{R}$, $\sigma_i^2 = E\{X_{1i}^2\} - \mu_i^2 \in \mathbf{R}_+$, $c_{ij} = E\{(X_{1i}-\mu_i)(X_{1j}-\mu_j)\}$, $j, i\in\{1,\dots,k\}$, $j\neq i$. *Then* (2.1) *holds with* $A_{ni} = \sqrt{n}\sigma_i$ *and* $B_{ni} = n\mu_i$, $n\in\mathbf{N}$, $i\in\{1,\dots,k\}$. *The limiting covariance matrix* Γ *equals the correlation matrix of* $\mathbf{X}_1$. □

3. Limit distribution for one-dimensional renewal processes

Let $\{X_n, n\in\mathbf{N}\}$ be a sequence of independent $\mathbf{R}_+$-valued random variables. Denote by $\{N(t), t\in\mathbf{R}_+\}$ the associated renewal process defined by $N(t) = \sup\{n\in\mathbf{N} : X_1 + \ldots + X_n \leq t\}$, $t\in\mathbf{R}_+$. In the non-i.i.d. case for previous results concerning the asymptotic normality of $N(t)$ see Siegmund [11] or Niculescu [9]. Next we will assume there exist sequences $\{A_n, n\in\mathbf{N}\}\subset\mathbf{R}_+$ and $\{B_n, n\in\mathbf{N}\}\subset\mathbf{R}$ such that as $n\to\infty$

$$\frac{\sum_{s=1}^{n}(X_s - B_s)}{A_n} \Rightarrow V \tag{3.1}$$

where V is a standard normally distributed random variable. Note that (3.1) implies

$$\lim_{n\to\infty} A_n/A_{n+1} = 1 \quad \text{and} \quad \lim_{n\to\infty} B_n/A_n = 0. \tag{3.2}$$

For every $x\in\mathbf{R}$ denote by $[x]$ the greatest integer contained in x.

THEOREM 2. *Assume* (3.1) *holds true and there exist functions* $U, f : \mathbf{R}\to\to\mathbf{R}_+$ *such that for all* $y\in\mathbf{R}_+$

(i) $$\lim_{x\to\infty} x + yf(x) = \infty,$$

(ii) $$\lim_{x\to\infty} f(x+yf(x))/f(x) = 1,$$

(iii) $$\lim_{x\to\infty} \frac{A_{[x+yf(x)]}}{A_{[x]}} = a(y) \in \mathbb{R}_+,$$

(iv) $$\lim_{x\to\infty} \frac{\sum_{s=1}^{[x+yf(x)]} B_s - U(x)}{A_{[x]}} = h(y) \in \mathbb{R}.$$

Then for every $z \in \mathbb{R}$

$$\lim_{x\to\infty} P\{N(U(x)) > zf(x) + x\} = \Phi(-h(z)/a(z)),$$

where $\Phi(\cdot)$ *is the standard normal distribution function.*

Proof. Observe that for x large enough

$$P\{N(U(x)) > zf(x)+x\} = P\left\{\frac{\sum_{s=1}^{[zf(x)+x]+1}(X_s - B_s)}{A_{[zf(x)+x]+1}} \leqq -\frac{\sum_{s=1}^{[zf(x)+x]+1} B_s - U(x)}{A_{[zf(x)+x]+1}}\right\}.$$

Using (i)–(iv) and (3.2) we deduce

$$\frac{\sum_{s=1}^{[zf(x)+x]+1} B_s - U(x)}{A_{[zf(x)+x]+1}} =$$

$$= \frac{\sum_{s=1}^{[zf(x)+x]} B_s - U(x)}{A_{[x]}} \frac{A_{[x]}}{A_{[zf(x)+x]}} \frac{A_{[zf(x)+x]}}{A_{[zf(x)+x]+1}} + \frac{B_{[zf(x)+x]+1}}{A_{[zf(x)+x]+1}} \to \frac{h(z)}{a(z)}$$

as $x \to \infty$. Now the desired result is a simple consequence of (3.1) and the continuity of Φ.

Since in most cases we will assume $A_n^2 = \sum_{s=1}^{n} E\{X_s^2\}$, we restrict attention to nondecreasing sequences $\{A_n, n \in \mathbb{N}\}$. Now we discuss the limit functions appearing in (iii) and (iv) of Theorem 2.

Proposition 4. *Suppose the assumptions in Theorem 2 are satisfied. Then*

(a) $a(x) = e^{Cx}$ *for some* $C \geqq 0$; *and*

(b) $$h(x) = \begin{cases} A + Bx & \text{if } C = 0, \\ A + B(1 - e^{Cx}) & \text{if } C > 0, \end{cases}$$

for some constants $A, B \in \mathbf{R}$.

PROOF. (a) From (i)–(iii) it follows that $a(\cdot)$ is nondecreasing and continuous a.e. For almost all $z, y \in \mathbf{R}$ we deduce

$$a(y) = \lim_{x\to\infty} \frac{A_{[x+zf(x)+yf(x+zf(x))]}}{A_{[x+zf(x)]}} =$$

$$= \lim_{x\to\infty} \frac{A_{[x+(z+yf(x+zf(x))/f(x))f(x)]}}{A_{[x]}} \frac{A_{[x]}}{A_{[x+zf(x)]}} =$$

$$= a(z+y)/a(z).$$

Hence $a(z+y) = a(z)a(y)$ a.e. This in fact shows that $a(\cdot)$ is continuous and that there exist a constant $C \geqq 0$, such that $a(x) = e^{Cx}$, $x \in \mathbf{R}$.

(b) Let $A = h(0)$ and $H(x) = h(x) - A$, $x \in \mathbf{R}$. For all $z, y \in \mathbf{R}$, as in part (a) we obtain $H(z+y) - H(z) = a(z)H(y)$.

If $C = 0$ then $H(z+y) = H(z) + H(y)$ for all $z, y \in \mathbf{R}$, which implies the existence of a constant $B \in \mathbf{R}$ such that $H(x) = Bx$, $x \in \mathbf{R}$.

If $C > 0$ then, for every $z, y \in \mathbf{R}$, $H(z)(1 - e^{Cz}) = H(y)(1 - e^{Cy})$, which is equivalent to the existence of a constant $B \in \mathbf{R}$ such that $H(x) = B(1 - e^{Cx})$, $x \in \mathbf{R}$. □

REMARK 2. Condition (ii) on f can be motivated as follows. If $\{A_n, n \in \mathbf{N}\}$ is such that

$$\lim_{x\to\infty} \frac{A_{[x+yf(x)]}}{A_{[x]}} = e^{Cy}, \quad y \in \mathbf{R},$$

for some $C \in \mathbf{R}_+$, and $A : \mathbf{R} \to \mathbf{R}_+$ is defined by $A(x) = A_{[x]} 1_{[1,\infty)}(x)$, $x \in \mathbf{R}$, then $A(\cdot)$ belongs to the class of functions $\Gamma(f(x)/C)$ introduced by de Haan [4], and consequently $A(\cdot)$ automatically satisfies (ii). □

4. Limit distribution for multi-dimensional renewal processes

Let $\alpha \in \mathbf{R}$. A sequence $\{s_n, n \in \mathbf{N}\} \subset \mathbf{R}_+$ is said to belong to the class $\mathcal{RV}_\alpha$ if and only if, as $n \to \infty$, $s_n \sim n^\alpha L(n)$, where L is a slowly varying function, i.e. L is a positive and measurable function such that for all $t > 0$, $L(tx) \sim L(x)$ as $x \to \infty$.

PROPOSITION 5. *Let* $\mathbf{Z}_n = (Z_{n1}, \dots, Z_{nk})$, $n \in \mathbf{N}$, *be a sequence of independent random vectors such that* $E\{Z_{ni}\} = 0$ *and* $A_{ni}^2 = \sum_{s=1}^{n} E\{Z_{si}^2\} \in \mathbf{R}_+$ *for all* $i \in \{1, \dots, k\}$ *and* $n \in \mathbf{N}$. *Suppose* (2.5) *and* (2.6) *of Proposition 2 are satisfied for some unit-diagonal covariance matrix* $\Gamma = (\Gamma_{ij})_{i,j=1}^{k}$ *and* $\{A_{ni}, n \in \mathbf{N}\} \in \mathcal{RV}_{1/2}$, $i \in \{1, \dots, k\}$. *Let* $\mathrm{K} = (K_1, \dots, K_k) \in \mathbf{R}_+^{k-1} \times \{1\}$. *Then, as*

$n \to \infty$,

$$\left(\frac{1}{A_{n_1(n),1}}\sum_{s=1}^{n_1(n)} Z_{s1}, \dots, \frac{1}{A_{n_{k-1}(n),k-1}}\sum_{s=1}^{n_{k-1}(n)} Z_{s,k-1}, \frac{1}{A_{nk}}\sum_{s=1}^{n} Z_{sk}\right) \Rightarrow \mathbf{V}$$

for all sequences $\{n_i(n), n \in \mathbb{N}\} \subset \mathbb{N}$ *such that* $n_i(n) \sim K_i n$ *as* $n \to \infty$, $i \in \in \{1, \dots, k-1\}$, *where the random vector* $\mathbf{V}$ *has a multivariate normal distribution with zero mean vector and the covariance matrix* $\Gamma(\mathbf{K}) = (\gamma_{ij})_{i,j=1}^{k}$ *defined by* $\gamma_{ij} = \Gamma_{ij} \min(\sqrt{K_i/K_j}, \sqrt{K_j/K_i})$, $i, j \in \{1, \dots, k\}$.

PROOF. With the observation that the proof for dimension $k > 2$ only reproduces the similar one for the bivariate case, we restrict ourselves to the case $k = 2$. First consider $K_1 > 1$. Then $n_1(n) > n$ for n large enough. We have

$$\left(\frac{1}{A_{n_1(n),1}}\sum_{s=1}^{n_1(n)} Z_{s1}, \frac{1}{A_{n2}}\sum_{s=1}^{n} Z_{s2}\right) = \left(\left(\frac{A_{n1}}{A_{n_1(n),1}} - \frac{1}{\sqrt{K_1}}\right)\frac{1}{A_{n1}}\sum_{s=1}^{n} Z_{s1}, 0\right) +$$

$$+\left(\frac{1}{A_{n1}\sqrt{K_1}}\sum_{s=1}^{n} Z_{s1}, \frac{1}{A_{n2}}\sum_{s=1}^{n} Z_{s2}\right) + \left(\frac{1}{A_{n_1(n),1}}\sum_{s=n+1}^{n_1(n)} Z_{s1}, 0\right) =$$

$$= I_4 + I_5 + I_6, \quad \text{say.}$$

From Proposition 2 it follows

$$\left(\frac{1}{A_{n1}}\sum_{s=1}^{n} Z_{s1}, \frac{1}{A_{n2}}\sum_{s=1}^{n} Z_{s2}\right) \Rightarrow \mathbf{V} = (V_1, V_2)$$

as $n \to \infty$, where $\mathbf{V}$ is normally distributed with zero mean vector and covariance matrix $\Gamma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, $\rho = \Gamma_{12}$. Then, as $n \to \infty$,

$$I_5 \Rightarrow \left(\frac{1}{\sqrt{K_1}} V_1^{\cdot}, V_2^{\cdot}\right), \quad \text{where} \quad (V_1^{\cdot}, V_2^{\cdot}) \stackrel{\mathcal{D}}{=} (V_1, V_2).$$

Secondly, since

$$\frac{1}{A_{n1}}\sum_{s=1}^{n} Z_{s1} \Rightarrow V_1 \quad \text{and} \quad \frac{A_{n1}}{A_{n_1(n),1}} \to \frac{1}{\sqrt{K_1}},$$

it follows $I_4 \xrightarrow{P} (0,0)$ as $n \to \infty$. Finally using (2.5) and $\{A_{n1}, n \in \mathbb{N}\} \in \mathcal{RV}_{1/2}$, it follows as $n \to \infty$

$$I_6 \Rightarrow (V_1^{\cdot\cdot} \sqrt{(K_1 - 1)/K_1}, 0)$$

where $V_1^{\cdot\cdot}$ is standard normally distributed and it is independent of $(V_1^{\cdot}, V_2^{\cdot})$. Combining the results on I_4, I_5 and I_6 the proof is concluded for the case $K_1 > 1$.

Next suppose $K_1 = 1$. A similar decomposition holds with I_6 replaced by

$$I_6^{\cdot} = \frac{(-1)^{\delta}}{A_{n_1(n),1}}\Big(\sum_{s=1+\min(n,n_1(n))}^{\max(n,n_1(n))} X_{s1}, 0\Big),$$

where $\delta = 0$ or 1 according to $n_1(n) > n$ or $n_1(n) \leqq n$. Since for every $\varepsilon \in \mathbb{R}_+$, $u, v \in \mathbb{N}$, $u < v$,

$$P\Big\{\Big|\sum_{s=u+1}^{v} X_{s1}\Big| > \varepsilon\Big\} \leqq \frac{A_{v1}^2 - A_{u1}^2}{\varepsilon^2}$$

then using $\min\{n, n_1(n)\} \sim \max\{n, n_1(n)\}$, as $n \to \infty$, and $\{A_{n1}, n \in \mathbb{N}\} \in \mathcal{RV}_{1/2}$, we obtain $I_6^{\cdot} \xrightarrow{P} (0,0)$, which ends the proof. □

THEOREM 3. *Let $\{\mathbf{X}_n, n \in \mathbb{N}\}$ be a sequence of independent random vectors with values in $\mathbb{R}_+^k$ such that $\mu_{ni} = E\{X_{ni}\} \in \mathbb{R}_+$ and $A_{ni}^2 = \sum_{s=1}^{n} E\{(X_{si} - \mu_{si})^2\} \in \mathbb{R}_+$, $i \in \{1, \ldots, k\}$, $n \in \mathbb{N}$. Suppose (2.7), (2.8) and (2.9) hold true and $\{A_{ni}, n \in \mathbb{N}\} \in \mathcal{RV}_{1/2}$, $i \in \{1, \ldots, k\}$. Assume there exists functions $U_i : \mathbb{R} \to \mathbb{R}_+$ such that for all $y \in \mathbb{R}$*

$$\lim_{x\to\infty} \frac{\sum_{s=1}^{[x+yA_{[x],i}]} \mu_{si} - U_i(x)}{A_{[x],i}} = h_i(y), \tag{4.1}$$

$i \in \{1, \ldots, k\}$. Then for all $\mathbf{z} = (z_1, \ldots, z_k) \in \mathbb{R}^k$

$$\lim_{x\to\infty} P\{\mathbf{N}(\mathbf{U}(x)) > \mathbf{z} \,\square\, \mathbf{A}_{[x]} + x\mathbf{1}\} = \Phi\Big\{\mathop{\times}_{i=1}^{k} (-\infty, -h_i(z_i)]\Big\},$$

where Φ is the same multivariate normal distribution function as in Proposition 1, $\mathbf{U}(x) = (U_1(x), \ldots, U_k(x))$, $\mathbf{A}_n = (A_{n1}, \ldots, A_{nk})$, $n \in \mathbb{N}$, and $\mathbf{z} \,\square\, \mathbf{A}_{[x]} + x\mathbf{1} = (z_1 A_{[x],1} + x, \ldots, z_k A_{[x],k} + x)$ whatever $x \in [1, \infty)$.

PROOF. The proof is similar to that of Theorem 2. Indeed we can take $f_i(x) \equiv A_{[x],i}$; then (i), (ii) and (iii) in Theorem 2 are satisfied since $\{A_{ni}, n \in \mathbb{N}\} \in \mathcal{RV}_{1/2}$. Also, Proposition 4 shows that $h_i(x) = C_i + xD_i$ for some constants $D_i \in \mathbb{R}_+$ and $C_i \in \mathbb{R}$, $i \in \{1, \ldots, k\}$. Since for every $i, j \in \{1, \ldots, k\}$, $i \neq j$,

$$\lim_{x\to\infty} \frac{x + z_i A_{[x],i}}{x + z_j A_{[x],j}} = 1$$

the desired result follows from Proposition 5. □

COROLLARY 2. *If in Theorem 3 we assume there exist constants* $a_i \in \in \mathbf{R}_+$, $i \in \{1, \dots, k\}$, *and a function* $B : \mathbf{R}_+ \to \mathbf{R}_+$ *such that, as* $n \to \infty$, $A_{ni} \sim a_i B(n)$ *for each* $i \in \{1, \dots, k\}$, *then for all* $\mathbf{z} = (z_1, \dots, z_k) \in \mathbf{R}^k$ *and* $z \in \mathbf{R}$

$$\text{(a)}\quad \lim_{x\to\infty} P\{\mathbf{N}(\mathbf{U}(x)) > B(x)\mathbf{z} + x\mathbf{1}\} = \Phi\left\{\mathop{\times}_{i=1}^{k}\left(-\infty, -C_i - \frac{D_i}{a_i} z_i\right]\right\};$$

$$\text{(b)}\quad \lim_{x\to\infty} P\{\inf_{1\leq i\leq k}(N_i(U_i(x))) > zB(x) + x\} = \Phi\left\{\mathop{\times}_{i=1}^{k}\left(-\infty, -C_i - \frac{D_i}{a_i} z\right]\right\};$$

$$\text{(c)}\quad \lim_{x\to\infty} P\{\sup_{1\leq i\leq k}(N_i(U_i(x))) < zB(x) + x\} = \Phi\left\{\mathop{\times}_{i=1}^{k}\left(-C_i - \frac{D_i}{a_i} z, \infty\right)\right\};$$

where the constants C_i *and* D_i *are those from the proof of Theorem 3,* $i \in \{1, \dots, k\}$.

COROLLARY 3. *Assume in Theorem 3 that for every* $i \in \{1, \dots, k\}$ *there exist* $\sigma_i, \mu_i \in \mathbf{R}_+$ *such that*

$$A_{ni}^2 \sim n\sigma_i^2, \quad \text{as } n \to \infty; \tag{4.2}$$

and

$$\sqrt{n}\Big(\frac{1}{n}\sum_{s=1}^{n}\mu_{si} - \mu_i\Big) \to 0 \quad \text{as } n \to \infty. \tag{4.3}$$

Then (4.1) *holds with* $h_i(y) = y\mu_i$, $i \in \{1, \dots, k\}$, *and then for every* $\mathbf{z} = (z_1, \dots, z_k) \in \mathbf{R}^k$

$$\lim_{x\to\infty} P\{\mathbf{N}(x\mu) > \mathbf{z}\sqrt{x} + x\mathbf{1}\} = \Phi\left\{\mathop{\times}_{i=1}^{k}\left(-\infty_i, -\frac{\mu_i}{\alpha_i} z_i\right]\right\},$$

where $\mu = (\mu_1, \dots, \mu_k)$.

PROOF. We only have to verify that (4.1) holds true for $U_i(x) = \mu_i x$ and $h_i(y) = \mu_i y$. Note that

$$\frac{\sum\limits_{s=1}^{[x+yA_{[x],i}]} \mu_{si} - \mu_i x}{A_{[x],i}} =$$

$$= \frac{[x + yA_{[x],i}]}{A_{[x],i}}\left\{\frac{1}{[x + yA_{[x],i}]}\sum_{s=1}^{[x+yA_{[x],i}]} \mu_{si} - \mu_i\right\} + \frac{[x + yA_{[x],i}] - x}{A_{[x],i}}\mu_i.$$

Using (4.2) and (4.3) it follows that this converges to $y\mu_i$ as $x \to \infty$. □

COROLLARY 4. *In the i.i.d. case, if assumptions in Theorem 3 are satisfied then for every* $\mathbf{z} = (z_1, \dots, z_k) \in \mathbb{R}^k$

$$\lim_{x\to\infty} P\{\mathrm{N}(x\mu) > \mathbf{z}\sqrt{x} + x\mathbf{1}\} = \Phi\left\{\mathop{\times}_{i=1}^{k}\left(-\infty_i, -\frac{\mu_{i1}}{\sigma_{i1}} z_i\right]\right\}. \qquad \square$$

REFERENCES

[1] AHMAD, I. A., The exact order of normal approximation in bivariate renewal theory, *Adv. in Appl. Probab.* **13** (1981), 113–128. *MR* **82b**:60111

[2] CHOW, Y. S. and TEICHER, H., *Probability Theory. Independence, interchangeability, martingales,* Springer-Verlag, New York–Heidelberg, 1978. *MR* **80a**:60004

[3] CSENKI, A., An invariance principle in k-dimensional extended renewal theory, *J. Appl. Probab.* **16** (1979), 567–574. *MR* **80i**:60123

[4] HAAN, L. DE, *On regular variation and its applications to the weak convergence of sample extremes,* Mathematical Centre Tracts **32**, Mathematisch Centrum, Amsterdam, 1970. *MR* **44** #3370

[5] HAAN, L. DE, OMEY, E. and RESNICK, S., Domains of attraction and regular variation in IR^d, *J. Multivariate Anal.* **14** (1984), 17–33. *MR* **85e**:60025

[6] HUNTER, J. J., Renewal theory in two dimensions: asymptotic results, *Advances in Appl. Probability* **6** (1974), 546–562. *MR* **49** #11649b

[7] MIHOC, G. and NICULESCU, S. P., *Procese stohastice de reînnoire* (Renewal processes), Editura Academiei Republicii Socialiste România, Bucharest, 1982. *MR* **84k**:60110

[8] NICULESCU, S. P., On the asymptotic distribution of multivariate renewal processes, *J. Appl. Probab.* **21** (1984), 639–645. *MR* **85i**:60076

[9] NICULESCU, S. P., On general renewal processes, *Stud. Cerc. Mat.* **40** (1988), 507–514. *MR* **90f**:60065

[10] NICULESCU, S. P. and OMEY, E., On the exact order of normal approximation in multivariate renewal theory, *J. Appl. Probab.* **22** (1985), 280–287. *MR* **86i**:60073

[11] SIEGMUND, D. O., On the asymptotic normality of one-sided stopping rules, *Ann. Math. Statist.* **39** (1968), 1493–1497. *MR* **37** #7043

(Received September 16, 1990)

CANADA CENTRE FOR INLAND WATERS
NATIONAL WATER RESEARCH INSTITUTE
ROOM L464
867 LAKESHORE ROAD
P.O. BOX 5050
BURLINGTON, ONTARIO
L7R 4A6
CANADA

ECONOMISCHE HOGESCHOOL SINT-ALOYSIUS
STORMSTRAAT 2
B–1000 BRUSSEL
BELGIUM

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 189–207

RESULTS AND PROBLEMS ON STRONGLY CONNECTED MOORE AUTOMATA

A. ÁDÁM and I. BABCSÁNYI

To the memory of Professor Tibor Gallai

Introduction

In the present article the study of the questions left open in the previous paper [4] is continued in another direction than it was done in [5].

After considering the congruences of an automaton with *only one* input symbol in [3], the investigation (by algebraic methods) of the question of simplicity (i.e. the lack of nontrivial congruences) of automata having *an arbitrary number* of input signs was initiated in [4]; the papers [4]–[5] deal with the reduction of this problem to the particular class consisting of *strongly connected* automata.

Three special problems were raised in [4, p.171] after the main result. The second and third of these were studied in Section 6 of [4] and in [5]. The first problem — to find a criterion of simplicity within the class of strongly connected automata — seems to be the most serious one. We make in the present work the first move toward the analysis of it (following the methods used in [4]–[5]). The results stated in the sequel are still far from a complete solution of the question. Among them, we assert some (easily provable) sufficient conditions for a strongly connected automaton in order to be non-simple. We show by an example that the system of these conditions is not necessary; this fact leads us to open problems which seem to be difficult. The third section of the paper explains also interesting examples, showing that various imaginable causes, implying non-simplicity, may really occur.

We consider always automata in the narrower sense that the input stimuli form a *free* semigroup (not an arbitrary one). As we regard the notion of simplicity, the output function plays an essential role.[1]

1991 *Mathematics Subject Classification.* Primary 68Q70.
Key words and phrases. Simplicity of automata, distinguishable and indistinguishable state pairs, automorphism.

The research of the authors has been partially supported by the Hungarian National Foundation for Scientific Research Grant (OTKA) No. 1909 and No. 1965, respectively.

[1]Each automaton having at least two states and only one output symbol is obviously non-simple. Each automaton with a bijective output function is simple.

More or less independently of [3]–[5], the authors have also written some other papers on several sides of the topics of semigroup-theoretical aspects of automata and the questions of their behaviour. Let us now mention the articles [1] and [6] where automata have been presented in terms of certain partitions of free semigroups; furthermore the work [2] in which automata have been constructed by a combinatorial procedure and the notion of distinguishability of states was introduced.

In Sections 4–5 of this paper the automorphisms of automata are touched. Our approach differs from that of Karpov [9] who gave a construction for determining the automorphism group of an automaton provided that its congruences are known. We think that it is not easier to get an overview of the congruences (in case of the majority of non-simple automata) than to establish the automorphism group.

1. Preliminaries

In the present paper finite Moore automata $\mathbf{A} = (A, X, Y, \delta, \lambda)$ are considered. We denote by $F(X)$ the set of all input words (including the empty word e). The notions of connectedness, congruence, simplicity, (in)distinguishable state pair are used in the same sense as in the previous article [4]. We recall that $\mathbf{A}$ is called *strongly connected* if for each state pair a, b there is an input word p such that $\delta(a, p) = b$ (or, equivalently, if $\mathbf{A}$ has no proper subautomaton; we understand by subautomaton the A-subautomaton in sense of § 1.3 of [8]).

Since the congruences and simplicity of automata satisfying $|X| = 1$ have been treated already in [3], we regard in what follows automata fulfilling $|X| \geqq 2$. Also $|Y| \geqq 2$ is supposed (cf. Footnote 1).

In the sequel state pairs a, b will be often referred to. Such a pair is considered as an ordered one,[2] and we speak of a *proper* pair if we want to emphasize that $a \neq b$ is required.

It is known that the indistinguishability of states[3] of $\mathbf{A}$ is an equivalence relation and the partition $\pi_{\max}$, corresponding to it, is the largest congruence of $\mathbf{A}$. The smallest congruence of $\mathbf{A}$ equals, of course, the smallest partition o of the set of states. $\pi_{\max}$ and o are also called the *maximal congruence* and the *trivial congruence* of $\mathbf{A}$, respectively. $\mathbf{A}$ is simple precisely when $\pi_{\max} = o$ (or, equivalently, if each proper state pair is distinguishable).

A subset J of A is called *strongly connected* if for each pair $a(\in J)$, $b(\in J)$ there is a sequence $a(= a_0), x_1, a_1, x_2, \dots, x_k, b(= a_k)$ such that $a_1, a_2, \dots, a_{k-1}$ belong to J and $\delta(a_{i-1}, x_i) = a_i$ whenever $1 \leqq i \leqq k$ $(x_i \in X)$.

---

[2] The ordering is frequently indifferent since the distinguishability of two states and the set $H_{a,b}$ — to be introduced later — are defined symmetrically.

[3] This is defined for a state pair a, b by the fulfilment of $\lambda(\delta(a, p)) = \lambda(\delta(b, p))$ for each input word p.

A mapping α of the state set A (of **A**) into itself is called an *endomorphism* if

$$\alpha(\delta(a,x)) = \delta(\alpha(a),x) \quad \text{and} \quad \lambda(a) = \lambda(\alpha(a))$$

are satisfied for every choice of $a(\in A)$ and $x(\in X)$. An endomorphism is said to be an *automorphism* if it is a bijective mapping (of A onto itself). When we speak of a nontrivial endomorphism (or automorphism), we mean that it differs from the identical mapping of A.

A mapping γ is called a *partial quasi-isomorphism* (of an automaton A) if the following five conditions are fulfilled:

(1) The definition domain J and the (precise) range K of γ satisfy the formulae

$$J \cup K \subseteqq A, \quad J \cap K = \emptyset, \quad |J| \geqq 2.$$

(2) γ is a bijective mapping (between J and K).[4]
(3) J is a strongly connected subset of A.
(4) The inequality

$$\delta(a,x) \neq \delta(\gamma(a),x)$$

implies

$$\delta(a,x) \in J \& \delta(\gamma(a),x) \in K \& \gamma(\delta(a,x)) = \delta(\gamma(a),x)$$

for each $a(\in J)$ and $x(\in X)$.

(5) $\lambda(a) = \lambda(\gamma(a))$ for each $a(\in J)$.

Let a state pair a,b be considered. Denote by $H_{a,b}$ the set of all input words p such that $\delta(a,p) \neq \delta(b,p)$. Evident consequences of the definition of $H_{a,b}$ are:

$px \in H_{a,b}$ implies $p \in H_{a,b}$ (where $p \in F(X)$, $x \in X$),

$H_{a,b} = H_{b,a}$,

$H_{a,b} = \emptyset$ whenever $a = b$.

Let all the state pairs (of A) be partitioned into three types in sense of the following rules:

the pair a,b is of type (I) if $H_{a,b}$ is finite,

the pair a,b is of type (II) if $H_{a,b} = F(X)$,

the pair a,b is of type (III) if $H_{a,b}$ is infinite and properly included in $F(X)$.

The types (II), (III) contain proper state pairs only. We shall see examples in § 3 showing that any of the three types (and many of the imaginable combinations of them) is existing.

---

[4] Hence $|J| = |K|$.

2. The Main Question

The question of characterizing the simple automata among the strongly connected Moore ones was explicitly raised in [4] (p. 171). A complete solution of it would be a generalization of the last sentence of Corollary 2 in [3] (p. 262), dealing with strongly connected autonomous automata.[5] Now we express our basic question in the form of a compound statement.

MAIN QUESTION. Determine the logical connection of the following three conditions (A), (B), (C) for strongly connected Moore automata **A**:

(A) **A** is not simple.

(B) At least one of the subsequent three assertions is true in **A**:

- (b-1) There is an indistinguishable proper state pair of type (I).
- (b-2) There is an indistinguishable proper state pair of type (II).
- (b-3) There is an indistinguishable proper state pair of type (III).

(C) At least one of the subsequent three assertions is true in **A**:

- (c-1) There is a proper pair $a(\in A)$, $b(\in A)$ such that $\lambda(a)=\lambda(b)$ and $\delta(a,x)=\delta(b,x)$ for each $x(\in X)$.
- (c-2) **A** has a nontrivial automorphism.
- (c-3) There is a partial quasi-isomorphism between two appropriate subsets of A.

In what follows, we shall see that (A), (B) are equivalent, and we shall verify that (B) follows from (C) (see Propositions 1, 2, 5, 12 and Corollary 2). The implication (B) $\Rightarrow$ (C), however, fails to be true (in general), thus some important questions are left open (cf. Example 7 and Problems 1, 2, 4, 5).

Two results concerning the Main Question may be stated at once. (Observe that they are valid without assuming the strong connectedness.) The known facts about distinguishability, largest congruence and simplicity[6] imply immediately

PROPOSITION 1. *Conditions* (A) *and* (B) *in the Main Question are equivalent.* □

PROPOSITION 2. *Assertions* (b-1) *and* (c-1) *are equivalent.*

PROOF. Implication (c-1) $\Rightarrow$ (b-1) is obvious. Conversely, consider indistinguishable states c, d such that $c\neq d$ and $H_{c,d}$ is finite. Let p be an input word whose length is maximal (in $H_{c,d}$). It is easy to see that $a=\delta(c,p)$ and $b=\delta(d,p)$ fulfil (c-1). □

---

[5] It follows from this result that each indistinguishable proper state pair belongs to type (II) if $|X|=1$.

[6] These were referred to already in Section 1. For details see [2] (especially Section 5) and [4] (Proposition 5).

3. Some examples

EXAMPLE 1. Now we shall see a sequence of automata, depending on a parameter $w(\geqq 1)$. For any particular choice of w, put

$$A = \{a, b_1, b_2, \dots, b_w, c_1, c_2, \dots, c_w\}$$

(thus $|A| = 2w + 1$), $X = \{x_1, x_2\}$ and $Y = \{y_1, y_2\}$. The transition function is defined by

$$\delta(a, x_1) = b_1, \qquad \delta(a, x_2) = c_1,$$

$$\left.\begin{array}{l}\delta(b_i, x_1) = \delta(b_i, x_2) = b_{i+1} \\ \delta(c_i, x_1) = \delta(c_i, x_2) = c_{i+1}\end{array}\right\} \quad \text{if } 1 \leqq i \leqq w-1,$$

$$\delta(b_w, x_1) = \delta(b_w, x_2) = \delta(c_w, x_1) = \delta(c_w, x_2) = a,$$

and the output function is defined by

$$\lambda(a) = y_1, \qquad \lambda(b_i) = \lambda(c_i) = y_2 \qquad \text{if } 1 \leqq i \leqq w.$$

It can easily be seen that exactly the state pairs of form b_i, c_i are indistinguishable (where $1 \leqq i \leqq w$), every indistinguishable pair is of type (I) and H_{b_i, c_i} consists of all input words whose length does not exceed $w - i$. The assertion (c-1) holds for b_w, c_w.

One of these examples (for $w = 3$) can be seen in Fig. 1.

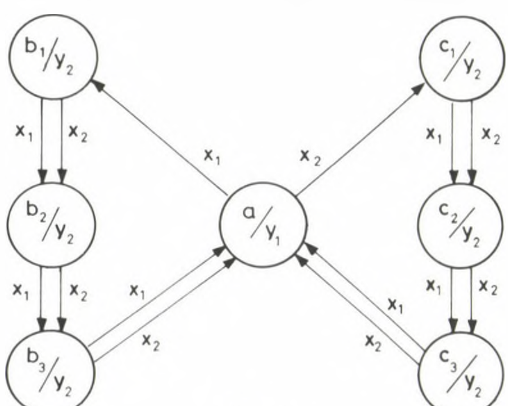


Fig. 1

In Examples 2–7 we shall write simply i instead of a_i.

EXAMPLE 2. Put $A = \{1, 2, \dots, 6\}$, $X = \{x_1, x_2\}$, $Y = \{y_1, y_2\}$; let δ, λ be defined by Table 1 (see Fig. 2).

Table 1

i	$\delta(i,x_1)$	$\delta(i,x_2)$	$\lambda(i)$
1	2	6	y_1
2	3	1	y_2
3	4	2	y_1
4	5	3	y_1
5	6	4	y_2
6	1	5	y_1

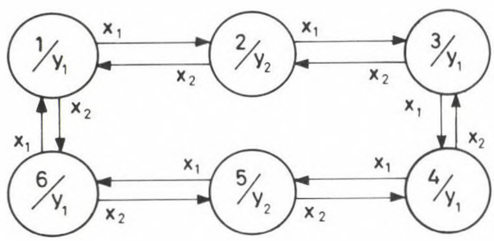


Fig. 2

There are three indistinguishable pairs: 1, 4; 2, 5; 3, 6; they are of type (II). The automaton has a nontrivial automorphism:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 1 & 2 & 3 \end{pmatrix}.$$

EXAMPLE 3. Put $A = \{1, 2, \ldots, 6\}$, $X = \{x_1, x_2\}$, $Y = \{y_1, y_2, y_3\}$; let δ, λ be defined by Table 2 (see Fig 3).

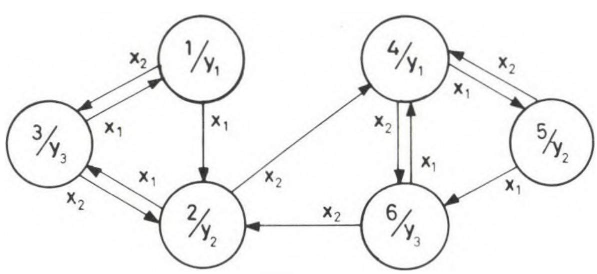


Fig. 3

Table 2

i	$\delta(i, x_1)$	$\delta(i, x_2)$	$\lambda(i)$
1	2	3	y_1
2	3	4	y_2
3	1	2	y_3
4	5	6	y_1
5	6	4	y_2
6	4	2	y_3

There are three indistinguishable pairs: 1, 4; 2, 5; 3, 6; they are of type (III). The mapping

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

is a partial quasi-isomorphism with $J = \{1, 2, 3\}$, $K = \{4, 5, 6\}$. Every state belongs to $J \cup K$.

EXAMPLE 4. We recall in Table 3 Example 3 of the paper [3] (pp. 278–279). There are three indistinguishable pairs: 2, 3; 4, 5; 6, 7; they are of type (III). The mapping

$$\begin{pmatrix} 2 & 4 & 6 \\ 3 & 5 & 7 \end{pmatrix}$$

is a partial quasi-isomorphism. The state 1 does not belong to $J \cup K$.

Table 3

i	$\delta(i, x_1)$	$\delta(i, x_2)$	$\lambda(i)$
1	2	3	y_1
2	4	4	y_2
3	5	5	y_2
4	6	6	y_3
5	7	7	y_3
6	2	1	y_4
7	3	1	y_4

EXAMPLE 5. Put $A = \{1, 2, \ldots, 5\}$, $X = \{x_1, x_2, x_3\}$, $Y = \{y_1, y_2\}$; let δ, λ be defined by Table 4 (see Fig. 4). It is easy to see that this automaton **A** has no nontrivial automorphism.[7] Introduce the following notation for subsets of A:

$$B = \{1,4\}, \quad C = \{2,3\}, \quad D = \{3,5\}, \quad E = \{1,2,4\}, \quad F = \{1,4,5\}.$$

A has two nontrivial congruences:

$$\pi_1 = \langle B, \{2\}, \{3\}, \{5\} \rangle ,$$
$$\pi_{\max} = \langle D, E \rangle .$$

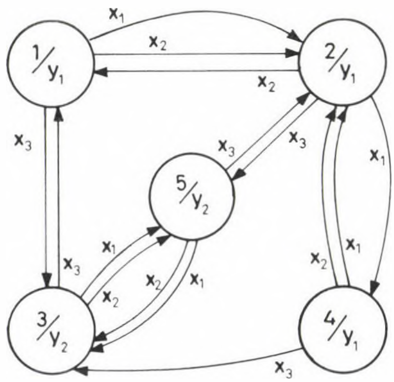


Table 4

i	$\delta(i, x_1)$	$\delta(i, x_2)$	$\delta(i, x_2)$	$\lambda(i)$
1	2	2	3	y_1
2	4	1	5	y_1
3	5	5	1	y_2
4	2	2	3	y_1
5	3	3	2	y_2

By regarding $\pi_{\max}$ it is clear that there are four indistinguishable state pairs. Among these, the pair 1, 4 belongs to type (I).

We are going to show that the indistinguishable pairs 1, 2; 2, 4; 3, 5 are of type (II). Indeed, consider the partition $\pi_2 = \langle C, F \rangle$ of the state set.

[7] This can be checked immediately, or follows from the fact that $|A|$ is prime by use of the assertions stated in the next section.

π_2 is not a congruence since it is not compatible with λ. However, π_2 is compatible with δ, therefore we can form the factor semiautomaton (i.e., automaton without output function) A/π_2. This is a semiautomaton having two states and satisfying $\delta'(d, x) \neq d$ for every choice of a state d and an input symbol $x(\in X)$ (where δ' is the transition function of A/π_2).

Let now b, c be selected arbitrarily out of the pairs 1, 2; 2, 4; 3, 5. One of b, c belongs to C, the other of them belongs to F. For any choice of $p(\in F(X))$, $\delta(b, p)$ and $\delta(c, p)$ are not equal because they lie in different classes of π_2.

EXAMPLE 6. Put $A = \{1, 2, \ldots, 8\}$, $X = \{x_1, x_2\}$, $Y = \{y_1, y_2\}$; let δ, λ be defined by Table 5 (see Fig. 5).

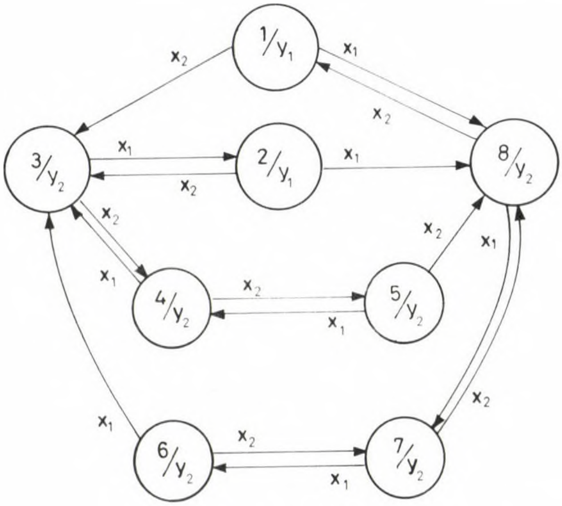


Fig. 5

Table 5

i	$\delta(i, x_1)$	$\delta(i, x_2)$	$\lambda(i)$
1	8	3	y_1
2	8	3	y_1
3	2	4	y_2
4	3	5	y_2
5	4	8	y_2
6	3	7	y_2
7	6	8	y_2
8	7	1	y_2

There are three indistinguishable state pairs: 1, 2; 4, 6; 5, 7. The pair

1, 2 is of type (I). The pairs 4, 6 and 5, 7 belong to type (III). The mapping

$$\begin{pmatrix} 4 & 5 \\ 6 & 7 \end{pmatrix}$$

is a partial quasi-isomorphism (with $J = \{4,5\}$, $K = \{6,7\}$). The states 1, 2, 3, 8 do not belong to $J \cup K$.

EXAMPLE 7. Put $A = \{1,2,3,4,5\}$, $X = \{x_1, x_2\}$, $Y = \{y_1, y_2\}$; let δ, λ be defined by Table 6 (see Fig. 6).

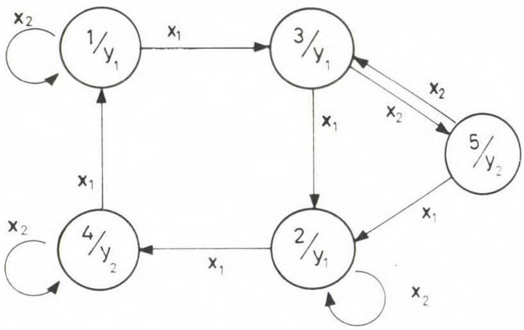


Fig. 6

Table 6

i	$\delta(i, x_1)$	$\delta(i, x_2)$	$\lambda(i)$
1	3	1	y_1
2	4	2	y_1
3	2	5	y_2
4	1	4	y_2
5	2	3	y_2

There are five indistinguishable state pairs. Among these, the pair 3, 5 belongs to type (III). One may check that the other four ones (the pairs 1, 2; 3, 4 and 4, 5) are of type (II). Condition (C) is not valid for this automaton.

4. On the automorphisms

We regard Proposition 5 and Problem 2 as the most important matter in this section. Also two other open questions are included, and we mention a few simply reachable related assertions on automorphisms which are more or less known in the literature.

PROPOSITION 3. *For any state pair a, b of a strongly connected automaton there exists at most one endomorphism α such that $\alpha(a) = b$.*

PROOF. Assume the existence of α, let c be an arbitrary state. There is an input word p fulfilling $\delta(a, p) = c$. The deduction

$$\alpha(c) = \alpha(\delta(a, p)) = \delta(\alpha(a), p) = \delta(b, p)$$

shows that the image $\alpha(c)$ is determined uniquely. □

COROLLARY 1 (Weeg [11], see Gécseg–Peák [8], Statement 5.5.5). *We have $\alpha(a) \neq a$ for each state a if α is a nontrivial endomorphism of a strongly connected automaton.* □

PROPOSITION 4 (Oehmke [10]). *Each endomorphism α of a strongly connected automaton is an automorphism.*

PROOF. Choose a state a arbitrarily. Our aim is to show that a belongs to the (precise) range of α. The first equality in the formula

$$\alpha(\delta(a, p)) = \delta(\alpha(a), p) = a$$

holds for every input word p, the second equality is true (by the strong connectedness) for at least one p. □

Recall the notation of the Main Question.

PROPOSITION 5. *The implication (c-2) $\Rightarrow$ (b-2) is valid for the strongly connected Moore automata* **A**. *More precisely: whenever α is a nontrivial automorphism and a is a state of* **A**, *then the pair $a, \alpha(a)$ is indistinguishable and belongs to type* (II).

PROOF. By Corollary 1 we have

$$\delta(a, p) \neq \alpha(\delta(a, p)) = \delta(\alpha(a), p)$$

for any $p(\in F(X))$, hence $a, \alpha(a)$ is a pair of type (II). The indistinguishability is a consequence of the deduction

$$\lambda(\delta(a, p)) = \lambda(\alpha(\delta(a, p))) = \lambda(\delta(\alpha(a), p)). \qquad \square$$

Example 7 (at the end of the preceding section) shows that implication (b-2) $\Rightarrow$ (c-2) cannot hold in full generality. Later we verify it under very particular circumstances (Proposition 9).

PROBLEM 1. Find (possibly wide) subclasses of the class of strongly connected automata in which (b-2) implies (c-2).

PROBLEM 2. Study the structural (e.g. symmetry) properties of the strongly connected automata which fulfil (b-2) but do not satisfy (C).

Denote by K_j the set of all states a of an automaton **A** which fulfil $\lambda(a) = y_j$ (where j can be $1, 2, \ldots, m$; $m = |Y|$). Next we state a fact which lies near to Theorem 2.9 in Deussen's paper [7] (where the strongly connected automata are called simple ones).

PROPOSITION 6. *Let $\mathfrak{G}$ be a subgroup of the full automorphism group $\mathfrak{A}(\mathbf{A})$ of a strongly connected automaton* $\mathbf{A}$. *Denote the order of $\mathfrak{G}$ by w. Consider the orbits $H_1, H_2, \ldots, H_t$ of $\mathfrak{G}$. Then the equalities*

$$|H_1| = |H_2| = \ldots = |H_t| = w$$

are valid and $t = |A|/w$, furthermore, w is a common divisor of the numbers $|K_1|, |K_2|, \ldots, |K_m|$.

PROOF. Fix an element a in an arbitrary orbit H_i. Let b run through all the states. If b belongs to H_i, then there is exactly one $\alpha(\in \mathfrak{G})$ such that $\alpha(a) = b$ (by Propositions 3 and 4); when $b \in A - H_i$, then $\alpha(a) \neq b$ for every $\alpha(\in \mathfrak{G})$. Thus $|H_i| = w$. It is clear that $wt = |A|$. The last conclusion follows by observing that any H_i is entirely included in some K_j because of the requirement $\lambda(a) = \lambda(\alpha(a))$ in the definition of endomorphisms. □

REMARK 1. Let $\mathbf{A}$ be a strongly connected automaton. Fix an input word p and define a mapping α_p of A into itself by $\alpha_p(a) = \delta(a, p)$. The mapping α_p is an automorphism of $\mathbf{A}$ if and only if p satisfies

$$(4.1) \quad \delta(a, px) = \delta(a, xp), \quad \lambda(a) = \lambda(\delta(a, p)) \qquad \text{for every } a(\in A),\ x(\in X).$$

The assertion of the remark becomes clear by noting that the formulae in (4.1) can be written as

$$\delta(\alpha_p(a), x) = \alpha_p(\delta(a, x)), \quad \lambda(a) = \lambda(\alpha_p(a)),$$

respectively. □

The following question can be raised on α_p (defined in Remark 1):

PROBLEM 3. Decide the validity of the following sentence: whenever α is an automorphism of a strongly connected automaton, then there is an input word p such that $\alpha = \alpha_p$.

EXAMPLE 8. The subsequent automaton shows that the answer to Problem 3 is negative when the requirement of strong connectedness is omitted: $A = \{a_1, a_2, a_3\}$, $X = \{x\}$, $Y = \{y_1, y_2\}$, $\delta(a_i, x) = a_3$ for $i \in \{1, 2, 3\}$, $\lambda(a_1) = \lambda(a_2) = y_1$, $\lambda(a_3) = y_2$. Indeed,

$$\alpha = \begin{pmatrix} a_1 & a_2 & a_3 \\ a_2 & a_1 & a_3 \end{pmatrix}$$

is an automorphism, but $\alpha \neq \alpha_p$ for any input word p.

5. Some consequences of the existence of an indistinguishable pair of second type

Denote by N the set of all states a of an automaton $\mathbf{A} = (A, X, Y, \delta, \lambda)$ such that the one-element set $\{a\}$ is a class of the maximal congruence $\pi_{\max}$

of **A**. We emphasize that Y is thought as the set of the actually occurring output symbols. We shall write $\operatorname{ind}\pi_{\max}$ for the index — i.e., the number of classes — of $\pi_{\max}$ and (sometimes) m for $|Y|$.

For a state a and an input word p let us denote by $k(a,p)$ the largest number k such that the states

$$a, \delta(a,p), \delta(a,p^2), \ldots, \delta(a,p^k)$$

are pairwise different. (Evidently, $0 \leqq k(a,p) < |A|$. Any $k(a,e)$ equals zero, but $p \neq e$ does not imply $k(a,e) > 0$ in general.) Denote by k_0 the minimum of the numbers $k(a,p)$ taken for all pairs $a(\in A)$, $p(\in F(X))$ such that $k(a,p) > 0$.

PROPOSITION 7. *Suppose that* **A** *is a strongly connected automaton having an indistinguishable state pair* c, d *of type* (II). *Then*

(a) $N = \emptyset$,

and

(b) *for any output symbol* y_i *there exist two states* a_i, b_i *such that* $\lambda(a_i) = \lambda(b_i) = y_i$ *and the pair* a_i, b_i *is indistinguishable and of type* (II).

PROOF. (a) Let a be an arbitrary state. There is a $p(\in F(X))$ such that $\delta(c,p) = a$. The pair $a, \delta(d,p)$ is proper and indistinguishable by the supposition posed on c and d. Thus $a \notin N$.

(b) Fix y_i. Let p_i be an input word fulfilling $\lambda(\delta(c,p_i)) = y_i$. Then the pair $a_i = \delta(c,p_i)$, $b_i = \delta(d,p_i)$ satisfies the conclusion. □

The next result has been got in common with F. Wettl:

PROPOSITION 8. *Suppose that a strongly connected automaton satisfies at least one of conditions* (c) *and* (d):

(c) *There is an indistinguishable state pair of type* (II).

(d) *For any input symbol* y_i *the states* a *fulfilling* $\lambda(a) = y_i$ *are not pairwise distinguishable.*

Then $|Y| \leqq |A|/(k_0 + 1)$.

PROOF. We form m sequences of states

$$\begin{array}{ll} (S_1) & a_{1,1}, a_{1,2}, \ldots, a_{1,s(1)}, \\ (S_2) & a_{2,1}, a_{2,2}, \ldots, a_{2,s(2)}, \\ & \ldots \\ (S_i) & a_{i,1}, a_{i,2}, \ldots, a_{i,s(i)}, \\ & \ldots \\ (S_m) & a_{m,1}, a_{m,2}, \ldots, a_{m,s(m)}, \end{array}$$

by the following two rules:

(1) For any i $(1 \leqq i \leqq m)$, the pair $a_{i,1}, a_{i,2}$ is proper, indistinguishable and satisfies $\lambda(a_{i,1}) = \lambda(a_{i,2}) = y_i$. (The possibility of this choice follows from (c) by Proposition 7, from (d) immediately.)

(2) The subsequences $a_{i,3}, \dots, a_{i,s(i)}$ are obtained in such a way that we consider an input word q_i satisfying $\delta(a_{i,1}, q_i) = a_{i,2}$, and we define $s(i), a_{i,j}$ by $s(i) = 1 + k(a_{i,1}, q_i)$ and $a_{i,j} = \delta(a_{i,1}, q_i^{j-1})$ (resp.).

It is evident that

$$\lambda(a_{i,1}) = \lambda(a_{i,2}) = \dots = \lambda(a_{i,s(i)}) = y_i$$

for each i, thus the sequences (S_i) are disjoint. Any (S_i) consists of pairwise different states by the definition of the number $k(a,p)$. Hence a state of $\mathbf{A}$ may occur at most once in the sequences, and the deduction

$$|A| \geqq \sum_{i=1}^{m} s(i) = \sum_{i=1}^{m} (1 + k(a_{i,1}, q_i)) \geqq m(1 + k_0) = |Y|(k_0 + 1)$$

is valid. □

Observe that the maximal possible value of $|Y|$ is $|A|/2$ if Proposition 8 applies.

PROPOSITION 9. *Let* $\mathbf{A}$ *be as in Proposition 7. If*

$$|Y| = |A|/2, \tag{5.1}$$

then

(e) *each class of* $\pi_{\max}$ *consists of exactly two states, and*

(f) $\mathbf{A}$ *has a nontrivial automorphism.*

PROOF. (e) We have

$$|A|/2 \geqq \operatorname{ind} \pi_{\max} \geqq |Y| \tag{5.2}$$

by the conclusion (a) of Proposition 7. Because of (5.1), equalities are true in (5.2).

(f) Denote the states of $\mathbf{A}$ by $a_1, b_1, a_2, b_2, \dots, a_m, b_m$ so that $\lambda(a_i) = \lambda(b_i) = y_i$ for every i $(1 \leqq i \leqq m)$. We can choose the subscripts in such a manner that the pair a_1, b_1 is indistinguishable and of type (II). The classes of $\pi_{\max}$ are obviously of form $\{a_i, b_i\}$. Consider the mapping

$$\alpha = \begin{pmatrix} a_1 & a_2 & \dots & a_m & b_1 & b_2 & \dots & b_m \\ b_1 & b_2 & \dots & b_m & a_1 & a_2 & \dots & a_m \end{pmatrix}.$$

Some of the properties of the automorphisms are clearly fulfilled by α, we have still to verify $\alpha(\delta(a,x)) = \delta(\alpha(a),x)$ for each $a (\in A)$ and $x (\in X)$.

First we show $\delta(a_i, x) \neq \delta(b_i, x)$ where $1 \leqq i \leqq m$. Let p be an input word such that $\delta(a_1, p) = a_i$. Then the deduction

$$\delta(a_i, x) = \delta(\delta(a_1, p), x) = \delta(a_1, px) \neq \delta(b_1, px) = \delta(\delta(b_1, p), x) = \delta(b_i, x)$$

is obvious, except the last step. $\delta(b_1, p) = b_i$ is true because

$$a_i = \delta(a_1, p) \equiv \delta(b_1, p) \qquad (\mathrm{mod}\ \pi_{\max})$$

and the pair a_1, b_1 is of type (II).

We have got that to each x and i there is a j such that

$$\{\alpha(\delta(a_i, x)),\ \alpha(\delta(b_i, x))\} = \{\delta(b_i, x), \delta(a_i, x)\} = \{a_j, b_j\}$$

(where $1 \leqq j \leqq m$; i and j are not necessarily different).

For completing the proof, let a state ($\in \{a_i, b_i\}$) and an input symbol be chosen. This can be done in four manners:

$$\alpha(\delta(a_i, x)) = a_j, \qquad \alpha(\delta(a_i, x)) = b_j$$
$$\alpha(\delta(b_i, x)) = a_j, \qquad \alpha(\delta(b_i, x)) = b_j.$$

By the analogy of these cases, it suffices to study the first possibility. We have then

$$\delta(a_i, x) = \alpha^2(\delta(a_i, x)) = \alpha(a_j) = b_j,$$

thus

$$\delta(\alpha(a_i), x) = \delta(b_i, x) \in \{a_j, b_j\} - \{b_j\} = \{a_j\},$$

hence

$$\alpha(\delta(a_i, x)) = a_j = \delta(\alpha(a_i), x). \qquad \square$$

6. On the simultaneous occurrence of indistinguishable state pairs belonging to different types

Let **A** be an automaton. In this section we denote the sets of indistinguishable proper state pairs (in A) of type (I) or (II) or (III) by S_1, S_2, S_3, respectively.

PROPOSITION 10. *Let* **A** *be a strongly connected automaton satisfying* $S_2 \neq \emptyset$. *If* $S_1 \cup S_3 \neq \emptyset$, *then* **A** *has at least five states.*

PROOF. Consider a pair c, d belonging to S_2 and a pair a, b lying in S_1 or S_3. There is a $p(\in F(X))$ such that $\delta(c, p) = a$. Denote $\delta(d, p)$ by f. The pair a, f belongs to S_2.

Choose a state g such that[8] $\lambda(g) \neq \lambda(a)$. There is a $q(\in F(X))$ such that $\delta(a, q) = g$. Denote $\delta(f, q)$ by h.

It is obvious that $\lambda(b) = \lambda(a) = \lambda(f)$. The pair g, h belongs to S_2 and $\lambda(g) = \lambda(h)$. We can observe that a, b, f, g, h are pairwise distinct states ($b \neq f$ because the pairs a, b and a, f belong to different types). $\square$

Example 5 shows that $|A| = 5$ is reachable under the suppositions of Proposition 10.

---

[8] Recall that the automata fulfilling $|Y| = 1$ were excluded in § 1.

PROPOSITION 11. *Let* **A** *be a strongly connected automaton fulfilling* $S_2 \neq \neq \emptyset$. *If the implication*

$$a \neq \delta(f,q) \Rightarrow \delta(a,q) \neq \delta(f,q^2) \tag{6.1}$$

is valid whenever the pair a, f *is an element of* S_2 *and* $q \in F(X)$, *then* $S_1 = \emptyset$.

PROOF. Assume that S_1 is non-empty, we are going to show the falsity of (6.1).

Consider again a pair c, d being in S_2. Now we can choose (by Proposition 2) a proper state pair a, b such that (c-1) is satisfied (i.e., $\lambda(a) = \lambda(b)$ and $\delta(a,x) = \delta(b,x)$ for every $x (\in X)$). Introduce p and f as in the proof of Proposition 10. Let q be an input word fulfilling $\delta(f,q) = b$. Denote $\delta(a,q)$ by g. It is easy to see that the pairs a, f and b, g are in S_2. Analogously to the preceding proof, $b \neq f$, hence $q \neq e$.

We can write q in the form $q = xr$ ($x \in X$, $r \in F(X)$). The deduction

$$\delta(a,q) = \delta(a,xr) = \delta(\delta(a,x),r) = \delta(\delta(b,x),r) = \\ = \delta(b,xr) = \delta(b,q) = \delta(\delta(f,q),q) = \delta(f,q^2)$$

is valid (the third equality follows from the choice of a and b), and $a \neq b = = \delta(f,q)$. □

We note that the implication (6.1) is fulfilled by any quasi-perfect automaton (for the definition of this class of automata see e.g. § 5.2 in [8]).

7. On partial quasi-isomorphisms

Recall the notion of partial quasi-isomorphism and the notations in the Main Question (Sections 1–2).

PROPOSITION 12. *The implication (c-3)* $\Rightarrow$ *(B) is valid for the strongly connected Moore automata* **A**. *More precisely: whenever* γ *is a partial quasi-isomorphism of* **A** *(from* J *to* K*) and* $a \in J$, *then the pair* $a, \gamma(a)$ *is indistinguishable.*

PROOF. Let us fix an arbitrary element a of J, assume that $p (= x_1 x_2 \dots x_k)$ runs through all the input words. We distinguish two cases.

Case 1. The inequality

$$\delta(a, x_1 x_2 \dots x_j) \neq \delta(\gamma(a), x_1 x_2 \dots x_j) \tag{7.1}$$

holds for each j where $0 \leqq j \leqq k$. It follows by induction on j that the states on the left-hand and right-hand sides of (7.1) belong to J and K, respectively; furthermore

$$\delta(\gamma(a), x_1 x_2 \dots x_j) = \gamma(\delta(a, x_1 x_2 \dots x_j))$$

for each j. Consequently

$$\lambda(\delta(a,p)) = \lambda(\gamma(\delta(a,p))) = \lambda(\delta(\gamma(a),p)). \tag{7.2}$$

Case 2. There is an h such that

$$\delta(a, x_1x_2 \dots x_h) = \delta(\gamma(a), x_1x_2 \dots x_h) \tag{7.3}$$

where $1 \leqq h \leqq k$. Then (7.3) holds also when the subscript h is replaced by an arbitrary j fulfilling $h < j \leqq k$; especially, $\delta(a,p) = \delta(\gamma(a),p)$, therefore

$$\lambda(\delta(a,p)) = \lambda(\delta(\gamma(a),p)). \tag{7.4}$$

By showing the validity of formulae (7.2) and (7.4) we have verified the indistinguishability of a and $\gamma(a)$. □

The next statement follows from Propositions 2, 5, 12:

COROLLARY 2. *The implication* (C) ⇒ (B) *is valid for the strongly connected Moore automata.* □

PROBLEM 4. Find (possibly wide) subclasses of the class of strongly connected automata in which the implications (b-3)⇒(c-3) and/or (c-3)⇒ ⇒(b-3) are true.

PROBLEM 5. Study the structural properties of the strongly connected automata that fulfil (b-3) but do not satisfy (C).

8. Further considerations

We know that the indistinguishability of two states is a congruence relation. Now a fact will be stated which is somewhat related to this (without supposing the strong connectedness).

REMARK 2. Consider two states a, b of an automaton **A** and two input words p, q. If the state pairs a, b and $\delta(a,p)$, $\delta(a,q)$ are indistinguishable, then the pair $\delta(b,p)$, $\delta(b,q)$ is indistinguishable, too.

The remarks holds since the deduction

$$\begin{aligned}\lambda(\delta(\delta(b,p),r)) &= \lambda(\delta(b,pr)) = \lambda(\delta(a,pr)) = \lambda(\delta(\delta(a,p),r)) = \\ &= \lambda(\delta(\delta(a,q),r)) = \lambda(\delta(a,qr)) = \lambda(\delta(b,qr)) = \lambda(\delta(\delta(b,q),r))\end{aligned}$$

is true for every choice of $r (\in F(X))$. □

Recall the definitions of $k(a,p)$ and k_0 (in Section 5).

PROPOSITION 13. *Let* **A** *be a strongly connected automaton. If* $|Y| > |A| - k_0$, *then* **A** *is simple.*

PROOF. Assume that **A** is not simple, our aim is to verify $|Y| \leqq |A| - k_0$. We can choose an indistinguishable proper pair a, b. There is an input word p such that $\delta(a,p) = b$, clearly $k(a,p) > 0$. We have $\delta(a, p^{i+1}) = \delta(b, p^i)$, thus $\delta(a, p^i)$ and $\delta(a, p^{i+1})$ are indistinguishable for each $i (\geqq 0)$.

The states in the sequence

$$a, \delta(a,p), \delta(a,p^2), \ldots, \delta(a, p^{k(a,p)})$$

are indistinguishable and pairwise different. Denote their set by B. The output symbol $\lambda(b)$ is common for the elements of B. Consequently,

$$|Y| \leqq 1 + |A - B| = 1 + |A| - |B| = 1 + |A| - (k(a,p)+1) = |A| - k(a,p) \leqq |A| - k_0$$

where the last inequality follows from the definition of k_0. □

REFERENCES

[1] ÁDÁM, A., Automata-leképezések, félcsoportok, automaták (Automaton mappings, semigroups, automata), *Mat. Lapok* **19** (1968), 327–343 (in Hungarian[9]). *MR* **40** #5350; *Zbl* **179**, 23

[2] ÁDÁM, A., On the question of description of the behaviour of finite automata, *Studia Sci. Math. Hungar.* **13** (1978), 105–124. *MR* **83i**:68079

[3] ÁDÁM, A., On the congruences of finite autonomous Moore automata, *Acta Cybernet. (Szeged)* **7** (1986), 259–279. *MR* **88d**:68075

[4] ÁDÁM, A., On simplicity-critical Moore automata, I, *Acta Math. Hungar.* **52** (1988), 165–174. *MR* **89i**:68098

[5] ÁDÁM, A., On simplicity-critical Moore automata, II, *Acta Math. Hungar.* **54** (1989), 291–296. *MR* **91a**:68197

[6] BABCSÁNYI, I., On the simplicity of cyclic Mealy automata, *Pure Math. and Appl. A* **1** (1991), 275–286. *MR* **93b**:68051

[7] DEUSSEN, P., On the algebraic theory of finite automata, *ICC Bull.* **4** (1965), 231–264. *MR* **33** #3844

[8] GÉCSEG, F. and PEÁK, I., *Algebraic theory of automata*, Akadémiai Kiadó, Budapest, 1972. *MR* **48** #10701

[9] (KARPOV, YU. G.) КАРПОВ, Ю. Г., О группе автоморфизмов конечного автомата, *Автоматика и Телемеханика* **1973**, no. 8, 70–74. *MR* **52** #16152

[9A] KARPOV, YU. G., Group of automorphisms of a finite automaton, *Automat. Remote Control* **1973**, no. 8, Part 1, 1261–1265.

[10] OEHMKE, R. H., On the structure of an automaton and its input semigroup, *J. Assoc. Comput. Mach.* **10** (1963), 521–525. *MR* **29** #4646

[11] WEEG, G. P., The group and semigroup associated with automata, *Proc. Sympos. Math. Theory of Automata* (1962), Polytechnic Press, Brooklyn, N.Y., 1963, 257–266. *MR* **30** #5899

(Received October 9, 1990)

[9] With a detailed English abstract, contained in the *Zentralblatt für Math.*, too.

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

BUDAPESTI MŰSZAKI EGYETEM
KÖZLEKEDÉSMÉRNÖKI KAR
MATEMATIKA TANSZÉK
MŰEGYETEM RKP. 9
H–1111 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 209–212

ON A GENERALIZATION OF AN OLD THEOREM OF ERDÖS

K. KOVÁCS

In 1946 Erdős [1] proved the following theorem:

THEOREM 1. *If the real-valued additive function f is monotonic then* $f(n) = c\log n$.

We need two of the several generalizations of the above theorem for our purpose.

THEOREM 2 ([2]). *If the real-valued additive function f is monotonic on a set of upper density one then $f(n) = c\log n$ for all $n \in \mathbb{N}$.*

THEOREM 3 ([3]). *Let $a > 0$ and b denote coprime integers. If $\|f(an+ +b)\|$ is monotonic ($\|\cdot\|$ denotes the Euclidean norm) for an additive function $f : \mathbb{N} \to \mathbb{R}^k$, then $f(n) = c\log n$ ($c \in \mathbb{R}^k$) for all $(n, a) = 1$.*

In this paper we examine some special cases of the following problems: What can we say if $\sum_{i=1}^{m} c_i f(a_i n + b_i)$ is monotonic for a real-valued additive function f or $\sum_{i=1}^{m} c_i \|f(a_i n + b_i)\|$ is monotonic for an additive function $f : \mathbb{N} \to \to \mathbb{R}^k$? $f(n)$ will be probably $c\log n$ for all $(n, \prod_{i=1}^{m} a_i) = 1$.

THEOREM 4. *Let a and b denote different integers. If*

$$f(n+a) - f(n+b) \geqq 0 \qquad (\text{or } \leqq 0, \text{ or monotonic}) \tag{1}$$

for a real-valued completely additive function on a set of upper density one then $f(n) = c\log n$ for all $n \in \mathbb{N}$.

THEOREM 5. *Let a and b denote different integers. If $f(n+a) + f(n+b)$ is monotonic for a real-valued completely additive function on a set of upper density one then $f(n) = c\log n$ for all $n \in \mathbb{N}$.*

1991 *Mathematics Subject Classification.* Primary 11A25.
Key words and phrases. Characterization of additive functions.

Supported by the Hungarian National Foundation for Scientific Research Grant No. 1901.

0081-6906/94/$ 4.00 ©1994 *Akadémiai Kiadó, Budapest*

THEOREM 6. *Let d denotes a fixed integer. If $\sum_{k=1}^{m} f(n+kd)$ is monotonic for a real-valued completely additive function on a set of upper density one then $f(n) = c\log n$ for all $n \in \mathbb{N}$.*

THEOREM 7. *Let $a \in \mathbb{N}$, $b, d \in \mathbb{Z}$ such that $a \mid b-d$ and $(b,d) = 1$. If*

$$f(an+b) - f(an+d) \geqq 0 \qquad (\text{or } \leqq 0, \text{ or monotonic}) \tag{2}$$

for $n \geqq n_0$ then $f(n) = c\log n$ for all $(n, b-d) = 1$.

COROLLARY. Theorem 7 implies a stronger result than Theorem 4 on the set of the natural numbers. If $f(n+b) - f(n+d)$ is monotonic for an additive function then $f(n) = c\log n$ for all $(n, b-d) = 1$.

REMARK. If f is completely additive in Theorem 7 then the condition $(b,d) = 1$ is not necessary.

THEOREM 8. *Let $a \in \mathbb{N}$, $b > d$ integers such that $a \mid b-d$ and $(b,d) = 1$. If $f(an+b) + f(an+d)$ is monotonic for $n \geqq n_0$ for a real-valued additive function then $f(n) = c\log n$ for all $(n, 2(b-d)) = 1$.*

COROLLARY. Theorem 8 implies a stronger result than Theorem 5 on the set of the natural numbers: *If $f(n+a) + f(n+b)$ is monotonic for an additive function then $f(n) = c\log n$ for all $(n, 2(b-d)) = 1$.*

THEOREM 9. *Let a and b denote different integers. If $\|f(n+a)\| - \|f(n+b)\| \geqq 0$ (or $\leqq 0$, or monotonic) for $n \geqq n_0$ then $f(n) = c\log n$.*

THEOREM 10. *Let $a \in \mathbb{N}$, $b, d \in \mathbb{Z}$ such that $a \mid b-d$ and $(b,d) = 1$. If $\|f(an+b)\| - \|f(an+d)\| \geqq 0$ (or $\leqq 0$, or monotonic) for $n \geqq n_0$ then $f(n) = c\log n$ for all $(n, a) = 1$.*

PROOF of THEOREM 4. Replacing n by $|a-b|n - b$ in (1) we get

$$f(n + \operatorname{sgn}(a-b)) - f(n) \geqq 0$$

on a set of upper density one. Theorem 2 yields our result. The other two cases in (1) can be treated similarly. (If the difference in (1) is monotonic then we get easily that the difference is $\geqq 0$ or $\leqq 0$ for $n \geqq n_0$.)

PROOF of THEOREM 5. We can assume that

$$F(n) = f(n+a) + f(n) \qquad (\text{where } a > 0)$$

is monotonic on a set of upper density one. $F(n) \geqq F(n-a)$ implies

$$f(n+a) \geqq f(n-a).$$

Replacing n by an we get

$$f(n+1) \geqq f(n-1).$$

Replacing n by $2n+1$ we have

$$f(n+1) \geqq f(n).$$

This can be ensured on a set of upper density one. By Theorem 2 we get our result.

PROOF of THEOREM 6. The case $d=0$ is Theorem 1. For $d \neq 0$ let us replace n by $|d|n$. We get that

$$F(n) = \sum_{k=1}^{m} f(n+(\operatorname{sgn} d)k)$$

is monotonic on a set of upper density one. The inequality $F(n+1) \geqq F(n)$ yields

$$f(n+m+1) \geqq f(n+1) \qquad \text{or} \qquad f(n-m+1) \geqq f(n-1).$$

By Theorem 4 we get that $f(n) = c \log n$.

PROOF of THEOREM 7. Replacing n by $|b-d|/a$ in (2) we have that

$$f(|b-d|n+b) - f(|b-d|n+d) \geqq 0 \qquad (\leqq 0 \text{ or monotonic}).$$

In each case f and even $|f|$ is monotonic on the arithmetical sequence $|b-d|n+b$. Using that $(b-d,b) = (b,d) = 1$ we have $f(n) = c\log n$ for all $(n, b-d) = 1$ by Theorem 3.

If f is completely additive we can take out $k = (b-d,b)$. Then we can apply Theorem 3 using that $\left(\frac{b-d}{k}, \frac{b}{k}\right) = 1$.

If $a=1$ and f is additive then the condition $a \mid b-d$ is trivially satisfied. If b and d are not coprime then replacing n by $n+t$ we can ensure that $b+t, d+t$ are positive and the bigger one is a prime. This yields that $(d+ + t, b+t) = 1$. We get $f(n) = c\log n$ for all $(n, b-d) = 1$.

REMARK. In Theorem 7 we cannot write $f(n) = c\log n$. The choice $f(p^\alpha) = c\log p^{\alpha-\alpha_0} + f(p^{\alpha_0})$ with $p^{\alpha_0} \parallel b-d$ contradicts this assertion.

PROOF of THEOREM 8. We get as in the proof of Theorem 7 that

$$G(n) = f((b-d)n+b) + f((b-d)n+d)$$

is monotonic. $G(n) \geqq G(n-1)$ implies

$$f((b-d)n+b) - f((b-d)n-b+2d) \geqq 0.$$

Applying Theorem 7 we get $f(n) = c\log n$ for all $(n, 2(b-d)) = 1$.

Theorem 10 can be proved similarly to Theorem 7 again.

Theorem 9 is a special case of Theorem 10 if $(b,d) = 1$. This can be achieved by a suitable replace of n by $n+t$. (See the proof of Theorem 7.)

REFERENCES

[1] Erdős, P., On the distribution function of additive functions, *Ann. of Math.* (2) **47** (1946), 1–20. *MR* **7**—416

[2] Birch, B. J., Multiplicative functions with non-decreasing normal order, *J. London Math. Soc.* **42** (1967), 149–151. *MR* **34** #2535

[3] Kovács, K., On the characterization of additive functions on residue classes, *Acta Math. Hungar.* **50** (1987), 123–125. *MR* **88e**:11007

(Received November 6, 1990)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR
ALGEBRA ÉS SZÁMELMÉLETI TANSZÉK
MÚZEUM KRT. 6–8
H–1088 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 213–231

EIGENVALUE PROBLEM FOR SOME NONLINEAR ELLIPTIC PARTIAL DIFFERENTIAL EQUATION

GABRIELLA BOGNÁR

0. Introduction

A well-known problem in mathematical physics is the determination of the principal frequency of a simply supported homogeneous membrane [5], [22], [19]. The function $u = u(x, y)$ describing the motion of the membrane satisfies the differential equation

$$\Delta u + \lambda u = 0 \tag{0.1}$$

in the bounded domain $D \in R^2$ with the boundary condition

$$u|_{\partial D} = 0$$

for fixed membrane (Dirichlet problem). It is known that this problem has eigensolutions $u_i \in L^2(D)$ and the corresponding eigenvalues are λ_i $(i = 1, 2, \dots)$ such that $\lambda_i \to \infty$ when $i \to \infty$. The eigenvalue λ_1 is the first (smallest, principal) one (see [5]).

For a given domain D the determination of the value λ_1 is known only for some special domains [16]. These are circle, circular segment, rectangle and three types of triangles with angles (60°, 60°, 60°), (45°, 45°, 90°), (30°, 60°, 90°). In general case the eigenvalue λ_1 can be obtained as the solution of the following variational problem. Let the function $f = f(x, y)$ be L^2-integrable, vanishing on ∂D, not identically zero in D and the partial derivatives f_x, f_y exist almost everywhere then

$$\lambda_1 = \min R\,[f],$$

where the ratio

$$R\,[f] = \frac{\displaystyle\iint_D (f_x^2 + f_y^2)\,dx\,dy}{\displaystyle\iint_D f^2\,dx\,dy} \tag{0.2}$$

1991 *Mathematics Subject Classification.* Primary 35J60; Secondary 35J40, 35J55, 35P30.

Key words and phrases. Nonlinear partial differential equations, eigenvalue problem, smallest eigenvalue.

is called the Rayleigh quotient. For the first eigenfunction u_1 the equality $\lambda_1 = R[u_1]$ holds concerning the linear partial differential equation (0.1) [8].

It is known that there is a connection between the eigenvalues and the shape of the domain. E.g. such a connection is the Rayleigh conjecture according to which for all membranes with a given area the circle has the minimum principal frequency. The method of proving this conjecture is the rearrangement method [6], [12], [18], [20]. This method was applied successfully also for the differential equation

$$\Delta u + f(u) = 0,$$

see Kawohl [11]. The Rayleigh quotient of the equation $\Delta u + \lambda f(u) = 0$ has the form

$$R[u] = \frac{\iint\limits_D (u_x^2 + u_y^2)\,dx\,dy}{\iint\limits_D uf(u)\,dx\,dy},$$

where $f(u)$ is continuous, $f(u) > 0$ for $u > 0$, $f(0) = 0$ and it satisfies the Lipschitz condition [13].

There are more possibilities for the generalization of the Rayleigh quotient to extend the space L^2 to fields L^{p+1}. Ôtani [14] and De Thelin [21] introduced the quotient

$$R[v] = \left(\frac{\|\nabla v\|_{L^{p+1}}}{\|v\|_{L^{p+1}}}\right)^{p+1} \qquad \text{in } W_0^{1,p+1}(D)\setminus\{0\}, \tag{0.3}$$

where $W_0^{1,p+1}(D)$ denotes the Banach space of functions $v \in C_0^1(D)$ with the norm

$$\|v\| = \left\{\iint\limits_D \left[|v|^{p+1} + |v_x|^{p+1} + |v_y|^{p+1}\right]\,dx\,dy\right\}^{\frac{1}{p+1}}.$$

In (0.3) $\|v\|_{L^{p+1}}$ and $\|\nabla v\|_{L^{p+1}}$ denote the $L^{p+1}(D)$ norm of v and of $|\nabla v| = \left[\sum\limits_{i=1}^{N}\left(\frac{\partial v}{\partial x_i}\right)^2\right]^{\frac{1}{2}}$, respectively. The Euler–Lagrange equation which corresponds to the variational problem of minimizing (0.3) has the form

$$\begin{aligned} -\operatorname{div}(|\nabla u|^{p-1}\nabla u) &= \lambda|u|^{p-1}u && \text{in } D \in R^N, \text{ if } 0 < p < \infty, \\ u &= 0 && \text{on } D. \end{aligned}$$

In this paper we give another generalization of (0.2) by modifying the norm of ∇v in (0.3) as follows:

$$\|\nabla v\|_{L^{p+1}} = \left[\iint\limits_D \left(\sum_{i=1}^{N}\left|\frac{\partial v}{\partial x_i}\right|^{p+1}\right)dx\,dy\right]^{\frac{1}{p+1}}. \tag{0.4}$$

where $p > 0$ is real. We also determine the corresponding Euler–Lagrange equation and examine some properties of the new Rayleigh quotient.

1. Preliminaries and results

Let D be a bounded domain in R^2, $p > 0$ be real and let the set F_D be defined as the set of the real functions $u = u(x,y) : D \to R$ satisfying the following conditions:

(I) $u|_{\partial D} = 0$,

(II) u is absolutely continuous for fixed x with respect to y and for fixed y with respect to x,

(III) $0 < \iint\limits_D |u_x|^{p+1}\,dx\,dy < \infty$ and $0 < \iint\limits_D |u_y|^{p+1}\,dx\,dy < \infty$.

Using the norm (0.4) we generalize the Rayleigh quotient (0.2) by

$$R[u] = \frac{\iint\limits_D \left[|u_x|^{p+1} + |u_y|^{p+1}\right] dx\,dy}{\iint\limits_D |u|^{p+1}\,dx\,dy}, \qquad u \in F_D. \tag{1.1}$$

The Euler–Lagrange equation of this variational problem is

$$\frac{\partial}{\partial x}\left(\left|\frac{\partial u}{\partial x}\right|^{p-1}\frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left(\left|\frac{\partial u}{\partial y}\right|^{p-1}\frac{\partial u}{\partial y}\right) + \lambda|u|^{p-1}u = 0, \qquad (x,y) \in D, \tag{1.2}$$

$$u = 0, \qquad (x,y) \in \partial D. \tag{1.3}$$

If problems (1.2) and (1.3) have a solution u then cu ($c =$ constant) is also a solution. The solutions can be made unique by the introduction of a suitable norm. Clearly, if $p = 1$, equation (1.2) is equivalent to the linear equation (0.1). As in the linear case if there exist constants λ_i and corresponding functions u_i ($i = 1, 2, \ldots$) for which problems (1.2) and (1.3) are satisfied then we call them eigenvalues and eigenfunctions, respectively. When u_x and u_y exist and are continuous, moreover $(u_x)^{\overset{*}{p}}$ and $(u_y)^{\overset{*}{p}}$ are differentiable almost everywhere in D, then u is called the classical solution of eigenvalue problems (1.2), (1.3).

The function $u \in W_0^{1,p+1}(D)$ is a weak solution of problems (1.2) and (1.3) if

$$\iint\limits_D \left(u_x^{\overset{*}{p}} v_x + u_y^{\overset{*}{p}} v_y\right) dx\,dy - \lambda \iint\limits_D u^{\overset{*}{p}} v\,dx\,dy = 0 \tag{1.4}$$

for all $v \in W_0^{1,p+1}(D)\backslash\{0\}$, where

$$x^{\overset{*}{p}} := \begin{cases} x^p & \text{if } x \geqq 0, \\ -|x|^p & \text{if } x < 0. \end{cases}$$

We introduce the value Λ by

$$\Lambda = \inf_{v \in F_D} R[v]. \tag{1.5}$$

The value Λ exists because, by (1.1), $R[v] > 0$. Clearly, $\Lambda \geqq 0$. In fact we shall show that $\Lambda > 0$. In the linear case it is known that $\Lambda = \lambda_1$ where λ_1 is the smallest eigenvalue [5]. We guess that $\Lambda = \lambda_1$ if $p \neq 1$. The aim of this paper is to prove someway this conjecture, see Remark 1.1. This conjecture is based on two known particular cases [3].

1. When D is a rectangle $\{(x,y) \mid 0 < x < a, 0 < y < b\}$ the eigenfunctions (classical solutions) are of product form

$$u_{k,l} = A_{k,l} S_p\left(\frac{k\hat{\pi}}{a}x\right) S_p\left(\frac{l\hat{\pi}}{b}y\right),$$

and the corresponding eigenvalues can be given in the form

$$\lambda_{k,l} = p\hat{\pi}^{p+1}\left(\frac{k^{p+1}}{a^{p+1}} + \frac{l^{p+1}}{b^{p+1}}\right), \qquad k,l = 1,2,\ldots,$$

where the function $S_p = S_p(x)$ is the generalized sine function [3].

2. It was also shown that problems (1.2) and (1.3) have classical solutions of radial form

$$v = v(\varrho),$$

if D is bounded by the central symmetric convex curve

$$|x|^{\frac{1}{p}+1} + |y|^{\frac{1}{p}+1} = 1. \tag{1.6}$$

In this case the partial differential equation (1.2) is equivalent to the equation

$$\frac{1}{\varrho}\frac{d}{d\varrho}\left[\varrho\left(\frac{d}{d\varrho}v\right)^{\overset{*}{p}}\right] + \lambda v^{\overset{*}{p}} = 0$$

and the boundary condition (1.3) is

$$v(1) = 0.$$

The first eigenvalue λ_1 and the corresponding eigenfunction are given approximately in [3] as a function of p. The curve defined by (1.6) was called figuratrix by H. Rund [17, p. 25] and isoperimetrix by H. Busemann [4, p. 168]. This curve plays the same role in the case of (1.2) as the unit circle in the linear case (0.1).

In this work we shall prove for Λ the following results.

THEOREM 1.1. *Let the domain D be bounded in the (x, y)-plane and denote the greatest diameter of D in the direction parallel to x-axis and to y-axis by d_x and d_y, respectively. Then*

$$\Lambda \max(d_x^{p+1}, d_y^{p+1}) > 1. \tag{1.7}$$

THEOREM 1.2. *For every λ_i of the eigenvalue problem* (1.2) *and* (1.3) *the relations $\lambda_i = R[u_i]$, $\lambda_i \geqq \Lambda$ $(i = 1, 2, \dots)$ hold.*

THEOREM 1.3. *Let u_1 be an eigenfunction of problems* (1.2) *and* (1.3) *with eigenvalue λ_1 such that $u_1 > 0$ in D. Then $\Lambda = \lambda_1$ and for each eigenvalue λ_i $(i \neq 1)$ the relation $\lambda_i > \lambda_1$ holds.*

REMARK 1.1. In the two special cases mentioned earlier the classical solutions of (1.2) and (1.3) have the property required in Theorem 1.3. We conjecture that $u_1 > 0$ is always satisfied. Instead of this, however, we can prove the following theorem.

THEOREM 1.4. *If $p \geqq 1$ then the eigenvalue problems* (1.2) *and* (1.3) *have a nonnegative nontrivial solution $u_1 \in W_0^{1,p+1}(D) \backslash \{0\}$ and $\lambda_1 = \Lambda$.*

The proofs of these theorems are given in the next section.

In Section 3 we shall deal with the Steiner and Schwarz symmetrizations of the domain D. Here we prove that the smallest eigenvalue λ_1 cannot be increased under these symmetrizations (see Theorems 3.1, 3.2).

We introduce the following notations:

$$D(u) = \iint_D \left(|u_x|^{p+1} + |u_y|^{p+1}\right)\, dx\, dy, \tag{1.8.}$$

$$\|u\|_{p+1}^{p+1} = \|u\|_{L^{p+1}(D)}^{p+1} = \iint_D |u|^{p+1}\, dx\, dy. \tag{1.9}$$

We shall use the following inequalities:

for the *reals* X, Y and $p > 0$

$$|X|^{p+1} + p|Y|^{p+1} - (1+p)XY^{\overset{*}{p}} = 0, \tag{1.10}$$

where equality holds if and only if $X = Y$ [10, p. 61];

for the *vectors* $\mathbf{a}_i \in R^3$ $(i = 1, 2, \dots, d)$ the triangular inequality in Minkowskian metric [10, p. 30]

$$\|\mathbf{a}_1 + \mathbf{a}_2 + \dots + \mathbf{a}_d\|_{p+1} \leqq \|\mathbf{a}_1\|_{p+1} + \|\mathbf{a}_2\|_{p+1} + \dots + \|\mathbf{a}_d\|_{p+1}; \tag{1.11}$$

for the *functions* $u, v \in L^{p+1}(D)$ the Clarkson inequality [1, p. 37]:

$$\left\|\frac{u+v}{2}\right\|_{p+1}^{p+1} + \left\|\frac{u-v}{2}\right\|_{p+1}^{p+1} = \frac{1}{2}\|u\|_{p+1}^{p+1} + \frac{1}{2}\|v\|_{p+1}^{p+1}, \qquad 1 \leqq p < \infty; \tag{1.12}$$

for the *geometric data* of D the generalized isoperimetric inequality in Minkowskian metric [9]:

$$L^2 - 4PA \geqq 0, \qquad P = 2\frac{p}{p+1} B\left(\frac{p}{p+1}, \frac{p}{p+1}\right), \tag{1.13}$$

where L is the length of the boundary of the domain D in Minkowskian metric, A is the usual area of D and P is a constant depending on the metric. In our case the indicatrix will be the curve given by $|x|^{p+1} + |y|^{p+1} = 1$. In (1.13) the equality holds if and only if the domain D is bounded by the so-called curve (c_ϱ):

$$|x|^{\frac{1}{p}+1} + |y|^{\frac{1}{p}+1} = \varrho^{\frac{1}{p}+1}, \qquad \varrho \in R^+. \tag{1.14}$$

REMARK 1.2. The Minkowskian length of this curve is $2P\varrho$ and the area bounded by curve (c_ϱ) is $P\varrho^2$.

2. Proofs

In the proof of Theorem 1.1 we shall give a generalization of a fundamental integral inequality formulated by H. Poincaré.

PROOF of THEOREM 1.1. Let the continuous function $u \in F_D$ be defined on D. We consider a subdomain D_γ with polygonal boundary ∂D_γ such that D_γ tends to D as $\gamma \to 0$. For sufficiently small γ we have $|u| < \varepsilon$ on ∂D_γ and $\varepsilon \to 0$. Through each point P in D_γ we draw a line $y =$ constant in the direction of x up to the intersection point P' of the line and ∂D_γ. Let $u(P') = \varepsilon(P)$ then $\varepsilon(P)$ is a piecewise continuous function and

$$|\varepsilon(P)| < \varepsilon. \tag{2.1}$$

Since

$$|u(P) - u(P')| = \left| \int_{P'}^{P} u_x \, dx \right|$$

by the Hölder inequality we obtain

$$|u(P) - \varepsilon(P')|^{p+1} \leqq d_x^p \int_{P'}^{P} |u_x|^{p+1} \, dx \leqq d_x^p \int_I |u_x|^{p+1} \, dx,$$

where I is the total intersection of the line $y =$ constant with D_γ, and d_x is the greatest diameter of the domain D in the direction of x. Integrating it over I we find that

$$\int_I |u(P) - \varepsilon(P)|^{p+1} \, dx \leqq d_x^{p+1} \int_I |u_x|^{p+1} \, dx.$$

Now we integrate with respect to y in D_γ. Drawing a line x = constant through P' we obtain in a similar manner

$$\iint_{D_\gamma} |u(P) - \varepsilon(P)|^{p+1}\, dx\, dy \leqq K \iint_{D_\gamma} \left[|u_x|^{p+1} + |u_y|^{p+1}\right]\, dx\, dy,$$

where $K = \max\left\{d_x^{p+1}, d_y^{p+1}\right\}$. By (2.1) and the triangle inequality we get

$$\sqrt[p+1]{\iint_{D_\gamma} |u(P)|^{p+1}\, dx\, dy} \leqq$$

$$\leqq \sqrt[p+1]{\iint_{D_\gamma} \varepsilon^{p+1}\, dx\, dy} + \sqrt[p+1]{K \iint_{D_\gamma} \left[|u_x|^{p+1} + |u_y|^{p+1}\right]\, dx\, dy}.$$

Since ε can be chosen arbitrarily small and $D_\gamma \to D$ when $\gamma \to 0$, therefore

$$\iint_D |u|^{p+1}\, dx\, dy \leqq K \iint_D \left[|u_x|^{p+1} + |u_y|^{p+1}\right]\, dx\, dy. \tag{2.2}$$

From (2.2) it follows

$$R[u] \geqq \frac{1}{K} > 0$$

and by (1.5) we get (1.7).

REMARK 2.1. Inequality (2.2) is the generalization of Poincaré inequality.

PROOF of THEOREM 1.2. We consider u_i as a nontrivial solution of (1.2), (1.3) and denote the corresponding eigenvalue by λ_i $(i = 1, 2, \dots)$. Thus we have the equation

$$\frac{\partial}{\partial x}\left(\frac{\partial u_i}{\partial x}\right)^{\overset{*}{p}} + \frac{\partial}{\partial y}\left(\frac{\partial u_i}{\partial y}\right)^{\overset{*}{p}} + \lambda_i u_i^{\overset{*}{p}} = 0.$$

Multiplying by u_i and integrating by parts over D we have

$$||(u_i)_x||_{p+1}^{p+1} + ||(u_i)_y||_{p+1}^{p+1} = \lambda_i ||u_i||_{p+1}^{p+1},$$

consequently

$$\lambda_i = R[u_i].$$

From (1.5) we obtain

$$\Lambda = \inf_{u \in F_D} R[u] \leqq R[u_i] = \lambda_i.$$

PROOF of THEOREM 1.3. Clearly, $u_1 \in F_D$ and the eigenfunction u_1 has no zero in D. As in Theorem 1.1 we consider the subdomain D_γ of D. Applying the method of Beesack [2] we substitute $X = u_x$ and $Y = u\dfrac{(u_1)_x}{u_1}$ in inequality (1.10) and integrate it over D_γ we find

$$\iint\limits_{D_\gamma} \left[|u_x|^{p+1} + p\left|u\frac{(u_1)_x}{u_1}\right|^{p+1} - (1+p)u_x u^{\overset{*}{p}}\left(\frac{(u_1)_x}{u_1}\right)^{\overset{*}{p}} \right] dx\,dy \geqq 0. \tag{2.3}$$

Through every point $P \in D$ we draw a line $y = \text{const}$. The intersection points of this line with D have coordinates $x_1, x_2 (\geqq x_1)$. Consequently, the intersection points by D_γ have coordinates $x_1 + \gamma$, $x_2 - \gamma$. We integrate the third term in (2.3) by parts for fixed y, so we obtain

$$\int\limits_{x_1+\gamma}^{x_2-\gamma} (1+p)u_x u^{\overset{*}{p}}\left(\frac{(u_1)_x}{u_1}\right)^{\overset{*}{p}} dx =$$

$$= \left[|u|^{p+1}\left(\frac{(u_1)_x}{u_1}\right)^{\overset{*}{p}}\right]_{x_1+\gamma}^{x_2-\gamma} - \int\limits_{x_1+\gamma}^{x_2-\gamma} |u|^{p+1}\frac{\left\{[(u_1)_x]^{\overset{*}{p}}\right\}_x}{u_1^{\overset{*}{p}}}\,dx + p\int\limits_{x_1+\gamma}^{x_2-\gamma}\left|u\frac{(u_1)_x}{u_1}\right|^{p+1} dx.$$

By Condition II of the set F_D and the Hölder inequality we obtain

$$|u(x_1+\gamma)|^{p+1} = \left|\int\limits_{x_1}^{x_1+\gamma} u_x\,dx\right|^{p+1} \leqq \gamma^p \int\limits_{x_1}^{x_1+\gamma} |u_x|^{p+1}\,dx,$$

therefore

$$\lim_{\gamma\to 0}\left[|u|^{p+1}\left(\frac{(u_1)_x}{u_1}\right)^{\overset{*}{p}}\right]_{x_1+\gamma} \leqq \lim_{\gamma\to 0}\int\limits_{x_1}^{x_1+\gamma}|u_x|^{p+1}\,dx = 0.$$

In a similar way we get

$$\lim_{\gamma\to 0}\left[|u|^{p+1}\left(\frac{(u_1)_x}{u_1}\right)^{\overset{*}{p}}\right]^{x_2-\gamma} \leqq 0,$$

so

$$\left[|u|^{p+1}\left(\frac{(u_1)_x}{u_1}\right)^{\overset{*}{p}}\right]_{x_1+\gamma}^{x_2-\gamma} = 0.$$

From (2.3) we have

$$\iint_D \left[|u_x|^{p+1} + |u|^{p+1} \frac{\{[(u_1)_x]^{p}\}_x}{u_1^{\overset{*}{p}}} \right] dx\, dy = 0. \tag{2.4}$$

Similarly, substituting $X = u_y$ and $Y = u\dfrac{(u_1)_y}{u_1}$ to inequality (1.10) we find that

$$\iint_D \left[|u_y|^{p+1} + |u|^{p+1} \frac{\left\{[(u_1)_y]^{\overset{*}{p}}\right\}_y}{u_1^{\overset{*}{p}}} \right] dx\, dy \geqq 0. \tag{2.5}$$

We add the inequalities above and we find that

$$\iint_D \left[|u_x|^{p+1} + |u_y|^{p+1} + |u|^{p+1} \frac{\left\{[(u_1)_x]^{\overset{*}{p}}\right\}_x + \left\{[(u_1)_y]^{\overset{*}{p}}\right\}_y}{u_1^{\overset{*}{p}}} \right] dx\, dy \geqq 0. \tag{2.6}$$

Since u_1 solves (1.2) with λ_1 therefore

$$\iint_D \left[|u_x|^{p+1} + |u_y|^{p+1} - \lambda_1 |u|^{p+1} \right] dx\, dy \geqq 0. \tag{2.7}$$

Equality holds if and only if there is equality in (2.4) and (2.5) that is

$$u_x = u\frac{(u_1)_x}{u_1}, \qquad u_y = u\frac{(u_1)_y}{u_1}$$

and

$$\left(\frac{u}{u_1}\right)_x = \left(\frac{u}{u_1}\right)_y = 0$$

so $u = cu_1$ ($c = $ constant). From (2.7) it follows that

$$R[u] \geqq \lambda_1,$$

consequently,

$$\Lambda \geqq \lambda_1.$$

Theorem 1.2 implies that $\Lambda = \lambda_1$ and for every eigenvalue λ_i the inequality $\lambda_i > \lambda_1$ $(i \neq 1)$ holds.

COROLLARY 2.1. *If $u_1 > 0$ in D then u_1 has no local minimum in D.*

PROOF. We assume that the first eigenfunction $u_1 \in F_D$ has a local minimum in D. Then there exists the domain $D' \subset D$ in which u_1 is less than $\min u_1|_{\partial D'}$. In D' we substitute the function u_1 by a plane parallel to the plane (x, y) in height $\min u_1|_{\partial D'}$. By this method we get a new eigenfunction u_1' in D and a corresponding Rayleigh quotient $R[u_1']$. Comparing $R[u_1']$ with $R[u_1]$ we find that in $R[u_1']$ the numerator is less and the denominator is greater than in $R[u_1]$ (see (1.1)), hence

$$R[u_1'] < R[u].$$

Applying Theorems 1.2 and 1.3 we get that λ_1 is not the smallest eigenvalue. It is a contradiction.

PROOF of THEOREM 1.4. We choose a minimizing sequence $\{u_m\}_{m=1}^{\infty} \subset \subset W_0^{1,p+1}(D)$ such that $\|u_m\|_{p+1}^{p+1} = 1$, $D(u_m) \to \Lambda$ and $p \geqq 1$. Since $W_0^{1,p+1}(D)$ is compactly imbedded in $L^{p+1}(D)$ [1, p. 144] there exists a subsequence $\{u_{m_k}\}_{k=1}^{\infty}$ and $u_{m_k} \to u$ with $\|u\|_{p+1}^{p+1} = 1$ as $k \to \infty$ in $L^{p+1}(D)$. By (1.12) we have for any l, n and $u_l, u_n \in \{u_{m_k}\}$

$$D\left(\frac{u_l - u_n}{2}\right) + D\left(\frac{u_l + u_n}{2}\right) \leqq \frac{1}{2}D(u_l) + \frac{1}{2}D(u_n).$$

We obtain by (1.5)

$$0 < D\left(\frac{u_l - u_n}{2}\right) \leqq \frac{1}{2}D(u_l) + \frac{1}{2}D(u_n) - \Lambda\left\|\frac{u_l + u_n}{2}\right\|_{p+1}^{p+1}.$$

As $l \to \infty$, $n \to \infty$ we get

$$\frac{1}{2}D(u_l) + \frac{1}{2}D(u_n) - \Lambda\left\|\frac{u_l + u_n}{2}\right\|_{p+1}^{p+1} \to 0.$$

Now we see that for $\{u_{m_k}\}_{k=1}^{\infty}$

$$\left\{\frac{\partial}{\partial x}u_{m_k}\right\}_{k=1}^{\infty} \to v_1 \qquad \text{and} \qquad \left\{\frac{\partial}{\partial y}u_{m_k}\right\}_{k=1}^{\infty} \to v_2$$

are Cauchy sequences in $L^{p+1}(D)$. It is easy to show that

$$\frac{\partial u}{\partial x} = v_1 \qquad \text{and} \qquad \frac{\partial u}{\partial y} = v_2$$

almost everywhere in D. Hence $u_m \to u$ in $W_0^{1,p+1}(D)$ and $D(u) = \Lambda$ with $\|u\|_{p+1}^{p+1} = 1$. Furthermore, since $|u|$ also satisfies (1.1), $|u|$ becomes a nonnegative nontrivial solution of the eigenvalue problem (1.2), (1.3).

3. Application of the symmetrizations

Let us consider the bounded domain D and the continuous function $u(x,y)$ on D satisfying $u_{|\partial D} = 0$. We introduce the level set D_c of u for $c \in R$ by

$$D_c = \{(x,y) \in \bar{D} \mid u(x,y) \geqq c\}.$$

Clearly,

$$D_{c'} \supset D_{c''} \qquad \text{if } u_{\min} \leqq c' < c'' \leqq u_{\max}$$

and

$$D_c = D \qquad \text{if } c < u_{\min},$$
$$D_c = \emptyset \qquad \text{if } c > u_{\max}.$$

By these properties we can reconstruct the function u from the level sets $\{D_c\}$.

There are several rearrangements, see [11]. We shall deal with two types of rearrangements on the plane. We shall denote the rearrangement of G by $G^{(r)}$. These rearrangements will have the property: if $G_1 \subset G_2$ then $G_1^{(r)} \subset G_2^{(r)}$.

Let the continuous function u be defined on D and its level sets be denoted by $D_c (c \in R)$. Using the rearrangement of sets D_c we obtain $D_c^{(r)}$. The family of the level sets $D_c^{(r)}$ determines uniquely the function $u^{(r)}$ on $D^{(r)}$ such that the level sets of $u^{(r)}$ are $D_c^{(r)}$. The function $u^{(r)}$ is called the rearrangement of function u. The rearrangements have the following properties (see [11, p. 20]):

(a) if u is Lipschitz continuous then $u^{(r)}$ is also Lipschitz continuous;

(b) the functions u and $u^{(r)}$ are equimeasurable;

(c) the mapping $u \to u^{(r)}$ is order preserving.

Steiner symmetrizations

The Steiner symmetrization of the first kind is a geometrical transformation connected with a certain line on the plane. First we shall choose the x-axis as the line of symmetrization.

Let D be a bounded (open) domain in the plane (x,y). Let $D(x') = D \cap \{(x',y) \mid y \in R\}$ be defined for any $x' \in R$. $D(x')$ is empty when x' tends to $+\infty$ or $-\infty$ and the points x' for which $D(x') \neq \emptyset$ belong to an open interval I. Clearly, $D(x')$ is a union of open line sections. We denote the sum of the lengths of these sections by $|D(x')|$. Let the Steiner symmetrization $D^{(x)}$ of D with respect to the x-axis be defined by

$$D^{(x)} = \bigcup_{x' \in I} D^{(x)}(x'),$$

where

$$D^{(x)}(x') = \begin{cases} \{(x', y) \in R^2 \mid |y| < \frac{1}{2}|D(x')|\} & \text{if } x' \in I, \\ \emptyset & \text{if } x' \notin I. \end{cases}$$

Let the continuous function $u(x, y)$ be given by its level sets D_c. The function $u^{(x)}$ — called the Steiner symmetrization of u with respect to the x-axis — is reconstructed from the level sets $D_c^{(x)}$.

Let σ be a two-dimensional manifold in R^3. If σ can be given by the equation $z = u(x, y)$ with continuous first partial derivatives u_x, u_y and R_{xy} is orthogonal projection of σ onto the plane (x, y) then we define the surface area of σ in Minkowskian metric as follows

$$S_{xy} = \iint\limits_{R_{xy}} \left[1 + |u_x|^{p+1} + |u_y|^{p+1}\right]^{\frac{1}{p+1}} dx\, dy.$$

This surface area can be based in classical sense by subdividing σ into h small triangles $\sigma_1, \sigma_2, \ldots, \sigma_h$ with surface areas $\Delta S_1, \Delta S_2, \ldots, \Delta S_h$ and by forming the sum

$$\sum_{i=1}^{h} \Delta S_i.$$

By defining suitably the areas of σ_i $(i = 1, 2, \ldots, h)$ this sum tends to a limit as $h \to \infty$ and

$$S_{xy} = \lim_{h \to \infty} \sum_{i=1}^{h} \Delta S_i.$$

Let σ be given by the equation $y = v(x, z)$ and let R_{xz} be its projection on the plane (x, z). If v has continuous first partial derivatives on R_{xz} then

$$S_{xz} = \iint\limits_{R_{xz}} \left[1 + |v_x|^{p+1} + |v_z|^{p+1}\right]^{\frac{1}{p+1}} dx\, dz.$$

Let σ be given by the equation $x = w(y, z)$, and R_{yz} by its projection on the plane (y, z). If w has continuous first partial derivatives on R_{yz} then

$$S_{yz} = \iint\limits_{R_{yz}} \left[1 + |w_y|^{p+1} + |w_z|^{p+1}\right]^{\frac{1}{p+1}} dy\, dz.$$

If the first partial derivatives of u, v and w are not equal to zero then

$$S_{xy} = S_{xz} = S_{yz}. \tag{3.1}$$

Equality (3.1) holds when we consider the surface area of a part of σ.

We shall prove the following lemma concerning the surface area of a two-dimensional manifold:

LEMMA 3.1. *Under Steiner symmetrization the surface area of a two-dimensional manifold with Minkowskian metric does not increase.*

PROOF. We consider a two-dimensional manifold σ having sufficiently smooth surface. Let a straight line parallel to the y-axis intersect σ at the points: (x, y_1, z), (x, y_2, z), $\ldots$, (x, y_{2k}, z) where $y_1 > y_2 > \cdots > y_{2k}$, $k \geqq 1$, and the Steiner symmetrization of σ at the points $(x, -y, z)$, (x, y, z), where $y = \frac{1}{2} \sum_{m=1}^{2k} (-1)^{m-1} y_m$.

The surface area of σ can be expressed in Minkowskian metric as follows:

$$S = \frac{1}{2} \iint \sum_{m=1}^{2k} \left[1 + \left| \frac{\partial y_m}{\partial x} \right|^{p+1} + \left| \frac{\partial y_m}{\partial z} \right|^{p+1} \right]^{\frac{1}{p+1}} dx\, dz,$$

where the domain of integration is the orthogonal projection of σ onto the plane of symmetrization (x, z). The surface area of the Steiner symmetrization of σ has the form

$$S^{(x)} = \iint \left[1 + \left| \frac{\partial y}{\partial x} \right|^{p+1} + \left| \frac{\partial y}{\partial z} \right|^{p+1} \right]^{\frac{1}{p+1}} dx\, dz =$$

$$= \iint \left[1 + \left| \frac{1}{2} \sum_{m=1}^{2k} (-1)^{m-1} \frac{\partial y_m}{\partial x} \right|^{p+1} + \left| \frac{1}{2} \sum_{m=1}^{2k} (-1)^{m-1} \frac{\partial y_m}{\partial z} \right|^{p+1} \right]^{\frac{1}{p+1}} dx\, dz.$$

By applying the Minkowski inequality (1.11) we obtain

$$\left(k + \left| \frac{1}{2} \sum_{m=1}^{2k} (-1)^{m-1} \frac{\partial y_m}{\partial x} \right|^{p+1} + \left| \frac{1}{2} \sum_{m=1}^{2k} (-1)^{m-1} \frac{\partial y_m}{\partial z} \right|^{p+1} \right)^{\frac{1}{p+1}} \leqq$$

$$\leqq \frac{1}{2} \sum_{m=1}^{2k} \left(1 + \left| \frac{\partial y_m}{\partial x} \right|^{p+1} + \left| \frac{\partial y_m}{\partial z} \right|^{p+1} \right)^{\frac{1}{p+1}},$$

hence

$$S \geqq S^{(x)}.$$

REMARK 3.1. If $k > 1$ then $S > S^{(x)}$.

In the case $p = 1$ G. Pólya and G. Szegő [15] examined the first eigenvalue connecting D and $D^{(x)}$. They proved that the Steiner symmetrization diminished the first eigenvalue. Our aim is to generalize this statement for the case $p > 0$.

THEOREM 3.1. *If the first eigenvalue problem* (1.2), (1.3) *is solvable and there exist smallest eigenvalues* λ_1 *for* D *and* $\lambda_1^{(x)}$ *for* $D^{(x)}$ *and* $\Lambda^{(x)} = \lambda_1^{(x)}$ *then* $\lambda_1 \geqq \lambda_1^{(x)}$.

PROOF. Let the first eigenfunction $u_1 \in F_D$ be given with level sets D_c. We consider the two-dimensional manifold σ given by the equation $z = u_1(x, y)$ and $R_{xy} = D$. The Steiner symmetrization of the level sets D_c with respect to the x-axis transforms D into $D^{(x)}$ and u_1 into $u_1^{(x)}$. By Lemma 3.1, the surface area of σ is diminished under symmetrization. Since the area of D and $D^{(x)}$ are equal, therefore we have by (3.1) that

$$\begin{aligned}(3.2)\qquad &\iint\limits_D \left[1 + |u_{1x}|^{p+1} + |u_{1y}|^{p+1}\right]^{\frac{1}{p+1}} dx\,dy \geqq \\ &\geqq \iint\limits_{D^{(x)}} \left[1 + \left|u_{1x}^{(x)}\right|^{p+1} + \left|u_{1y}^{(x)}\right|^{p+1}\right]^{\frac{1}{p+1}} dx\,dy.\end{aligned}$$

Applying inequality (3.2) to δu_1 instead of u_1 — where $\delta > 0$ is arbitrary constant — we get

$$\begin{aligned}&\iint\limits_D \left[1 + \delta^{p+1}\left(|u_{1x}|^{p+1} + |u_{1y}|^{p+1}\right)\right]^{\frac{1}{p+1}} dx\,dy \geqq \\ &\geqq \iint\limits_{D^{(x)}} \left[1 + \delta^{p+1}\left(\left|u_{1x}^{(x)}\right|^{p+1} + \left|u_{1y}^{(x)}\right|^{p+1}\right)\right]^{\frac{1}{p+1}} dx\,dy.\end{aligned}$$

Since

$$(1 + \delta^{p+1}x)^{\frac{1}{p+1}} \approx 1 + \frac{1}{p+1}\delta^{p+1}x + \dots$$

we have

$$\begin{aligned}&\iint\limits_D \left[1 + \frac{1}{p+1}\delta^{p+1}\left(|u_{1x}|^{p+1} + |u_{1y}|^{p+1}\right) + \dots\right] dx\,dy \geqq \\ &\geqq \iint\limits_{D^{(x)}} \left[1 + \frac{1}{p+1}\delta^{p+1}\left(\left|u_{1x}^{(x)}\right|^{p+1} + \left|u_{1y}^{(x)}\right|^{p+1}\right) + \dots\right] dx\,dy.\end{aligned}$$

Let us observe that the area of D and $D^{(x)}$ are equal. By the limiting procedure $\delta \to \infty$ we obtain

$$(3.3)\qquad \iint\limits_D \left(|u_{1x}|^{p+1} + |u_{1y}|^{p+1}\right) dx\,dy \geqq \iint\limits_{D^{(x)}} \left(\left|u_{1x}^{(x)}\right|^{p+1} + \left|u_{1y}^{(x)}\right|^{p+1}\right) dx\,dy.$$

By Property (C) in [11, p. 22] we obtain

$$\iint_D |u_1|^{p+1}\,dx\,dy = \iint_{D^{(x)}} \left|u_1^{(x)}\right|^{p+1} dx\,dy. \tag{3.4}$$

By Theorem 1.2 we have

$$\lambda_1 = \frac{\displaystyle\iint_D \left(|u_{1x}|^{p+1} + |u_{1y}|^{p+1}\right) dx\,dy}{\displaystyle\iint_D |u_1|^{p+1}\,dx\,dy}.$$

Making use (3.3) and (3.4) we get

$$\lambda_1 \geqq \frac{\displaystyle\iint_{D^{(x)}} \left(\left|u_{1x}^{(x)}\right|^{p+1} + \left|u_{1y}^{(x)}\right|^{p+1}\right) dx\,dy}{\displaystyle\iint_{D^{(x)}} \left|u_1^{(x)}\right|^{p+1} dx\,dy} \geqq \inf_{v\in F_{D^{(x)}}} R[v] = \Lambda^{(x)}.$$

By our assumption $\Lambda^{(x)} = \lambda_1^{(x)}$, hence the theorem follows.

Remark 3.2. Denote the level sets by $D_c^{(x)(y)}$ if the level sets $D_c^{(x)}$ is further symmetrized with respect to $x = 0$. For the smallest eigenvalue corresponding to the new central-symmetric level sets $D_c^{(x)(y)}$ of $u^{(x)(y)}$ we have the inequalities

$$\lambda_1 \geqq \lambda_1^{(x)} \geqq \lambda_1^{(x)(y)}, \qquad \lambda_1 \geqq \lambda_1^{(y)} \geqq \lambda_1^{(y)(x)}.$$

However, we also have the relation $D^{(x)(y)} = D^{(y)(x)}$, hence $\lambda_1^{(x)(y)} = \lambda_1^{(y)(x)}$.

Schwarz symmetrization

The central-symmetric Schwarz symmetrization is the most frequently used type of symmetrization. We define the Schwarz symmetrization $D^{(0)}$ of $D \subset R^2$ by

$$D^{(0)} = \begin{cases} \text{the curve } (c_\varrho) \text{ of the same area as } D & \text{if } D \neq \emptyset, \\ \emptyset & \text{if } D = \emptyset, \end{cases}$$

where the curve (c_ϱ) is given by (1.13) (see also Remark 1.2). Let the function u be given by its level sets D_c. The function $u^{(0)}$ called the Schwarz symmetrization of u with respect to the origin is reconstructed from the level sets $D_c^{(0)}$.

THEOREM 3.2. *If the first eigenvalue problem* (1.2), (1.3) *is solvable and there exist smallest eigenvalues* λ_1 *for* D *and* $\lambda_1^{(0)}$ *for* $D^{(0)}$ *then* $\lambda_1 > \lambda_1^{(0)}$ *unless* ∂D *is a curve* (c_ϱ).

PROOF. We consider the level sets

$$D_c = \{(x,y) \in \bar{D},\ u_1(x,y) \geqq c\}, \qquad c \in R,$$

of the first eigenfunction u_1. Now we introduce instead of the coordinates x, y the new coordinates w and s.

The intersection of the plane $w = c$ with the surface of $w = u_1$ gives the level sets D_c. Therefore we get $0 \leqq w \leqq a$, where a is the maximum value of u_1. Let the coordinate s be the arc length of the level line from 0 to the total length $L(w)$ of ∂D_c. We have the following relations:

$$\frac{\partial u_1}{\partial w} = 1, \qquad \frac{\partial u_1}{\partial s} = 0$$

and

$$\left|\frac{\partial x}{\partial s}\right|^{p+1} + \left|\frac{\partial y}{\partial s}\right|^{p+1} = 1,$$

that is

$$\frac{\partial u_1}{\partial x}\frac{\partial x}{\partial w} + \frac{\partial u_1}{\partial y}\frac{\partial y}{\partial w} = 1,$$
$$\frac{\partial u_1}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial u_1}{\partial y}\frac{\partial y}{\partial s} = 0.$$

For the new coordinates we have the Jacobian

$$\Delta = \begin{vmatrix} \dfrac{\partial x}{\partial w} & \dfrac{\partial x}{\partial s} \\[2ex] \dfrac{\partial y}{\partial w} & \dfrac{\partial y}{\partial s} \end{vmatrix}.$$

Hence we obtain

$$|(u_1)_x|^{p+1} + |(u_1)_y|^{p+1} = \frac{1}{|\Delta|^{p+1}}$$

so

$$D(u_1) = \int_{w=0}^{a} \int_{s=0}^{L(w)} \frac{1}{|\Delta|^p}\, ds\, dw.$$

We denote by $A(w)$ the area of D_c that is

$$A(w) = \iint_{D_c} dx\, dy = \int_{w}^{a} \int_{s=0}^{L(w)} |\Delta|\, ds\, dw,$$

and $A(0) = \text{mes}\, D$, $A(a) = 0$. We have

$$A'(w) = -\int_{s=0}^{L(w)} |\Delta|\, ds \tag{3.5}$$

and by Hölder inequality and (1.12) we obtain

$$\int_{s=0}^{L(w)} |\Delta|\, ds \left(\int_{s=0}^{L(w)} \frac{ds}{|\Delta|^p} \right)^{\frac{\overset{*}{1}}{p}} \geqq |L(w)|^{\frac{1}{p}+1} \geqq [4PA(w)]^{\frac{p+1}{2p}}. \tag{3.6}$$

Applying (3.5) we get

$$\int_{s=0}^{L(w)} \frac{ds}{|\Delta|^p} \geqq [-A'(w)]^{-\overset{*}{p}} [4PA(w)]^{\frac{p+1}{2}},$$

therefore

$$D(u_1) \geqq (4P)^{\frac{p+1}{2}} \int_{w=0}^{a} [-A'(w)]^{-\overset{*}{p}} [A(w)]^{\frac{p+1}{2}}\, dw. \tag{3.7}$$

Since $u_1^{(0)}$ is symmetric function we can write

$$u_1^{(0)}(x, y) = v(\varrho).$$

By Remark 1.2 we have

$$D\left(u_1^{(0)}\right) = \int_{w=0}^{a} \int_{s=0}^{L(w)} \left| \frac{dv}{d\varrho} \right|^p ds\, dw = \int_{w=0}^{a} 2P\varrho \left| \frac{dw}{d\varrho} \right|^p dw =$$

$$= 2^{p+1} P \int_{w=0}^{a} \varrho^{p+1} \left| \frac{dw}{d(\varrho^2)} \right|^p dw.$$

Similarly

$$\varrho = \left(\frac{A(w)}{P} \right)^{\frac{1}{2}} \qquad \text{and} \qquad \left| \frac{dw}{d(\varrho^2)} \right| = -\frac{P}{A'(w)}$$

then we obtain

$$D\left(u_1^{(0)}\right) = (4P)^{\frac{p+1}{2}} \int_{w=0}^{a} [-A'(w)]^{-\overset{*}{p}} [A(w)]^{\frac{p+1}{2}}\, dw. \tag{3.8}$$

Consequently, by (3.7), we obtain

$$D(u_1) \geqq D\left(u_1^{(0)}\right).$$

Making use of property (C) in [11, p. 22] we get

$$\|u_1\|_{p+1}^{p+1} = \iint_{D^{(0)}} |u_1^{(0)}|^{p+1}\, dx\, dy. \tag{3.9}$$

Therefore, by (3.9), we have

$$\lambda_1 = \frac{D(u_1)}{\|u_1\|_{p+1}^{p+1}} \geqq \frac{\iint_{D^{(0)}} \left(|u_{1x}^{(0)}|^{p+1} + |u_{1y}^{(0)}|^{p+1}\right) dx\, dy}{\iint_{D^{(0)}} |u_1^{(0)}|^{p+1}\, dx\, dy} \geqq \inf_{v \in F_{D^{(0)}}} R[v] = \Lambda^{(0)}.$$

By Theorem 1.3, $\Lambda^{(0)} = \lambda_1^{(0)}$ therefore $\lambda_1 \geqq \lambda_1^{(0)}$.

From the isoperimetric inequality (1.13) it follows that in (3.7) we have equality (also in (3.6), respectively) if the level lines are curves (c_ϱ). Then $\lambda_1 > \lambda_1^{(0)}$ unless ∂D is a curve (c_ϱ).

COROLLARY. *For the curve* (c_ϱ) *the equality*

$$\lambda_1^{(0)} = \left(\frac{j_0}{\varrho}\right)^{p+1} = \frac{P^{\frac{p+1}{2}} j_0^{p+1}}{A^{\frac{p+1}{2}}}$$

holds (see [3]) *where* j_0 *is the first positive zero of the generalized Bessel function of the first kind and* P *is defined in* (1.13). *Hence we get for the first eigenvalue the following lower bound:*

$$\lambda_1 \geqq \frac{P^{\frac{p+1}{2}} j_0^{p+1}}{A^{\frac{p+1}{2}}}.$$

REFERENCES

[1] ADAMS, A. R., *Sobolev spaces*, Pure and Applied Mathematics, Vol. 65, Academic Press, New York, 1975. *MR* **56** #9247

[2] BEESACK, P. R., Extensions of Wirtinger's inequality, *Trans. Roy. Soc. Canada* (3) **53** (1959), 21–30. *Zbl* **93**, 60

[3] BOGNÁR, G., The eigenvalue problem of a quasilinear partial differential equation, *NME Közl.*, Miskolc (to appear).

[4] BUSEMANN, H., The isoperimetric problem in the Minkowski plane, *Amer. J. Math.* **69** (1947), 863–871.

[5] COURANT, R. and HILBERT, D., *Methods of mathematical physics*, Vols. I, II, John Wiley, New York, 1953. *MR* **16**-426

[6] FABER, G., Beweis, dass unter allen homogenen Membranen von gleicher Fläche und gleicher Spannung die kreisförmige den tiefsten Grundton gibt, *Sitzungsberichte math.-phys. Kl. Bayerischen Akad. Wiss. München* (1923), 169–172. *Jb. Fortschritte Math.* **49**, 342

[7] GELFAND, I. M. and FOMIN, S. V., *Variational calculus*, Gosudarstv. Izdat. Fiz.-Mat. Lit., Moscow, 1961 (in Russian). *MR* **28** #3352

[8] GILBARG, D. and TRUDINGER, N. S., *Elliptic partial differential equations of second order*, 2nd ed., Grundlehren der math. Wissenschaften, **224**, Springer-Verlag, Berlin–Heidelberg–New York–Tokyo, 1983. *MR* **86c**:35035

[9] GUGGENHEIMER, H., On plane Minkowski geometry, *Geom. Dedicata* **12** (1982), 371–381. *MR* **85a**:52007

[10] HARDY, G. H., LITTLEWOOD, J. E. and PÓLYA, G., *Inequalities*, 2nd ed., Cambridge University Press, Cambridge, 1952. *MR* **13**-727

[11] KAWOHL, B., *Rearrangements and convexity of level sets in PDE*, Lecture Notes in Mathematics, Vol. 1150, Springer-Verlag, Berlin–Heidelberg–New York, 1985. *MR* **87a**:35001

[12] KRAHN, E., Über eine von Rayleigh formulierte Minimaleigenschaft des Kreises, *Math. Ann.* **94** (1924), 97–100.

[13] LEVINSON, N., Positive eigenfunctions for $\Delta u+\lambda f(u)=0$, *Arch. Rational Mech. Anal.* **11** (1962), 258–272. *MR* **26** #2751

[14] ÔTANI, M., Existence and nonexistence of nontrivial solutions of some nonlinear degenerate elliptic equations, *J. Funct. Anal.* **76** (1988), 140–159. *MR* **89c**: 35067

[15] PÓLYA, G. and SZEGŐ, G., *Isoperimetric inequalities in mathematical physics*, Princeton Univ. Press, Princeton, 1951.

[16] RAYLEIGH, I. W. S., *The theory of sound*, 2nd ed., London, 1896.

[17] RUND, H., *The differential geometry of Finsler spaces*, Die Grundlehren der mathematischen Wissenschaften, Bd. 101, Springer-Verlag, Berlin–Göttingen–Heidelberg, 1959. *MR* **21** #4462

[18] SCHWARZ, B., Bounds for the principal frequency of a nonhomogeneous membrane and for the generalized Dirichlet integral, *Pacific J. Math.* **7** (1957), 1653–1676. *MR* **19** #1180

[19] SZABÓ, I., *Höhere technische Mechanik*, Springer-Verlag, Berlin–Heidelberg–New York, 1972.

[20] SZEGŐ, G., Über eine Verallgemeinerung des Dirichletschen Integrals, *Math. Z.* **52** (1950), 676–685. *MR* **12** #703

[21] THELIN, F. DE, Sur l'espace propre associé à la première valeur propre du pseudolaplacien, *C. R. Acad. Sci. Paris Sér. I* **303** (1986), 355–358. *MR* **87i**:35147

[22] TYIHONOV, A. N. and SZAMARSZKIJ, A. A., *A matematikai fizika differenciálegyenletei*, Akadémiai Kiadó, Budapest, 1956 (in Hungarian).

(Received December 18, 1990)

MISKOLCI EGYETEM
MATEMATIKAI INTÉZETE
EGYETEMVÁROS
H-3515 MISKOLC
HUNGARY

Typeset by TypoTEX Ltd., Budapest
PRINTED IN HUNGARY
Akadémiai Kiadó és Nyomda Vállalat, Budapest

RECENTLY ACCEPTED PAPERS

TEMESVÁRI, Á. H., Die dünnste 8-fache gitterförmige Kreisüberdeckung der Ebene

DAO, Q. T., A minimal condition for stochastic approximation

BÖRÖCZKY, K. and SOLTAN, V., Translational and homothetic clouds for a convex body

HURELBAATAR, G., On the almost sure central limit theorem for Φ-mixing random variables

DEÁK, J., Extending a family of Cauchy structures in a limit space I–II

CONTENTS

HU ISSN 0081-6906

Studia Scientiarum Mathematicarum Hungarica

VOLUME 29
NUMBERS 3–4
1994

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. It 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-1117 Budapest, Prielle Kornélia u. 19-35

Manuscripts and editorial correspondence should be addressed to

J. Merza
Managing Editor

P.O. Box 127
H-1364 Budapest

Tel.: (36)(1) 118-2875 Fax: (36)(1) 117-7166
e-mail: h3299mer @ ella.hu

Subscription information

Subscription price for Volume 29 (1994) in 4 issues: \$ 92.00 including normal postage, airmail delivery plus \$ 20.00. For 1995 volumes 30-31 are scheduled for publication. The subscription price is \$ 95.00 each, air delivery plus \$ 20.00 per volume.

Orders should be addressed to

AKADÉMIAI KIADÓ
P.O.Box 245
H-1519 Budapest

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 233–239

AN EQUICONVERGENCE THEOREM

M. B. TAHIR

The equiconvergence theorems play (as is well known) an important role in the theory of expansions. Recently a general equiconvergence theorem was published in [1] by Horváth, Joó and Komornik for the one-dimensional Schrödinger operator without any restriction on the distribution of the eigenvalues on the complex plane, generalizing some known classical results of the field. The proof uses some estimates of [2] given by Joó.

The aim of the present paper is to extend the result of [1] for the operator $Lu := u^{(4)}$ using the results of [3–7].

Let G be an arbitrary open interval on the real line. Let $(u_\alpha) \subset L^2(G)$ be a complete and minimal system of eigenfunctions of the operator L. As usual, a function $u \not\equiv 0$ is called an eigenfunction of the operator L with some eigenvalue $\lambda \in \mathbb{C}$ if $u \in H^4_{\text{loc}}(G)$ and

$$Lu = \lambda u \quad \text{a.e. on} \quad G.$$

Choose the 4th roots μ_i of λ such that

$$\operatorname{Re}\mu_1 \geqq \operatorname{Re}\mu_2 \geqq 0 \geqq \operatorname{Re}\mu_3 \geqq \operatorname{Re}\mu_4$$

and denote $\mu := \mu_2$, $\varrho := \operatorname{Re}\mu \geqq 0$, $\nu := |\operatorname{Im}\mu| = \operatorname{Re}\mu_1$. We know (see Komornik [4], Joó [5]) that

$$\begin{vmatrix} u(x) & u(x-t)+u(x+t) & u(x-2R)+u(x+2R) \\ 1 & 2\cos\mu t & 2\cos 2\mu R \\ 1 & 2\operatorname{ch}\mu t & 2\operatorname{ch} 2\mu R \end{vmatrix} = 0.$$

Indeed, expanding it according to the first row we get

$$4u(x)\,(\cos\mu t\operatorname{ch} 2\mu R - \operatorname{ch}\mu t\cos 2\mu R) +$$

$$+[u(x-t)+u(x+t)]2\,(\cos 2\mu R - \operatorname{ch} 2\mu R) +$$

$$+[u(x-2R)+u(x+2R)]2\,(\operatorname{ch}\mu t - \cos\mu t) = 0,$$

1980 *Mathematics Subject Classification* (1985 Revision). Primary 34B25; Secondary 33A99.

Key words and phrases. Eigenfunction, eigenfunction expansion, Riesz basis.

and hence, dividing by $e^{2\operatorname{Re}\mu_1 R} = e^{2\nu R}$ we obtain the following generalization of the Titchmarsh formula:

$$(1)\qquad \begin{aligned}&[u(x+t)+u(x-t)-2\operatorname{ch}\mu t\cdot u(x)]f_1(\mu,R)=\\ &=4u(x)f_0(\mu,R,t)+[u(x+2R)+u(x-2R)]f_2(\mu,R,t),\end{aligned}$$

where

$$f_1(\mu,R):=\frac{2(\operatorname{ch}2\mu R-\cos 2\mu R)}{e^{2\operatorname{Re}\mu_1 R}},$$

$$f_0(\mu,R,t):=\frac{\operatorname{ch}2\mu R(\cos\mu t-\operatorname{ch}\mu t)}{e^{2\operatorname{Re}\mu_1 R}},$$

$$f_2(\mu,R,t):=\frac{2(\operatorname{ch}\mu t-\cos\mu t)}{e^{2\operatorname{Re}\mu_1 R}}.$$

For $0\leqq t$ we have

$$|\cos\mu t-\operatorname{ch}\mu t|\leqq\begin{cases} c\,(|\mu|t)^2\leqq c|\mu|t & \text{if } t\leqq\dfrac{1}{|\mu|},\\ ce^{\operatorname{Re}\mu_1 t}\leqq c|\mu|te^{\operatorname{Re}\mu_1 t} & \text{if } t\geqq\dfrac{1}{|\mu|},\end{cases}$$

i.e.

$$(2)\qquad |\cos\mu t-\operatorname{ch}\mu t|\leqq c|\mu|te^{\operatorname{Re}\mu_1 t}\quad\text{if}\quad t\geqq 0,$$

and hence we obtain the following estimates of Komornik

$$(3)\qquad \left|\frac{f_0(\mu,R,t)}{t}\right|\leqq c|\mu|e^{\operatorname{Re}(2\mu R+\mu_1(t-2R))},$$

$$(4)\qquad \left|\frac{f_2(\mu,R,t)}{t}\right|\leqq c|\mu|e^{\operatorname{Re}\mu_1(t-2R)}.$$

We need the following lower estimate for the integral of f_1: for any $R_0>0$ and $R\in(R_0/2,R_0)$ there exists $A(R_0)>0$ and $\delta(R_0)>0$ such that

$$(5)\qquad \int_{\frac{3}{4}R}^{R}|f_1(\mu,r)|\,dr\geqq\delta>0\quad\text{if}\quad|\mu|\geqq A.$$

To prove (5) it is enough to use the identity

$$\frac{1}{2}f_1(\mu,r)=\frac{\operatorname{ch}2\mu r-\cos 2\mu r}{e^{2\nu r}}=$$

$$= \frac{e^{2(\varrho \pm i\nu)r} + e^{-2(\varrho \pm i\nu)r} - e^{2i(\varrho \pm i\nu)r} - e^{-2i(\varrho \pm i\nu)r}}{2e^{2\nu r}},$$

then we have

$$\left| \frac{\operatorname{ch} 2\mu r - \cos 2\mu r}{e^{2\nu r}} - \left(e^{2(\varrho - \nu \pm i\nu)r} - e^{\mp 2i\varrho r} \right) \right| \leqq \frac{2}{e^{2\nu r}}.$$

The difference of the angles in the exponents is $\pm 2i(\nu + \varrho)r$. If $\nu + \varrho$ or which is the same $|\mu|$ is large enough comparing with R, then the difference of the exponents is at least $\frac{1}{2}$ on a segment of length $R/4$ and the estimate (5) follows.

Now we are in the position to estimate the expansion of w below with respect to the system (u_α). Let, as usual, $\nu > 0$ and write instead of ν, ϱ, μ in the following ν_α, ϱ_α, μ_α, respectively, let further

$$w_R(t) := \begin{cases} \frac{\sin \nu(x-t)}{\pi(x-t)} & \text{if } |x-t| \leqq R, \\ 0 & \text{if } |x-t| > R, \end{cases}$$

$$w := D_{R_0}(w_R) := \frac{2}{R_0} \int\limits_{R_0/2}^{R_0} w_R dR.$$

We know [1]:

$$\langle u_\alpha, w \rangle - \delta(\nu, \nu_\alpha) u_\alpha(x) =$$

$$= D_{R_0}\left(\int\limits_0^R \frac{\sin \nu t}{\pi t} [u_\alpha(x+t) + u_\alpha(x-t)] dt - \delta(\nu, \nu_\alpha) u_\alpha(x) \right) =$$

$$= D_{R_0}\left(\int\limits_0^R \frac{\sin \nu t}{\pi t} 2 \operatorname{ch} \mu_\alpha t dt - \delta(\nu, \nu_\alpha) \right) u_\alpha(x) +$$

$$+ D_{R_0}\left(\int\limits_0^R \frac{\sin \nu t}{\pi t} [u_\alpha(x+t) + u_\alpha(x-t) - 2 \operatorname{ch} \mu_\alpha t \, u_\alpha(x)] dt \right) =$$

$$=: D u_\alpha(x) + B.$$

In (1) write r in place of R, take the absolute value and integrate in r

from $3R/4$ to R. Taking into account (3–5) we obtain for $0 \leqq t \leqq R$

(6)
$$\begin{aligned}
&\left|\frac{u_\alpha(x+t)+u_\alpha(x-t)-2\operatorname{ch}\mu_\alpha t\, u_\alpha(x)}{t}\right| \leqq \\
&\leqq c|\mu_\alpha| \int_{\frac{3}{4}R}^{R} e^{\operatorname{Re}(\mu_1 t+2r(\mu-\mu_1))} dr |u_\alpha(x)| + \\
&+c\|u_\alpha\|_{L^\infty(x-2R,x+2R)}|\mu_\alpha| \int_{\frac{3}{4}R}^{R} e^{\operatorname{Re}\mu_1(t-2r)} dr \leqq \\
&\leqq c|\mu_\alpha| e^{-0.5\operatorname{Re}\mu_1 R} e^{\frac{3}{2}\operatorname{Re}\mu R}|u_\alpha(x)| + \\
&+c|\mu_\alpha| e^{-0.5\operatorname{Re}\mu_1 R}\|u_\alpha\|_{L^\infty(x-2R,x+2R)} \quad (|\mu_\alpha|>A).
\end{aligned}$$

Let $R_0/2 < R < R_0$, then according to (6) we have

(7)
$$|B| \leqq c|\mu_\alpha| e^{-\nu_\alpha R_0/4}\|u_\alpha\|_{L^\infty(K_{2R_0})} e^{\frac{3}{2}\varrho_\alpha R_0}, \quad \alpha \in K$$

further (Joó [2], Lemma 3.2):

(8)
$$|D| \leqq c\frac{(1+\varrho_\alpha)^2 e^{\varrho_\alpha R_0}}{1+(\nu-\nu_\alpha)^2} \leqq c\frac{e^{\frac{3}{2}\varrho_\alpha R_0}}{1+(\nu-\nu_\alpha)^2}.$$

We know (Komornik [7], (3)) that for any compact intervals K, K' with $K \subset K' \subset G$, $K \subset \operatorname{int} K'$ there exists $\varepsilon_0 > 0$ such that

(9)
$$\|u_\alpha\|_{L^\infty(K)} e^{\varepsilon_0 \varrho_\alpha} \leqq \frac{1}{\varepsilon_0}\|u_\alpha\|_{L^2(K')}.$$

Now choose $R_0 > 0$ such that $3R_0 < \varepsilon_0$. Then, according to our estimates, we have for $|\mu_\alpha| > A$:

(10)
$$\begin{aligned}
&|\langle u_\alpha, w\rangle - \delta(\nu,\nu_\alpha)u_\alpha(x)| \leqq \\
&\leqq c\left(\frac{1}{1+(\nu-\nu_\alpha)^2} + |\mu_\alpha| e^{-\nu_\alpha R_0/4}\right)\|u_\alpha\|_{L^\infty(K_{2R_0})} e^{\frac{3}{2}\varrho_\alpha R_0} \leqq \\
&\leqq c\left(\frac{1}{1+(\nu-\nu_\alpha)^2} + |\mu_\alpha| e^{-\nu_\alpha R_0/4}\right)\|u_\alpha\|_{L^2(K')}.
\end{aligned}$$

Indeed, apply (9) for a bigger $\hat{K}$ in place of K and apply (9) for $\hat{K} \subset \subset \operatorname{int} K'$ and assume also $2R_0 < \operatorname{dist}(K, \partial\hat{K})$, i.e. $K_{2R_0} \subset \hat{K}$, namely in this case

$$\|u_\alpha\|_{L^\infty(K_{2R_0})} \leqq \|u_\alpha\|_{L^\infty(\acute{K})} \leqq ce^{-3R_0\varrho_\alpha}\|u_\alpha\|_{L^2(K')}.$$

Now we prove the estimate analogous to (10) for $|\mu_\alpha| \leqq A$. We know that

$$\langle u_\alpha, w\rangle = D_{R_0}\left(\int_0^R [u_\alpha(x+t) + u_\alpha(x-t)]\frac{\sin\nu t}{\pi t}dt\right) =$$

$$= D_{R_0}\left([u_\alpha(x+R) + u_\alpha(x-R)]\int_0^R \frac{\sin\nu t}{\pi t}dt\right) -$$

$$-D_{R_0}\left(\int_0^R [u'_\alpha(x+t) - u'_\alpha(x-t)]\left(\int_0^t \frac{\sin\nu\tau}{\pi\tau}d\tau\right)dt\right).$$

According to

$$\left|\int_0^R \frac{\sin\nu t}{\pi t}dt\right| = \left|\int_0^{\nu R} \frac{\sin u}{\pi u}du\right| \leqq c$$

we have

$$|\langle u_\alpha, w\rangle| \leqq c\left(\|u_\alpha\|_{L^\infty(K_{R_0})} + \|u'_\alpha\|_{L^\infty(K_{R_0})}\right).$$

On the other hand V. Komornik proved [6, Theorem 2] that

$$\|u'_\alpha\|_{L^\infty(K_{R_0})} \leqq c(1+|\mu_\alpha|)\|u_\alpha\|_{L^\infty(K_{R_0})},$$

consequently, we have

$$(11)\qquad |\langle u_\alpha, w\rangle| \leqq c\|u_\alpha\|_{L^\infty(K_{R_0})} \text{ if } |\mu_\alpha| \leqq A.$$

Now we are in the position to prove the following

THEOREM. *Let (u_α) be a complete and minimal system in $L^2(G)$ of eigenfunctions of the operator $Lu = u^{(4)}$ with arbitrary complex eigenvalues (λ_α), denote (v_α) the system biorthogonal to (u_α). Suppose*

$$(12)\qquad \sup_{t>0}\sum_{t\leqq\nu_\alpha\leqq t+1} 1 < \infty.$$

Denote

$$S_\nu(f,x) := \int_{x-R}^{x+R}\frac{\sin\nu(y-x)}{\pi(y-x)}f(y)dy, \qquad \sigma_\nu(f,x) := \sum_{\alpha<\nu}\langle f, v_\alpha\rangle\, u_\alpha(x).$$

Then the following statements are equivalent:

(a) *For any compact interval* $K \subset G$

$$\sup_{\alpha} \|v_\alpha\|_{L^2(G)} \|u_\alpha\|_{L^2(K)} < \infty.$$

(b) *For any compact interval* $K \subset G$ *and every* $f \in L^2(G)$

$$\lim_{\nu\to\infty} \sup_{x\in K} |S_\nu(f,x) - \sigma_\nu(f,x)| = 0.$$

(c) *For any compact interval* $K \subset G$ *and any* $f \in L^2(G)$

$$\lim_{\nu\to\infty} \|f - \sigma_\nu(f)\|_{L^2(K)} = 0.$$

PROOF. (a) $\Rightarrow$ (b). According to (10) and (11) we have

$$\sum |\langle u_\alpha, w\rangle - \delta(\nu,\nu_\alpha)u_\alpha(x)| \, \|v\|_{L^2(G)} \leqq$$

$$\leqq c \sum \frac{|\langle u_\alpha, w\rangle - \delta(\nu,\nu_\alpha)u_\alpha(x)|}{\|u_\alpha\|_{L^2(K')}} \leqq$$

$$\leqq c \sum_{\alpha=1}^{\infty} \left(\frac{1}{1+(\nu-\nu_\alpha)^2} + \frac{|\mu_\alpha|}{e^{\nu_\alpha R_0/4}} \right)$$

which, according to (12), can be estimated by a constant independent of ν, and hence for any fixed x and ν the series

$$F(x,y) := \sum [\langle u_\alpha, w\rangle - \delta(\nu,\nu_\alpha)u_\alpha(x)] v_\alpha(y)$$

is absolutely convergent in L^2_y for every fixed x, further

$$\int_G u_\alpha(y) F(x,y)\, dy = \langle u_\alpha, w\rangle - \delta(\nu,\nu_\alpha)u_\alpha(x)$$

and hence

$$F(x,y) = w - \sum_{\nu_\alpha<\nu}{}_{\alpha} u_\alpha(x)\overline{v}_\alpha(y) - \frac{1}{2}\sum_{\nu_\alpha=\nu}{}_{\alpha} u_\alpha(x)\overline{v}_\alpha(y),$$

i.e.

$$\sup_{x,\nu} \left\| w - \sum_{\nu_\alpha<\nu}{}_{\alpha} u_\alpha(x)\overline{v}_\alpha(y) - \frac{1}{2}\sum_{\nu_\alpha=\nu}{}_{\alpha} u_\alpha(x)\overline{v}_\alpha(y) \right\|_{L^2(G)} < \alpha.$$

Taking into account (11) we have

$$\left\| \sum_{\nu_\alpha=\nu} u_\alpha(x)\overline{v}_\alpha(y) \right\|_{L^2(G)} \leqq \sum_{\nu_\alpha=\nu} |u_\alpha(x)|\, \|\overline{v}_\alpha(y)\|_{L^2(G)} \leqq$$

$$\leqq c \sum_{\nu_\alpha=\nu} \|u_\alpha\|_{L^2(K')} \|v_\alpha\|_{L^2(G)} \leqq c$$

(we used also (12)). The other parts of the proof of (a) $\Rightarrow$ (b) and (b) $\Rightarrow$ (c) $\Rightarrow$ (a) are similar as in [1] hence we omit the details. $\square$

REFERENCES

[1] Horváth, M., Joó, I. and Komornik, V., An equiconvergence theorem, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **31** (1988), 19–26. *MR* **90j**:34039

[2] Joó, I., On the divergence of eigenfunction expansions, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **32** (1989), 3–36. *MR* **92d**:47030

[3] Joó, I. and Komornik, V., On the equiconvergence of expansions by Riesz bases formed by eigenfunctions of the Schrödinger operator, *Acta Sci. Math. (Szeged)* **46** (1983), 357–375. *MR* **85f**:35155

[4] Komornik, V., Local upper estimates for the eigenfunctions of a linear differential operator, *Acta Sci. Math. (Szeged)* **48** (1985), 243–256. *MR* **87j**:34046

[5] Joó, I., Remarks to a paper: „Upper estimates for the eigenfunctions of higher order of a linear differential operator" [Acta Sci. Math. (Szeged) **45** (1983), no. 1–4, 261–271; *MR* **84m**:34030] of V. Komornik, *Acta Sci. Math. (Szeged)* **47** (1984), 201–204. *MR* **85m**:34041

[6] Komornik, V., Upper estimates for the eigenfunctions of higher order of a linear differential operator, *Acta Sci. Math. (Szeged)* **45** (1983), 261–271. *MR* **84m**: 34030

[7] Komornik, V., On the equiconvergence of eigenfunction expansions associated with ordinary linear differential operators, *Acta Math. Hungar.* **47** (1986), 261–280. *MR* **87g**:34028

(Received January 2, 1991)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR
ALKALMAZOTT ANALÍZIS TANSZÉK
MÚZEUM KRT. 6–8
H–1088 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 241–244

ON A CONJECTURE OF ERDŐS ON BINOMIAL COEFFICIENTS

A. GRYTCZUK

1. Introduction

M. Wunderlich [4] attributes the following conjecture to P. Erdős: The equation

$$2\binom{x+n-1}{n}=\binom{y+n-1}{n} \tag{1.1}$$

where $n>1$, has only one solution in positive integers:

$$x=n, \qquad y=n+1. \tag{1.2}$$

A. Mąkowski [2] remarked that infinitely many counterexamples exist to this conjecture and he conjectured that it is probably true when we require

$$y-x\geqq 3. \tag{1.3}$$

The purpose of this paper is to give some new information about those conjectures. Namely we prove the following theorems:

THEOREM 1. *If $y-x=1$ then the equation* (1.1) *has only one solution in positive integers: $x=n$, $y=n+1$.*

THEOREM 2. *If $y-x=2$ then the equation* (1.1) *has infinitely many solutions in positive integers x,y,n and every such solution can be obtained from the solutions of the Pell's equation*

$$A^2-2B^2=-1$$

where $B=2u+3$, $x=u+1$, $y=x+2$, $n=\dfrac{A-B}{2}$.

THEOREM 3. *Let $n\geqq 3$ and $y-x=k\geqq 3$. Then the equation* (1.1) *has no solutions in positive integers x,y,n such that $x\leqq n+1$.*

THEOREM 4. *Let $n\geqq 3$ and $y-x=k\geqq 3$, where k is fixed integer and let*

$$F(n,x)=2\binom{x+n-1}{n}-\binom{x+k+n-1}{n} \tag{1.4}$$

1980 *Mathematics Subject Classification* (1985 Revision). Primary 10B15; Secondary 10B99.

Key words and phrases. Diophantine equations, binomial coefficients, Erdős' problem.

be an irreducible polynomial such that the leading form of $F(n,x)$ is not a constant multiple of a power of an irreducible polynomial. Then the equation

$$F(n,x)=0 \tag{1.5}$$

has only finite number of solutions in positive integers n,x such that $n<x-1$ and all such solutions satisfy

$$x \leqq (8k(k-1)k!)^{k^{2k^3}}. \tag{1.6}$$

2. Proof of Theorem 1

Let $x-1=u$ and $y-1=v$ then by the well-known formula on binomial coefficients it follows that equation (1.1) is equivalent to

$$2(u+1)(u+2)\dots(u+n)=(v+1)(v+2)\dots(v+n). \tag{2.1}$$

From (2.1) it follows that if $u \geqq v$ then (2.1) is impossible, therefore we have $v>u$ and $v-u \geqq 1$. If $y-x=1$ then $v-u=1$ and by (2.1) we get

$$2(u+1)=u+n+1. \tag{2.2}$$

From (2.2) we get $u=n-1$ and so $v=u+1=n$, $x=u+1=n$ and $y= =v+1=n+1$ which prove the theorem.

3. Proof of Theorem 2

Since $y-x=2$ then putting $x-1=u$, $y-1=v$ we get $v-u=2$ and by (2.1) we obtain

$$n^2+(2u+3)n-(u+1)(u+2)=0. \tag{3.1}$$

From (3.1) we have

$$(2u+3)^2+4(u+1)(u+2)=2(2u+3)^2-1=(2n+2u+3)^2 \tag{3.2}$$

and we have

$$A^2-2B^2=-1, \tag{3.3}$$

where $B=2u+3$ and $A=2n+2u+3$. From these the theorem follows.

We note that by a well-known result concerning Pell's equation we get $A_1=B_1=1$ and all solutions of (3.3) are given by the formula

$$A_{2m+1}+B_{2m+1}\sqrt{2}=(1+\sqrt{2})^{2m+1}. \tag{3.4}$$

4. Proof of Theorem 3

Let $y-x=k \geqq 3$ and $n \geqq 3$. Then putting $x-1=u$ and $y-1=v$ we get $v-u=k \geqq 3$ and by (2.1) it follows that

$$(4.1)\qquad 2(u+1)(u+2)\dots(u+k) = (u+1+n)(u+2+n)\dots(u+k+n).$$

From (4.1) we have

$$(4.2)\qquad 2 = \left(1+\frac{n}{u+1}\right)\left(1+\frac{n}{u+2}\right)\dots\left(1+\frac{n}{u+k}\right).$$

It is obvious that

$$(4.3)\qquad \frac{n}{u+1} > \frac{n}{u+2} > \dots > \frac{n}{u+k}.$$

From (4.2) and (4.3) we obtain

$$(4.4)\qquad 2 > \left(1+\frac{n}{u+k}\right)^k.$$

From (4.1) and the assumption of Theorem 3 we have $n \geqq u \geqq 3$ and since $k \geqq 3$ thus we get $\frac{n}{u+k} \geqq \frac{1}{k}$ and by (4.4) it follows that

$$2 < \left(1+\frac{1}{k}\right)^k \leqq \left(1+\frac{n}{u+k}\right)^k < 2,$$

which is impossible and the proof is complete.

5. Proof of Theorem 4

Let $x = u+1$, $y = v+1$, $n \geqq 3$ and $k \geqq 3$ be fixed integers. If $y - x = k$ then by (4.1) it follows that

$$(5.1)\qquad 2(u+1)(u+2)\dots(u+k) = n^k + A_1 n^{k-1} + \dots + A_k,$$

where

$$(5.2)\qquad \begin{aligned} A_1 &= (u+1)+(u+2)+\dots+(u+k) \\ &\vdots \\ A_k &= (u+1)(u+2)\dots(u+k). \end{aligned}$$

Let $\|F\|$ denote the height of the polynomial $F(n,u)$ given by (1.4) and (5.1) then by (5.1) and (5.2) we get

$$(5.3)\qquad \|F\| \leqq (k-1)k!.$$

Now by Runge's theorem [3] and some effective version of this theorem given by D. L. Hilliker and E. G. Straus (see [1], Theorem 4.9) we get that the equation $F(n,u)=0$ has only finite number of solutions in positive integers $n<u$ and all such solutions satisfy

$$u<(8k\|F\|)^{k^{2k^3}}. \tag{5.4}$$

From (5.4), (5.3) and by virtue of $u=x-1$ we obtain (1.6) and the proof is complete.

6. Remarks

First we remark that, by (3.4), all solutions of $A^2-2B^2=-1$ are given by the formulas

$$A_{2m+1}=3A_{2m-1}+4B_{2m-1}, \qquad B_{2m+1}=2A_{2m-1}+3B_{2m-1}, \tag{6.1}$$

where $A_1=B_1=1$ and $m=1,2,\dots$.

Formulas (6.1) are useful for the computation of solutions of (1.1) in the case $y-x=2$. From (6.1) we get

$$\begin{aligned} &\langle A_3,B_3\rangle=\langle 7,5\rangle, && \langle A_5,B_5\rangle=\langle 41,29\rangle,\\ &\langle A_7,B_7\rangle=\langle 239,169\rangle, && \langle A_9,B_9\rangle=\langle 1393,985\rangle, \quad \dots. \end{aligned} \tag{6.2}$$

In the case $\langle A_5,B_5\rangle=\langle 41,29\rangle$, by Theorem 2, we obtain $n=\frac{A-B}{2}=$ $=\frac{41-29}{2}=6$, $u=13$, $v=15$ and therefore we get $\langle x,y,n\rangle=\langle 14,16,6\rangle$.

This example was given by A. Mąkowski in [2].

NOTE added in proof. By using recent results given by A. Grytczuk and A. Schinzel ,,On Runge's theorem about Diophantine equations", Colloq. Math. Soc. J. Bolyai **60** (1992), 329–356 or by P. G. Walsh ,,A quantitative version of Runge's theorem on Diophantine equations", Acta Arith. LXII (1992), no. 2, 157–172 we can improve the estimate of Theorem 4 to $x\leqq$ $\leqq((4k^3)^{8k^2}(k-1)k!)^{96k^{11}}$ or $x\leqq(2k)^{18k^7}((k-1)k!)^{12k^6}$, respectively.

REFERENCES

[1] HILLIKER, D. L. and STRAUS, E. G., Determination of bounds for the solutions to those binary Diophantine equations that satisfy the hypotheses of Runge's theorem, *Trans. Amer. Math. Soc.* **280** (1983), 637–657. *MR* **85c**:11031

[2] MĄKOWSKI, A., Remark on a conjecture of Erdős on binomial coefficients, *Math. Comp.* **24** (1970), 705. *MR* **43** #6157

[3] RUNGE, C., Über ganzzahlige Lösungen von Gleichungen zwischen zwei Veränderlichen, *J. Reine Angew. Math.* **100** (1887), 425–435. *Jb. Fortschritte Math.* **19**, 76

[4] WUNDERLICH, M., Certain properties of pyramidal and figurate numbers, *Math. Comp.* **16** (1962), 482–486. *MR* **26** #6115

(Received February 2, 1991)

WYŻSZA SZKOŁA PEDAGOGICZNA
INSTYTUT MATEMATYKI
PL. SŁOWIANSKI 9
PL–65-069 ZIELONA GÓRA
POLAND

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 245–258

STRONGLY NONLINEAR ELLIPTIC SYSTEMS WITH NONLOCAL BOUNDARY CONDITIONS

I. M. HASSAN

In this paper the existence of the weak solution $u=(u_1,\dots,u_M)$ of the system

(0.1)
$$-\sum_{k=1}^{n}\partial_k\left[f_j^k(x,u,\partial_1u,\dots,\partial_nu)\right]+f_j^0(x,u,\partial_1u,\dots,\partial_nu)+g_j(x,u)=F_j$$

in Ω, $j=1,2,\dots,M$ with nonlinear and nonlocal boundary condition

(0.2) $$\partial^j_{\nu^*}u=h_{1,j}(x,u(x))+h_{2,j}(x,u(\Phi(x)))\qquad \text{on } \partial\Omega,\ j=1,\dots,M$$

will be proved, where $\Omega\subset\mathbf{R}^n$ is any domain with continuously differentiable and bounded boundary $\partial\Omega$,

$$\partial^j_{\nu^*}u:=\sum_{k=1}^{n}\left[f_j^k(x,u,\partial_1u,\dots,\partial_nu)\right]\nu_k,\qquad j=1,\dots,M,$$

and ν_k denote the coordinates of the normal unit vector on $\partial\Omega$. Φ is C^1-diffeomorphism in a neighbourhood of $\partial\Omega$ such that

$$S:=\Phi(\partial\Omega)\subset\bar{\Omega}.$$

$f_j^k(x,u,\partial_1u,\dots,\partial_nu)$, $h_{2,j}(x,u)$ have certain polynomial growth in u,∂_1u, $\dots$, ∂_nu but in the terms $g_j(x,u)$ and $h_{1,j}(x,u)$ no growth restriction is imposed with respect to u, but it is supposed that $g_j,h_{1,j}$ satisfy the sign conditions

$$g_j(x,\eta)\eta_j\geqq 0,\qquad h_{1,j}(x,\eta)\eta_j\leqq 0.$$

The weak solution of (0.1), (0.2) will be defined as follows. Assume that u is a classical solution of (0.1), (0.2) such that $h_{1,j}(x,u)\in L^1(\partial\Omega)$, $g_j(x,u)\in$ $\in L^1(\Omega)$. Multiply (0.1) by v then by the Gauss–Ostrogradskij theorem and

1991 *Mathematics Subject Classification.* Primary 35J55; Secondary 35J65.
Key words and phrases. Strongly nonlinear equations, nonlocal boundary conditions.

0081–6906/94/$ 4.00

by using an integral transformation we obtain

$$\sum_{j=1}^{M}\Big\{\sum_{k=1}^{n}\int_{\Omega} f_j^k(x,u,\partial_1 u,\dots,\partial_n u)\partial_k v_j+\int_{\Omega} f_j^0(x,u,\partial_1 u,\dots,\partial_n u)v_j-$$

$$(0.3)\quad -\int_{\partial\Omega} h_{1,j}(x,u)v_j\,d\sigma-\int_{S}\tilde{h}_{2,j}(x,u)v_j\left(\Phi^{-1}(x)\right)\,d\sigma_x+\int_{\Omega} g_j(x,u)v_j\Big\}=$$

$$=\sum_{j=1}^{M}\int_{\Omega} F_j v_j$$

for any $v \in W_p^1(\Omega)\times\ldots\times W_p^1(\Omega)$ with compact support if $v_j \in L^\infty(\Omega)$ and $v_j|_{\partial\Omega} \in L^\infty(\partial\Omega)$. Thus the weak solution of (0.1), (0.2) will be defined by (0.3). Conversely, if u is a solution of (0.3) which is sufficiently smooth, then u is a solution of (0.1), (0.2).

Nonlocal linear boundary value problems have been considered e.g. in [12] and [14]. The importance of nonlocal boundary value problems in plasma-physics has been emphasized in [13]. Nonlocal and nonlinear boundary value problems have been studied in [4], [6] and [8]. Nonlinear elliptic systems have been considered in [10].

1. Existence theorem

Denote by $W_p^1(\Omega)$ the Sobolev space of real-valued functions u, whose distributional derivatives of order $\leqq 1$ belong to $L^p(\Omega)$, $1<p<\infty$. If we define the norm in $W_p^1(\Omega)$ by

$$\|u\|:=\Big\{\sum_{|\alpha|\leqq 1}\int_{\Omega}|\partial^\alpha u|^p\Big\}^{1/p},$$

then $X:=W_p^1(\Omega)\times\ldots\times W_p^1(\Omega)$ is a reflexive Banach space. Denote by X' the dual space of X. The points $\xi=(\xi_0,\xi_1,\dots,\xi_n)\in\mathbf{R}^{(n+1)M}$, $(\xi_j = (\xi_j^1,\dots,\xi_j^M)\in\mathbf{R}^M)$ will be written in the form $\xi=(\eta,\zeta)$ where $\eta=\xi_0$, $\zeta=(\xi_1,\dots,\xi_n)$.

Assume that

(a) The functions $f_j^k:\Omega\times\mathbf{R}^{(n+1)M}\to\mathbf{R}$ satisfy the Carathéodory conditions, i.e. they are measurable in x for every fixed $\xi\in\mathbf{R}^{(n+1)M}$ and continuous in ξ for a.e. x in Ω. Similarly, the functions $h_{1,j}:\partial\Omega\times\mathbf{R}^M\to\mathbf{R}$, $\tilde{h}_{2,j}:S\times\mathbf{R}^M\to\mathbf{R}$ also satisfy the Carathéodory conditions.

(b) There exist a constant $c_1 > 0$ and a function $K_1 \in L^q(\Omega)$, $\left(\frac{1}{p}+\frac{1}{q}\right)=1$ such that

$$|f_j^k(x,\xi)| \leqq c_1|\xi|^{p-1} + K_1(x), \qquad j=1,\dots,M, \quad k=0,1,\dots,n$$

for all $\xi \in \mathbf{R}^{(n+1)M}$ and a.e. x in Ω.

(c) For all (η,ζ), (η,ζ') in $\mathbf{R}^{(n+1)M}$ with $\zeta \neq \zeta'$

$$\sum_{j=1}^{M}\sum_{k=1}^{n}\left[f_j^k(x,\eta,\zeta) - f_j^k(x,\eta,\zeta')\right](\xi_j - \xi_j') > 0.$$

(d) There exist constant c_2 and a function $K_2 \in L^1(\Omega)$ such that

$$\sum_{j=1}^{M}\sum_{k=0}^{n} f_j^k(x,\xi)\xi_j^k \geqq c_2|\xi|^p - K_2(x).$$

(e) For any $s > 0$, there exists $g_{j,s} \in L^1(\Omega)$ such that for a.e. $x \in \Omega$

$$|g_j(x,\eta)| \leqq g_{j,s}(x) \qquad \text{if } |\eta| \leqq s,\ j=1,\dots,M.$$

(f) For any $\eta \in \mathbf{R}^M$, and for a.e. x in Ω

$$g_j(x,\eta)\eta_j \geqq 0, \qquad j=1,\dots,M.$$

(g) For any $s > 0$, there is a function $h_{1,j,s} \in L^1(\partial\Omega)$ such that for a.e. x in $\partial\Omega$

$$|h_{1,j}(x,\eta)| \leqq h_{1,j,s}(x) \qquad \text{if } |\eta| \leqq s,\ j=1,\dots,M.$$

(h) For any $\eta \in \mathbf{R}$ and for a.e. x in $\partial\Omega$ we have

$$h_{1,j}(x,\eta)\eta_j \leqq 0, \qquad j=1,\dots,M.$$

(i) There exist constant $c_3 > 0$ and a fixed function $K_3 \in L^{1+1/\rho}(S)$ such that

$$\left|\tilde{h}_{2,j}(x,\eta)\right| \leqq c_3|\eta|^\rho + K_3(x)$$

where $0 < \rho < p-1$.

The main result of this paper is the following

THEOREM. *Suppose that the assumptions* (a)–(i) *are satisfied. Then for any* $F_j \in \left(W_p^1(\Omega)\right)'$ *there exists* $u \in X$ *such that*

$$\begin{aligned} &g_j(x,u) \in L^1(\Omega), && g_j(x,u)u_j \in L^1(\Omega), \\ &h_{1,j}(x,u) \in L^1(\partial\Omega), && h_{1,j}(x,u)u_j \in L^1(\partial\Omega) \end{aligned} \tag{1.1}$$

and

(1.2)

$$\sum_{j=1}^{M}\left\{\sum_{k=1}^{n}\int_{\Omega} f_j^k(x,u,\partial_1 u,\dots,\partial_n u)\partial_k v_j+\int_{\Omega} f_j^0(x,u,\partial_1 u,\dots,\partial_n u)v_j-\right.$$

$$\left.-\int_{\partial\Omega} h_{1,j}(x,u)v_j\,d\sigma-\int_{S}\tilde h_{2,j}(x,u)v_j\left(\Phi^{-1}(x)\right)\,d\sigma_x+\int_{\Omega} g_j(x,u)v_j\,dx\right\}=$$

$$=\sum_{j=1}^{M}\langle F_j,v_j\rangle$$

for all $v\in X$ *with compact support satisfying* $v\in L^\infty(\Omega)\times\ldots\times L^\infty(\Omega)$, $v_{|\partial\Omega}\in$ $\in L^\infty(\partial\Omega)\times\ldots\times L^\infty(\partial\Omega)$, *and for* $v=u$.

For any $u,v\in X$ let

$$\langle T(u),v\rangle:=\sum_{j=1}^{M}\sum_{k=1}^{n}\int_{\Omega} f_j^k(x,u,\partial_1 u,\dots,\partial_n u)\partial_k v_j+$$

(1.3)

$$+\sum_{j=1}^{M}\int_{\Omega} f_j^0(x,u,\partial_1 u,\dots,\partial_n u)v_j-\sum_{j=1}^{M}\int_{S}\tilde h_{2,j}(x,u)v_j\left(\Phi^{-1}(x)\right)\,d\sigma_x$$

and for any $\mu>0$ let

$$g_j^\mu(x,\eta):=\begin{cases} g_j(x,\eta) & \text{if } |\eta|\leqq\mu,\ |x|\leqq\mu\\ \mu\dfrac{g_j(x,\eta)}{|g_j(x,\eta)|} & \text{if } |\eta|>\mu,\ |x|\leqq\mu\\ 0 & \text{if } |x|>\mu\end{cases}\tag{1.4}$$

$$h_{1,j}^\mu(x,\eta):=\begin{cases} h_{1,j}(x,\eta) & \text{if } |\eta|\leqq\mu,\ x\in\partial\Omega\\ \mu\dfrac{h_{1,j}(x,\eta)}{|h_{1,j}(x,\eta)|} & \text{if } |\eta|>\mu,\ x\in\partial\Omega\end{cases}\tag{1.5}$$

$(j=1,\dots,M)$.

Define S_μ by

$$\langle S_\mu(u),v\rangle:=\sum_{j=1}^{M}\int_{\Omega} g_j^\mu(x,u)v_j-\sum_{j=1}^{M}\int_{\partial\Omega} h_{1,j}^\mu(x,u)v_j\,d\sigma.\tag{1.6}$$

Firstly we shall prove several lemmas (similar lemmas have been proved in [7], [9] and [11]).

LEMMA 1. *Operator $T+S_\mu : X \to X'$ is pseudomonotone.*

PROOF. By (a), (b), (i) and (1.4), (1.5) the operator $T+S_\mu$ is bounded. From (1.3) T can be written in the form $T=A-B$, where

$$\langle A(u),v\rangle := \sum_{j=1}^{M}\sum_{k=1}^{n}\int_\Omega f_j^k(x,u,\partial_1 u,\dots,\partial_n u)\partial_k v_j + $$
$$+\sum_{j=1}^{M}\int_\Omega f_j^0(x,u,\partial_1 u,\dots,\partial_n u)v_j,$$
$$\langle B(u),v\rangle := \sum_{j=1}^{M}\int_S \tilde{h}_{2,j}(x,u)v_j\left(\Phi^{-1}(x)\right)\,d\sigma_x.$$

From conditions (b), (d) and Carathéodory condition (see (a)) it follows that A is pseudomonotone (see [10]). Let (u^l) be a sequence such that (u^l) converges weakly in X to u and

$$\limsup_{l\to\infty}\langle T(u^l),u^l-u\rangle \leqq 0.$$

Firstly we shall prove that

$$\lim_{l\to\infty}\langle B(u^l),u^l-u\rangle = 0. \tag{1.7}$$

By a compact imbedding theorem there exists a subsequence $(\tilde{u}_j^l)$ of (u_j^l) such that for all j, $\tilde{u}^l_{j|\partial\Omega}$ converges to u_j in $L^{\tilde{q}}(\partial\Omega)$ where $\tilde{q}:=\rho+1<p$. By using condition (i), and Hölder's inequality (with $\frac{1}{\tilde{p}}+\frac{1}{\tilde{q}}=1$) we have

$$\begin{aligned}\left|\langle B(\tilde{u}^l),\tilde{u}^l-u\rangle\right| &= \sum_{j=1}^{M}\left|\int_S \tilde{h}_{2,j}(x,\tilde{u}^l)(\tilde{u}_j^l-u_j)\left(\Phi^{-1}(x)\right)\,d\sigma_x\right| \leqq \\ &\leqq \sum_{j=1}^{M}\left\{\int_S\left|\tilde{h}_{2,j}(x,\tilde{u}^l)\right|^{\tilde{p}}d\sigma_x\right\}^{1/\tilde{p}}\cdot\left\{\int_S\left|(\tilde{u}_j^l-u_j)\left(\Phi^{-1}(x)\right)\right|^{\tilde{q}}\right\}^{1/\tilde{q}} \leqq\end{aligned} \tag{1.8}$$

$$\leqq \text{const} \sum_{j=1}^{M} \left\{ \int_S \left| \bar{h}_{2,j}(x, \tilde{u}^l) \right|^{\bar{p}} d\sigma \right\}^{1/\bar{p}} \cdot \left\{ \int_{\partial\Omega} \left| \tilde{u}^l_j - u_j \right|^{\bar{q}} \right\}^{1/\bar{q}} \leqq$$

$$\leqq \text{const} \left\{ \int_S \left| c_3 |\tilde{u}^l|^{\rho} + K_3 \right|^{\bar{p}} d\sigma \right\}^{1/\bar{p}} \cdot \sum_{j=1}^{M} \| \tilde{u}^l_j - u_j \|_{L^{\bar{q}}(\partial\Omega)} =$$

$$= \text{const} \left\{ \int_S \left[c_3 |\tilde{u}^l|^{\rho} + K_3 \right]^{\rho+1/\rho} d\sigma \right\}^{\rho/\rho+1} \cdot \sum_{j=1}^{M} \| \tilde{u}^l_j - u_j \|_{L^{\bar{q}}(\partial\Omega)} \leqq$$

$$\leqq \text{const} \left\{ \left[\int_S |\tilde{u}^l|^{\rho+1} d\sigma \right]^{\rho/\rho+1} + \| K_3 \|_{L^{1+1/\rho}(S)} \right\} \cdot \sum_{j=1}^{M} \| \tilde{u}^l_j - u_j \|_{L^{\bar{q}}(\partial\Omega)} \leqq$$

$$\leqq \text{const} \left\{ \| \tilde{u}^l \|_X^{\varrho} + c \right\} \sum_{j=1}^{M} \| \tilde{u}^l_j - u_j \|_{L^{\bar{q}}(\partial\Omega)}.$$

In the last product the first term is bounded and the second term tends to 0, consequently,

$$\lim_{l \to \infty} \langle B(\tilde{u}^l), \tilde{u}^l - u \rangle = 0.$$

It is not difficult to show that (1.7) is true also for the original sequence. Further, we shall prove that

$$(1.9) \qquad B(u^l) \overset{w'}{\to} B(u) \qquad \text{in } X', \text{ i.e.}$$

for all $v \in X$

$$\langle B(u^l), v \rangle \to \langle B(u), v \rangle.$$

We have seen that there exists a subsequence $(\tilde{u}^l)$ of (u^l) such that $\tilde{u}^l_{|\partial\Omega}$ converges to u in $L^{\bar{q}}(\partial\Omega)$ and converges a.e. to u on $\partial\Omega$. Thus

$$\bar{h}_{2,j}(x, \tilde{u}^l) \to \bar{h}_{2,j}(x, u), \qquad \text{a.e. on } \partial\Omega.$$

Now we shall use Vitali's convergence theorem. By Hölder's inequality and the boundedness of the trace operator, we have

$$\left| \int_E \bar{h}_{2,j}(x, \tilde{u}^l) v_j \left(\Phi^{-1}(x) \right) d\sigma_x \right| \leqq$$

$$\leqq \left\{ \int_E |\bar{h}_{2,j}(x, \tilde{u}^l)|^{\bar{p}} d\sigma \right\}^{1/\bar{p}} \left\{ \int_E |v_j \left(\Phi^{-1}(x) \right)|^{\bar{q}} d\sigma_x \right\}^{1/\bar{q}} \leqq$$

$$\leqq \left\{ \int_{\partial\Omega} |\bar{h}_{2,j}(x, \tilde{u}^l)|^{\bar{p}} d\sigma \right\}^{1/\bar{p}} \left\{ \int_E |v_j \left(\Phi^{-1}(x) \right)|^{\bar{q}} \right\}^{1/\bar{q}} < c\varepsilon$$

if the measure of E is sufficiently small since

$$\int_{\partial\Omega} |\tilde{h}_{2,j}(x, u^l(x))|^{\tilde{p}} \, d\sigma_x$$

is bounded (see (1.8)). So it is not difficult to show that all conditions of Vitali's theorem are satisfied and thus we obtain

$$\lim_{l\to\infty} \langle B(\tilde{u}^l), v\rangle = \langle B(u), v\rangle,$$

and it is easy to show that the above equality is true also for the original sequence, i.e. we have (1.9).

Thus we have shown that if (u^l) converges weakly to u in X and

$$\limsup \langle T(u^l), u^l - u\rangle \leqq 0$$

then

$$\lim_{l\to\infty} \langle B(u^l), u^l - u\rangle = 0, \tag{1.10}$$

and $B(u^l)$ converges weakly to $B(u)$ in X', i.e.

$$B(u^l) \overset{w'}{\to} B(u) \qquad \text{in } X'. \tag{1.11}$$

From (1.10) it follows that

$$\limsup_{l\to\infty} \langle A(u^l), u^l - u\rangle \leqq 0.$$

Since A is pseudomonotone thus

$$\left(A(u^l)\right) \overset{w'}{\to} A(u) \qquad \text{in } X',$$

and by (1.11) $T(u^l)$ converges weakly to $T(u)$ in X'. By (1.10) we have

$$\lim_{l\to\infty} \langle T(u^l), u^l - u\rangle = 0.$$

So we have shown that T is pseudomonotone operator.

Now we prove that $T + S_\mu$ is pseudomonotone. Suppose that (u^l) converges weakly to u in X and $(T + S_\mu)(u^l)$ converges weakly in X' to some y,

$$\lim_{l\to\infty} \langle (T + S_\mu)(u^l), u^l - u\rangle \leqq 0. \tag{1.12}$$

Then (by compact imbedding theorems) there is a subsequence (u^{l_k}) of (u^l) such that

$$\lim_{k\to\infty}(u^{l_k})=u \qquad \text{a.e. in } \Omega \text{ and on } \partial\Omega.$$

Thus Lebesgue's dominated convergence theorem implies

$$\text{(1.13)} \qquad \begin{aligned} &\lim \|g_j^\mu(x,u^{l_k})-g_j^\mu(x,u)\|_{L^q(\Omega)}=0,\\ &\lim \|h_{1,j}^\mu(x,u^{l_k})-h_{1,j}^\mu(x,u)\|_{L^q(\partial\Omega)}=0,\end{aligned}$$

where q is defined by $\frac{1}{p}+\frac{1}{q}=1$, hence

$$\lim_{k\to\infty} S_\mu(u^{l_k})=S_\mu(u) \qquad \text{weakly in } X'$$

and so

$$\text{(1.14)} \qquad \lim_{k\to\infty} T(u^{l_k})=y-S_\mu(u) \qquad \text{weakly in } X'.$$

From the equality

$$\langle S_\mu(u^{l_k}),u^{l_k}-u\rangle=\langle S_\mu(u^{l_k})-S_\mu(u),u^{l_k}-u\rangle+\langle S_\mu(u),u^{l_k}-u\rangle$$

it follows that

$$\text{(1.15)} \qquad \lim_{k\to\infty}\langle S_\mu(u^{l_k}),u^{l_k}-u\rangle=0,$$

because by (1.13), the boundedness of $\|u^{l_k}-u\|_X$, $\|u_j^{l_k}-u_j\|_{L^p(\partial\Omega)}$ and Hölder's inequality

$$\lim_{k\to\infty}\langle S_\mu(u^{l_k})-S_\mu(u),u^{l_k}-u\rangle=0.$$

It is not difficult to show that (1.15) is true for the original sequence, too. Therefore (1.12) implies

$$\text{(1.16)} \qquad \limsup_{l\to\infty}\langle T(u^l),u^l-u\rangle\leqq 0.$$

Since T is pseudomonotone thus by (1.14) and (1.16) we have

$$T(u)=y-S_\mu(u),$$

i.e.

$$(T+S_\mu)(u)=y.$$

Further,

$$\lim_{l\to\infty}\langle T(u^l),u^l-u\rangle=0,$$

and so by (1.15)

$$\lim_{l\to\infty}\langle (T+S_\mu)(u^l),u^l-u\rangle=0,$$

which completes the proof of Lemma 1.

LEMMA 2. *Assume that (u^l) converges weakly to u in X and there is a constant c such that*

$$\sum_{j=1}^{M}\int_{\Omega} g_j^l(x,u^l)u_j^l - \sum_{j=1}^{M}\int_{\partial\Omega} h_{1,j}^l(x,u^l)u_j^l \leqq c. \tag{1.17}$$

Then

$$g_j(x,u)\in L^1(\Omega), \qquad g_j(x,u)u_j \in L^1(\partial\Omega),$$
$$h_{1,j}(x,u)\in L^1(\partial\Omega), \qquad h_{1,j}(x,u)u_j\in L^1(\partial\Omega)$$

for all $j=1,2,\ldots,M$ and there is a subsequence (u^{l_k}) of (u^l), $u\in X$ such that

$$\lim_{k\to\infty} u^{l_k} = u \qquad \text{a.e. in } \Omega \text{ and on } \partial\Omega, \tag{1.18}$$

further

$$\lim_{k\to\infty} \|g_j^{l_k}(x,u^{l_k}) - g_j(x,u)\|_{L^1(\partial\Omega)} = 0,$$
$$\lim_{k\to\infty} \|h_{1,j}^{l_k}(x,u^{l_k}) - h_{1,j}(x,u)\|_{L^1(\partial\Omega)} = 0,$$
$$j=1,2,\ldots,M.$$

PROOF. As (u^l) converges weakly to u in X thus (by a compact imbedding theorem) there exists a subsequence (u^{l_k}) of (u_l) with property (1.18). Since g_j, $h_{1,j}$ satisfy the Carathéodory condition, it is easy to show that

$$\begin{aligned} &\lim_{k\to\infty} g^{l_k}(x,u^{l_k}) = g(x,u) \qquad \text{for a.e. } x\in\Omega,\\ &\lim_{k\to\infty} h_1^{l_k}(x,u^{l_k}) = h_1(x,u) \qquad \text{for a.e. } x\in\partial\Omega. \end{aligned} \tag{1.19}$$

By (1.4), (1.5), (1.17) and assumptions (f), (h) we have

$$\int_{\Omega}[g_j(x,u^l)]u_j^l \leqq c, \qquad -\int_{\partial\Omega}[h_{1,j}^l(x,u^l)]u_j^l \leqq c. \tag{1.20}$$

Therefore by (1.19), (f) and (h) Fatou's lemma implies

$$g_j(x,u)u_j\in L^1(\Omega), \qquad h_{1,j}(x,u)u_j\in L^1(\partial\Omega).$$

For any $\delta>0$ we have by (e)

$$|g_j^{l_k}(x,u^{l_k})| \leqq g_{j,\delta^{-1}}(x) + \delta|g_j^{l_k}(x,u^{l_k})u_k^{l_k}|.$$

This implies that $g^{l_k}(x, u^{l_k})$ is equiintegrable, because by (1.20)

$$\int_E |g_j^{l_k}(x, u^{l_k})|\, dx < \varepsilon$$

if the measure of E is sufficiently small and there is a set A_ε of finite measure with

$$\int_{\Omega \setminus A_\varepsilon} |g_j^{l_k}(x, u^{l_k})|\, dx < \varepsilon.$$

By Vitali's theorem and (1.19) this shows that

$$g_j^{l_k}(x, u^{l_k}) \to g_j(x, u) \qquad \text{in } L^1(\Omega).$$

Similarly can be proved that

$$h_{1,j}^{l_k}(x, u^{l_k}) \to h_{1,j}(x, u) \qquad \text{in } L^1(\partial\Omega).$$

LEMMA 3. *The operator*

$$T + S_\mu : X \to X'$$

is coercive, i.e.

$$\lim_{\|u\| \to \infty} \frac{\langle (T + S_\mu)(u), u \rangle}{\|u\|} = +\infty.$$

PROOF. From (f) and (h) we obtain

$$\int_\Omega g_j^\mu(x, u) u_j \geqq 0, \qquad \int_{\partial\Omega} h_{1,j}^\mu(x, u) u_j\, d\sigma \leqq 0.$$

This implies that $\langle S_\mu(u), u \rangle \geqq 0$. Thus by using conditions (d) and (i) we obtain

$$\begin{aligned} \langle (T + S_\mu)(u), u \rangle &= \langle T(u), u \rangle + \langle S_\mu(u), u \rangle \geqq \langle T(u), u \rangle \geqq \\ &\geqq c_2 \|u\|_X^p - c_3 - c_4 \|u\|_X^{\rho+1} - c_5 \end{aligned} \tag{1.21}$$

(c_2–c_5 are positive constants). From this inequality and $\rho + 1 < p$ it follows that $T + S_\mu$ is coercive.

2. The proof of the theorem

By Lemmas 1–3 the operator $T+S_l$ is bounded, pseudomonotone and coercive for all $l=1,2,3,\ldots$. By using the well-known theory of pseudomonotone operators in reflexive Banach spaces (see e.g. [5]) we obtain that for any $F\in X'$ there exists $u^l\in X$ such that

$$(T+S_l)(u^l)=F. \tag{2.1}$$

By Lemma 3 the sequence (u^l) is bounded in X (see (1.21), where the constants are independent of μ). T is a bounded operator and so the sequence $(T(u^l))$ is bounded in X'. Since X is a reflexive Banach space, thus there exist a subsequence (u^{l_k}) of (u^l) and $u\in X$ such that

$$\begin{aligned} &\lim_{k\to\infty}(u^{l_k})=u \qquad \text{weakly in } X,\\ &\lim_{k\to\infty}T(u^{l_k})=y \qquad \text{weakly in } X', \end{aligned} \tag{2.2}$$

for some $y\in X'$. Combining the definition of S_l with (2.1) we find that

$$\begin{aligned} &\sum_{j=1}^{M}\int_\Omega g_j^{l_k}(x,u^{l_k})u_j^{l_k}-\sum_{j=1}^{M}\int_{\partial\Omega}h_{1,j}^{l_k}(x,u^{l_k})u_j^{l_k}\,d\sigma=\\ &=\langle S_{l_k}(u^{l_k}),u^{l_k}\rangle=\langle F,u^{l_k}\rangle-\langle T(u^{l_k}),u^{l_k}\rangle\leqq\\ &\leqq\|F\|_{X'}\|u^{l_k}\|_X+\|T(u^{l_k})\|_{X'}\|u^{l_k}\|_X<c. \end{aligned}$$

Thus, by Lemma 2,

$$\begin{aligned} &g_j(x,u)u_j\in L^1(\Omega), \qquad g_j(x,u)\in L^1(\Omega),\\ &h_{1,j}(x,u)u_j\in L^1(\partial\Omega), \qquad h_{1,j}(x,u)\in L^1(\partial\Omega) \end{aligned} \tag{2.3}$$

and there is a subsequence $(u^{l'_k})$ of (u^{l_k}) such that

$$\lim_{k\to\infty}(u^{l'_k})=u \quad \text{a.e. in } \Omega, \qquad \lim_{k\to\infty}(u^{l'_k})=u \quad \text{a.e. in } \partial\Omega, \tag{2.4}$$

and also

$$\begin{aligned} &\lim_{k\to\infty}\|g_j^{l'_k}(x,u^{l'_k})-g_j(x,u)\|_{L^1(\Omega)}=0,\\ &\lim_{k\to\infty}\|h_{1,j}^{l'_k}(x,u^{l'_k})-h_{1,j}(x,u)\|_{L^1(\partial\Omega)}=0,\\ &(j=1,\ldots,M). \end{aligned} \tag{2.5}$$

According to (2.1) for any $v\in X$

$$\langle (T+S_{l'_k})(u^{l'_k}),v\rangle=\langle F,v\rangle. \tag{2.6}$$

Consider in (2.6) a fixed $v \in X$ such that $v \in L^\infty(\Omega) \times \ldots \times L^\infty(\Omega)$ and $v_{|\partial\Omega} \in$ $\in L^\infty(\partial\Omega) \times \ldots \times L^\infty(\partial\Omega)$.

By using (2.3)–(2.6) we obtain as $k \to \infty$

$$\langle y, v\rangle + \sum_{j=1}^{M} \int_\Omega g_j(x,u)v_j - \sum_{j=1}^{M} \int_{\partial\Omega} h_{1,j}(x,u)v_j = \langle F, v\rangle. \tag{2.7}$$

Now we shall show that $y = T(u)$. Since T is pseudomonotone, it is sufficient to prove the following inequality:

$$\limsup_{k\to\infty} \langle T(u^{l'_k}), u^{l'_k} - u\rangle \leqq 0.$$

We have

$$\langle T(u^{l'_k}), u^{l'_k} - u\rangle = \langle T(u^{l'_k}), u^{l'_k}\rangle - \langle T(u^{l'_k}), u\rangle$$

and so by (2.2), (2.6) and Fatou's lemma

$$\limsup_{k\to\infty} \langle T(u^{l'_k}), u^{l'_k} - u\rangle = \limsup_{k\to\infty} \langle F - S^{l'_k}(u^{l'_k}), u^{l'_k}\rangle - \langle y, u\rangle \leqq$$

$$\leqq \langle F - y, u\rangle - \liminf_{k\to\infty} \left\{ \sum_{j=1}^{M} \int_\Omega g_j^{l'_k}(x, u^{l'_k}) u_j^{l'_k} - \sum_{j=1}^{M} \int_{\partial\Omega} h_{1,j}^{l'_k}(x, u^{l'_k}) u_j^{l'_k}\, d\sigma \right\} \leqq$$

$$\leqq \langle F - y, u\rangle - \sum_{j=1}^{M} \int_\Omega g_j(x,u)u_j + \sum_{j=1}^{M} \int_{\partial\Omega} h_{1,j}(x,u)u_j\, d\sigma.$$

Thus for any $w \in X \cap L^\infty(\Omega) \times \ldots \times L^\infty(\Omega)$, by using (2.7),

$$\begin{aligned} &\limsup_{k\to\infty} \langle T(u^{l'_k}), u^{l'_k} - u\rangle \leqq \langle F - y, u - w\rangle + \\ &+ \sum_{j=1}^{M} \int_\Omega [g_j(x,u)](w_j - u_j) - \sum_{j=1}^{M} \int_{\partial\Omega} [h_{1,j}(x,u)](w_j - u_j)\, d\sigma. \end{aligned} \tag{2.7'}$$

Since $\partial\Omega$ is bounded and continuously differentiable, thus $u \in X$ can be extended to $\mathbf{R}^n$ such that we obtain $u \in W_p^1(\mathbf{R}^n) \times \ldots \times W_p^1(\mathbf{R}^n)$. We know (see [11]) there is a sequence (w_j^l) in $W_p^1(\mathbf{R}^n) \cap L^\infty(\mathbf{R}^n)$ such that (w_j^l) converges to u_j in $W_p^1(\mathbf{R}^n)$ and a.e. in $\mathbf{R}^n$, further

$$|w_j^l(x)| \leqq |u_j(x)| \qquad \text{a.e. in } \mathbf{R}^n,\ j = 1, \ldots, M. \tag{2.8}$$

Now we show that for the trace of w^l and u

$$\left| w_{j|\partial\Omega}^l(x) \right| \leqq \left| u_{j|\partial\Omega}(x) \right| \qquad \text{for a.e. } x \in \partial\Omega. \tag{2.9}$$

We know that for a.e. $y \in \mathbb{R}^n$

$$-|u_j(y)| \leqq w_j^l(y) \leqq |u_j(y)|.$$

Thus for any $\eta_\varepsilon \in C_0(\mathbb{R}^n)$ with $\operatorname{supp} \eta_\varepsilon \subset \bar{B}_\varepsilon$, $\eta_\varepsilon \geqq 0$ and $\int \eta_\varepsilon = 1$ we have

$$-\int_{\mathbb{R}^n} |u_j(y)|\eta_\varepsilon(x-y)\,dy \leqq \int_{\mathbb{R}^n} w_j^l(y)\eta_\varepsilon(x-y)\,dy \leqq \int_{\mathbb{R}^n} |u_j(y)|\eta_\varepsilon(x-y)\,dy,$$

and so, by using the notation $v_\varepsilon(x) := \int_{\mathbb{R}^n} v(y)\eta_\varepsilon(x-y)\,dy$

$$(2.10) \qquad -|u_j|_\varepsilon\big|_{\partial\Omega} \leqq w_{j,\varepsilon}^l\big|_{\partial\Omega} \leqq |u_j|_\varepsilon\big|_{\partial\Omega}.$$

Since $(w_{j,\varepsilon}^l) \to w_j^l$ and

$$\left(|u_j|_\varepsilon\right) \to |u_j| \qquad \text{in } W_p^1(\mathbb{R}^n) \text{ as } \varepsilon \to +0$$

thus

$$w_{j,\varepsilon}^l\big|_{\partial\Omega} \to w_j^l\big|_{\partial\Omega},$$

and

$$|u_j|_\varepsilon\big|_{\partial\Omega} \to |u_j|\big|_{\partial\Omega} \qquad \text{in } L^1(\partial\Omega) \text{ as } \varepsilon \to +0.$$

Consequently, for a suitable sequence (ε_k) with $\lim_{k\to\infty}(\varepsilon_k) = 0$ we have

$$w_{j,\varepsilon_k}^l\big|_{\partial\Omega} \to w_j^l\big|_{\partial\Omega}, \qquad |u_j|_\varepsilon\big|_{\partial\Omega} \to |u_j|\big|_{\partial\Omega}$$

a.e. on $\partial\Omega$ as $k \to \infty$. Therefore from (2.10) we obtain

$$-\left|u_j\big|_{\partial\Omega}(x)\right| \leqq w_j^l\big|_{\partial\Omega}(x) \leqq \left|u_j\big|_{\partial\Omega}(x)\right|$$

which proves (2.9). By (2.8), (2.10) and Lebesgue's dominated convergence theorem we have

$$\langle F - y, u - w^l \rangle \to 0,$$

and

$$\int_\Omega [g_j(x,u)]w_j^l\,dx \to \int_\Omega [g_j(x,u)]u_j, \qquad \int_{\partial\Omega} [h_{1,j}(x,u)]w_j^l\,d\sigma \to \int_{\partial\Omega} [h_{1,j}(x,u)]u_j\,d\sigma,$$

since $g(x,u)u \in L^1(\Omega)$, $h_1(x,u)u \in L^1(\partial\Omega)$. Thus, from (2.7), it follows that

$$\limsup_{k\to\infty} \langle T(u^{l'_k}), u^{l'_k} - u \rangle \leqq 0.$$

Consequently, $y = T(u)$, and

$$\langle T(u^{l'_k}), u^{l'_k} - u\rangle \to 0.$$

Therefore from (2.7) we obtain (1.2) for all $v \in W_p^1(\Omega) \times \ldots \times W_p^1(\Omega)$ with $v \in L^\infty(\Omega) \times \ldots \times L^\infty(\Omega)$, $v_{|\partial\Omega} \in L^\infty(\Omega) \times \ldots \times L^\infty(\Omega)$. Setting $v = w^l$ in (2.7) we find that (1.2) is true also for $v = u$. So the proof of the existence theorem is complete.

REFERENCES

[1] ADAMS, R. A., *Sobolev spaces*, Pure and Applied Mathematics, Vol. 65, Academic Press, New York–London, 1975. *MR* **56** #9247

[2] BROWDER, F. E., Pseudo-monotone operators and nonlinear elliptic boundary value problems on unbounded domains, *Proc. Nat. Acad. Sci. U.S.A.* **74** (1977), 2659–2661. *MR* **56** #3469

[3] BROWDER, F. E., Non-local elliptic boundary value problems, *Amer. J. Math.* **86** (1964), 735–750. *MR* **30** #2215

[4] HASSAN, I. M., Nonlocal and strongly nonlinear third boundary value problems, *Studia Sci. Math. Hungar.* **27** (1992), 223–233.

[5] LIONS, J.-L., Quelques méthodes de résolution des problèmes aux limites non-linéaires, Dunod; Gauthier–Villars, Paris, 1969. *MR* **41** #4326

[6] SIMON, L., Nonlinear elliptic differential equations with nonlocal boundary conditions, *Acta. Math. Acad. Sci. Hungar.* **56** (1990), 343–352.

[7] SIMON, L., On strongly nonlinear elliptic equations in unbounded domains, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **28** (1985), 241–252. *MR* **87m**:35099

[8] SIMON, L., Strongly nonlinear elliptic variational inequalities with nonlocal boundary conditions, *Qualitative theory of differential equations* (Szeged, 1988), Colloq. Math. Soc. J. Bolyai, **53**, North-Holland, Amsterdam, 1990, 605–620. *MR* **91k**:35095

[9] SIMON, L., Variational inequalities for strongly nonlinear elliptic operators, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **29** (1986), 231–240. *MR* **88f**:35056

[10] SINGH, K., On quasilinear elliptic systems in $\mathbf{R}^n$, *Studia Sci. Math. Hungar.* **24** (1989), 307–314.

[11] WEBB, J. R. L., Boundary value problems for strongly nonlinear elliptic equations, *J. London Math. Soc.* (2) **21** (1980), 123–132. *MR* **82e**:35039

[12] БИЦАДЗЕ, А. В., К теории нелокальных краевых задач (On the theory of nonlocal boundary value problems), *Dokl. Akad. Nauk SSSR* **277** (1984), 17–19 (in Russian). *MR* **86b**:31005

[13] САМАРСКИЙ, А. А., О некоторых проблемах теории дифференциальных уравнений (Some problems of the theory of differential equations), *Differentsial'nye Uravneniya* **16** (1980), 1925–1935 (in Russian). *MR* **82d**:35003

[14] СКУБАЧЕВСКИЙ, А. Л., Эллиптические задачи с нелокальными условиями вблизи границы (Elliptic problems with nonlocal conditions near the boundary), *Mat. Sb. (N.S.)* **129** (1986), 279–302. *MR* **87h**:35089

(Received February 6, 1991)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR
ALKALMAZOTT ANALÍZIS TANSZÉK
MÚZEUM KRT. 6–8
H–1088 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 259–266

ON AN OPERATIONAL CALCULUS WITH WEIGHTING ELEMENT

H. WYSOCKI

Abstract

A model of the Bittner operational calculus, in which the Taylor rest of an element does not depend on derivatives of that element, has been constructed in this paper. The examples of this model has been given.

1. Preliminaries

The Bittner operational calculus [1,2] is referred to as the system

$$CO(L^0, L^1, S, T_q, s_q, q, Q),$$

where L^0 and L^1 are linear spaces over the same field Γ of scalars; the linear operation $S : L^1 \to L^0$ (denoted as $S \in L(L^1, L^0)$), called the (abstract) derivative, is a surjection. Moreover, Q is an arbitrary nonempty set of indices q for the operations $T_q \in L(L^0, L^1)$ such that $ST_q f = f$, $f \in \in L^0$, called integrals and for the operations $s_q \in L(L^1, L^1)$ such that $s_q x = = x - T_q S x$, $x \in L^1$, called limit conditions.

By induction we define a sequence of spaces L^n, $n \in N$ such that

$$L^n := \{x \in L^{n-1} : Sx \in L^{n-1}\}.$$

Then

$$\ldots \subset L^n \subset L^{n-1} \subset \ldots \subset L^1 \subset L^0$$

and

$$S^n(L^{m+n}) = L^m,$$

where

$$L(L^n, L^0) \ni S^n := \underbrace{S \circ S \circ \ldots \circ S}_{n\text{-times}}, \quad n \in N, \quad m \in N_0 := N \cup \{0\}.$$

The kernel of S, i.e. $\operatorname{Ker} S := \{c \in L^1 : Sc = 0\}$, is called the set of constants for the derivative S.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 44A40, Secondary 26C05.

Key words and phrases. Operational calculus, Taylor's formula.

0081–6906/94/$ 4.00

It can be shown that the limit conditions s_q, $q \in Q$ are projections of L^1 onto the subspace Ker S. Hence we have

$$s_q^2 x = s_q x \in \operatorname{Ker} S, \quad q \in Q, \ x \in L^1 \tag{1}$$

and

$$s_q c = c, \quad q \in Q, \ c \in \operatorname{Ker} S. \tag{2}$$

For the element $x \in L^n$, $n \in N$ the following Taylor formula holds:

$$x = s_q x + T_q s_q S x + \cdots + T_q^{n-1} s_q S^{n-1} x + T_q^n S^n x, \quad q \in Q. \tag{3}$$

The expression $T_q^n S^n x$ is called the n-th Taylor rest of the element x at the point q.

Let L^0 be an algebra and L^1 its subalgebra. We say that the derivative S satisfies the Leibniz condition if

$$S(x \cdot y) = Sx \cdot y + x \cdot Sy, \quad x, y \in L^1. \tag{4}$$

We say that the limit condition s_q is multiplicative if

$$s_q(x \cdot y) = s_q x \cdot s_q y, \quad x, y \in L^1. \tag{5}$$

It follows from (2), (4), (5) that [3]

$$s_q(c \cdot x) = c s_q x, \quad T_q(c \cdot f) = c T_q f, \quad q \in Q, \ c \in \operatorname{Ker} S, \ x \in L^1, \ f \in L^0. \tag{6}$$

We also infer from (4) that

$$(c, d \in \operatorname{Ker} S) \Rightarrow (cd \in \operatorname{Ker} S) \tag{7}$$

and thus Ker S is the subalgebra of L^0.

2. Integration by parts

For further discussion we shall assume that

— Q has more than one element,

— L^0 is a commutative algebra and L^1 is its subalgebra,

— the derivative S satisfies the Leibniz condition (4),

— the operations s_q, $q \in Q$, satisfy the multiplication condition (5).

The mapping $I_{q_1}^{q_2} \in L(L^0, L^0)$ described by the formula

$$I_{q_1}^{q_2} f := (T_{q_1} - T_{q_2}) f, \quad q_1, q_2 \in Q, \ f \in L^0 \tag{8}$$

is called the operation of definite integration.

It is easy to verify that the operation of definite integration has the following properties:

$$I_{q_1}^{q_2} = s_{q_2}T_q - s_{q_1}T_q, \quad I_{q_1}^{q_2} = s_{q_2}T_{q_1}, \quad I_{q_1}^{q} + I_{q}^{q_2} = I_{q_1}^{q_2}, \quad I_{q_1}^{q_2} = -I_{q_2}^{q_1},$$

where $q, q_1, q_2 \in Q$.

Moreover,

$$(9) \qquad I_{q_1}^{q_2}(L^0) \subset \operatorname{Ker} S.$$

Directly from (8) and (6) it follows that

$$(10) \qquad I_{q_1}^{q_2}(c \cdot f) = cI_{q_1}^{q_2}f, \quad q_1, q_2 \in Q, \ c \in \operatorname{Ker} S, \ f \in L^0.$$

Let $R_{q_1}^{q_2} \in L(L^1, \operatorname{Ker} S)$ be an operation such that

$$(11) \qquad R_{q_1}^{q_2}x := (s_{q_2} - s_{q_1})x, \quad q_1, q_2 \in Q, \ x \in L^1.$$

It is easy to prove that the Leibniz–Newton formula holds:

$$I_{q_1}^{q_2}Sx = R_{q_1}^{q_2}x, \quad q_1, q_2 \in Q, \ x \in L^1.$$

THEOREM 1. *The integration by parts formula holds:*

$$(12) \qquad I_{q_1}^{q_2}(x \cdot S^n y) = \sum_{i=0}^{n-1}(-1)^i R_{q_1}^{q_2}(S^i x \cdot S^{n-1-i}y) + (-1)^n I_{q_1}^{q_2}(S^n x \cdot y),$$
$$q_1, q_2 \in Q, \ x, y \in L^n, \ n \in N.$$

PROOF. We shall prove this formula by induction on $n \in N$. The case when $n = 1$ determines the statement of Theorem 2.3 [3]. Assuming the validity of the formula (12) for a fixed $n \in N$, for $x, y \in L^{n+1}$ we have

$$I_{q_1}^{q_2}(x \cdot S^{n+1}y) = I_{q_1}^{q_2}[x \cdot S^n(Sy)] =$$
$$= \sum_{i=0}^{n-1}(-1)^i R_{q_1}^{q_2}(S^i x \cdot S^{n-i}y) + (-1)^n I_{q_1}^{q_2}(S^n x \cdot Sy) =$$
$$= \sum_{i=0}^{n-1}(-1)^i R_{q_1}^{q_2}(S^i x \cdot S^{n-i}y) + (-1)^n[R_{q_1}^{q_2}(S^n x \cdot y) - I_{q_1}^{q_2}(S^{n+1}x \cdot y)] =$$
$$= \sum_{i=0}^{n}(-1)^i R_{q_1}^{q_2}(S^i x \cdot S^{n-i}y) + (-1)^{n+1} I_{q_1}^{q_2}(S^{n+1}x \cdot y).$$

An application of the induction principle finishes the proof for any $n \in N$. □

The other properties of the derivative S satisfying the Leibniz condition (4) are discussed in the works [9,10].

3. The operational calculus with weighting element

Assume more that L^1 is an algebra with unity e. It is easy to verify that $e \in \operatorname{Ker} S$ and

$$\text{(13)} \qquad \text{if } c \in \operatorname{Ker} S \text{ is invertible in } L^1, \text{ then } c^{-1} \in \operatorname{Ker} S.$$

Denote by $\operatorname{Inv}(\operatorname{Ker} S)$ the set of invertible elements in the algebra $\operatorname{Ker} S$.

Let an element $g \in L^0$ be given such that

$$g_q := I_q^{q_1} g \in \operatorname{Inv}(\operatorname{Ker} S), \quad q \in \hat{Q} := Q - \{q_1\},$$

where $q_1 \in Q$ is fixed.

The element g will be called the weighting element.

THEOREM 2. *The system* $(L^0, L^1, \hat{S}, \hat{T}_q, \hat{s}_q, q, \hat{Q})$, *where* $\hat{S} := S$ *and*

$$\hat{T}_q f := T_q f - (I_q^{q_1} g)^{-1} I_q^{q_1}(g \cdot T_q f), \quad q \in \hat{Q}, \ f \in L^0,$$
$$\hat{s}_q x := (I_q^{q_1} g)^{-1} I_q^{q_1}(g \cdot x), \quad q \in \hat{Q}, \ x \in L^1$$

forms an operational calculus.

PROOF. It is not difficult to notice that $\hat{S}, \hat{T}_q, \hat{s}_q, q \in \hat{Q}$ are linear operations. Moreover,

$$\hat{S}\hat{T}_q f = f, \quad q \in \hat{Q}, \ f \in L^0,$$

what follows from the axiom

$$S T_q f = f, \quad q \in Q, \ f \in L^0$$

and from properties (9), (13) and (7).

From the axiom

$$T_q S x = x - s_q x, \quad q \in Q, \ x \in L^1$$

and from (1), (10) we also obtain

$$\begin{aligned} \hat{T}_q \hat{S} x &= T_q S x - (I_q^{q_1} g)^{-1} I_q^{q_1}(g \cdot T_q S x) = \\ &= x - s_q x - g_q^{-1} I_q^{q_1}(g \cdot x) + g_q^{-1} \cdot g_q \cdot s_q x = \\ &= x - \hat{s}_q x, \qquad q \in \hat{Q}, \ x \in L^1. \end{aligned}$$

Thus $\hat{S}$ is the derivative, the operations $\hat{T}_q, q \in \hat{Q}$ are integrals and the operations $\hat{s}_q, q \in \hat{Q}$ are limit conditions. □

Using (6), (10) and (13) it is easy to verify that

$$\text{(14)} \quad \hat{s}_q(c \cdot x) = c\hat{s}_q x, \ \hat{T}_q(c \cdot f) = c\hat{T}_q f, \quad q \in \hat{Q}, \ c \in \operatorname{Ker} S, \ x \in L^1, \ f \in L^0.$$

THEOREM 3. *If the weighting element g satisfies the following conditions*

$$g \in L^{n-1}, \quad s_q S^i g = 0, \quad q \in Q, \quad i = 0, 1, \ldots, n-2, \quad n \geqq 2,$$

then the n-th Taylor rest $\hat{T}_q^n \hat{S}^n x$, $q \in \hat{Q}$ *does not depend on derivatives of an element* $x \in L^n$, $n \in N$.

PROOF. For $n = 1$ the theorem is evident. If $n > 1$ then from the assumptions of the weighting element g, from multiplication condition of the operations s_q, $q \in Q$ and from (11), (12) we get

$$\hat{s}_q \hat{S}^i x = \hat{s}_q S^i x = g_q^{-1} I_q^{q_1}(g \cdot S^i x) = (-1)^i g_q^{-1} I_q^{q_1}(S^i g \cdot x),$$
$$q \in \hat{Q}, \quad x \in L^n, \quad i = 1, 2, \ldots, n-1.$$

Hence and from Taylor formula (3) we obtain

$$\hat{T}_q^n \hat{S}^n x = x - \sum_{i=0}^{n-1} \hat{T}_q^i \hat{s}_q \hat{S}^i x = x - \sum_{i=0}^{n-1} (-1)^i \hat{T}_q^i [g_q^{-1} I_q^{q_1}(S^i g \cdot x)], \quad q \in \hat{Q}, \quad x \in L^n,$$

what means the proposition of the theorem. □

The last equality can be also rewritten in the form

$$\hat{T}_q^n \hat{S}^n x = x - g_q^{-1} \sum_{i=0}^{n-1} (-1)^i I_q^{q_1}(S^i g \cdot x) \cdot \hat{T}_q^i e, \quad q \in \hat{Q}, \quad x \in L^n.$$

4. Examples

A. Let be given an operational calculus, in which

$$L^n := C^n([a, b], R^1), \quad n \in N_0$$

and

$$Sx := \left\{ \frac{1}{w(t)} \frac{dx}{dt} \right\}, \quad T_{t_0} f := \left\{ \int_{t_0}^{t} w(\tau) f(\tau)\, d\tau \right\}, \quad s_{t_0} x := \{x(t_0)\},$$

where $t_0 = q \in Q := [a, b] \subset R^1$, $x = \{x(t)\} \in L^1$, $f = \{f(t)\} \in L^0$, $w = \{w(t)\} \in$ $\in L^0$ and $w(t) \neq 0$ for any $t \in [a, b]$.

For usual multiplication of functions, the spaces L^n, $n \in N_0$, are algebras such that $L^n \subset L^{n-1}$, $n \in N$, the derivative S satisfies the Leibniz condition and the operations s_{t_0}, $t_0 \in [a,b]$ are multiplicative. Let $g = \{g(t)\} \in L^0$ be a function such that

$$g_{t_0} = \int_{t_0}^{t_k} w(\tau) g(\tau)\, d\tau \neq 0,$$

where $t_k \in [a,b]$ is fixed.

From Theorem 2 it follows that the operations

$$\hat{T}_{t_0} f = \left\{ \int_{t_0}^{t} w(\tau) f(\tau)\, d\tau - \frac{1}{g_{t_0}} \int_{t_0}^{t_k} \left[w(t) g(t) \int_{t_0}^{t} w(\tau) f(\tau)\, d\tau \right] dt \right\},$$
$$f = \{f(t)\} \in L^0,$$
$$\hat{s}_{t_0} x = \left\{ \frac{1}{g_{t_0}} \int_{t_0}^{t_k} w(\tau) g(\tau) x(\tau)\, d\tau \right\}, \quad x = \{x(t)\} \in L^1$$

are integrals and limit conditions for the derivative $S = \frac{1}{w(t)} \frac{d}{dt}$, respectively.

The example of such operational calculus that $w = \{1\}$ and $g = \{1\}$ has been presented by Tasche [8] and generalized by Przeworska-Rolewicz [6] in case when $w = \{1\}$.

B. Let us consider the operational calculus [4, 5] with the derivative

$$Sx := \left\{ \alpha \frac{\partial x(z,t)}{\partial z} + \beta \frac{\partial x(z,t)}{\partial t} \right\}, \quad \alpha, \beta \in R^1, \ \beta \neq 0,$$

the integrals

$$T_{t_0} f := \left\{ \frac{1}{\beta} \int_{t_0}^{t} f\Big(z - \frac{\alpha}{\beta}(t-\tau), \tau\Big)\, d\tau \right\}$$

and limit conditions

$$s_{t_0} x := \left\{ x\Big(z - \frac{\alpha}{\beta}(t - t_0), t_0\Big) \right\},$$

where $t_0 = q \in Q := [a,b] \subset R^1$, $f = \{f(z,t)\} \in L^0 := C^1(R^1 \times [a,b], R^1)$, $x = = \{x(z,t)\} \in L^1 = \{f \in L^0 : Sf \in L^0\}$.

It is not difficult to verify that for the usual multiplication of functions of two variables the derivative S satisfies the Leibniz condition and the operations s_{t_0}, $t_0 \in [a,b]$ are multiplicative. In this case the operations $\hat{T}_{t_0}$ and

$\hat{s}_{t_0}$ have the following forms

$$\hat{T}_{t_0} f = \left\{ \frac{1}{\beta} \int_{t_0}^{t} f(z - \frac{\alpha}{\beta}(t-\tau), \tau)\, d\tau - \right.$$

$$\left. - \frac{1}{\beta g_{t_0}(z,t)} \int_{t_0}^{t_k} \left[g(z - \frac{\alpha}{\beta}(t-\tau), \tau) \int_{t_0}^{\tau} f(z - \frac{\alpha}{\beta}(t-u), u)\, du \right] d\tau \right\},$$

$$f = \{f(z,t)\} \in L^0,$$

$$\hat{s}_{t_0} x = \left\{ \frac{1}{g_{t_0}(z,t)} \int_{t_0}^{t_k} g(z - \frac{\alpha}{\beta}(t-\tau), \tau) x(z - \frac{\alpha}{\beta}(t-\tau), \tau)\, d\tau \right\},$$

$$x = \{x(z,t)\} \in L^1,$$

respectively, where $g = \{g(z,t)\} \in L^0$ is a function such that

$$g_{t_0}(z,t) = \int_{t_0}^{t_k} g(z - \frac{\alpha}{\beta}(t-\tau), \tau)\, d\tau \neq 0$$

for each $(z,t) \in R^1 \times [a,b]$ and fixed $t_k \in [a,b]$.

REFERENCES

[1] Bittner, R., Algebraic and analytic properties of solutions of abstract differential equations, *Rozprawy Mat.* **41** (1964), 63pp. *MR* **29** #6341

[2] Bittner, R., *Rachunek operatorów w przestrzeniach liniowych* (Operational calculus in linear spaces), PWN, Warszawa, 1974. (See *Zbl* **348** #44008.)

[3] Bittner, R. and Mieloszyk, E., About eigenvalues of differential equations in the operational calculus, *Zesz. Nauk. Politech. Gdańsk.* **285** (1978), *Mat.* 11, 87–99. *Zbl* **402** #34006

[4] Bittner, R. and Mieloszyk, E., Application of the operational calculus to solving non-homogeneous linear partial differential equations of the first order with real coefficients, *Zesz. Nauk. Politech. Gdańsk.* **345** (1982), *Mat.* 12, 33-45. *Zbl* **517** #35014

[5] Mieloszyk, E., Operational calculus in algebras, *Publ. Math. Debrecen* **34** (1987), 137–143. *MR* **88m**:44012

[6] Przeworska-Rolewicz, D., *Algebraic analysis*, PWN – Polish Scientific Publishers, Warszawa; D. Reidel Publ. Co., Dordrecht–Boston, MA, 1988. *MR* **89m**: 00005

[7] Sieńczewski, G., Deterministic models of stationary linear systems with right invertible operators and identification of their parameters, Institute of Mathematics, Polish Academy of Sciences, Warszawa, Preprint No 260, May 1982.

[8] Tasche, M., Algebraische Operatorenrechnung für einen rechtsinvertierbaren Operator, *Wiss. Z. Univ. Rostock* **23** (1974), 735–744.

[9] Wysocki, H., The result derivative; distributive results, *Acta Math. Hungar.* **53** (1989), 289–307. *MR* **90m**:44017

[10] Wysocki, H., Distributions of finite order in the operational calculus, *Publ. Math. Debrecen* **38** (1991), 49–68. *MR* **92f**:44014

(Received February 11, 1991)

KATEDRA MATEMATYKI
AKADEMIA MARYNARKI WOJENNEJ
PL–81-919 GDYNIA
POLAND

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 267–278

CONTINUOUS AND SYMMETRIC PRODUCTS IN THE BICYCLIC SEMIGROUP

A. PATELLI and B. PIOCHI

Abstract

A product $(i_1, j_1)(i_2, j_2)\dots(i_n, j_n)$ in the bicyclic semigroup $C(p,q)$ is said to be *continuous* if $|i_h - i_{h+1}| \leqq 1$ and $|j_h - j_{h+1}| \leqq 1$ for every $h = 1, \dots, n-1$ and to be *symmetric* if $i_h = j_{n-h+1}$ and $j_h = i_{n-h+1}$, for every $1 \leqq h \leqq n$. We give formulas to compute continuous products, when they contain no idempotents or they are symmetric and contain at most three idempotents.

The free semigroup on two generators p and q with the condition $pq = 1$ is called *bicyclic semigroup* and denoted by $C(p,q)$. Many properties of $C(p,q)$ are well known; it is also concerned with many problems about magnifying elements of a semigroup, as it was proved by Ljapin [3], Desq [1], Migliorini [5] and Gerente [2].

Some properties of $C(p,q)$ were studied by Migliorini [5], who proved that $C(p,q)$ is a part of a minimal semigroup associated with a magnifying element. One of these properties concerned the so-called *anti-diagonal, squared* and *pincer products*, which were proved to be idempotents. In this note we give an extension of such results to the more general case of a continuous or symmetric product of elements of $C(p,q)$.

1. Diagonalization of $C(p,q)$

The following arrangement of the elements of $C(p,q)$ is well known:

$$\begin{array}{c|cccccc}
 & 1 & 2 & 3 & \cdots & j & \cdots \\ \hline
1 & (1,1) & (1,2) & (1,3) & \cdots & (1,j) & \cdots \\
2 & (2,1) & (2,2) & (2,3) & \cdots & (2,j) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
i & (i,1) & (i,2) & (i,3) & \cdots & (i,j) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

Fig. 1.1

1991 *Mathematics Subject Classification.* Primary 20M10.
Key words and phrases. Semigroups, bicyclic semigroups, identities.

where the pair (i,j) is the element $q^i p^j$.

The following properties hold (see [5]):

(i) Every idempotent of $C(p,q)$ is an element (i,i);

(ii) every idempotent is the vertex of a subtable, namely the set of those elements (k,j) with $k,j \geqq i$: this is isomorphic to $C(p,q)$ (consider $(k,j) \longrightarrow \rightarrow (k-i, j-i)$).

In the subtable with vertex (i,i) one can consider some special "paths". Migliorini in [5] considered the *antidiagonal, squared, pincers products*, which are illustrated in Figs. 1.2, 1.3, 1.4. He proved that if (i,k) with $k>i$ is the first element of the product, then all of these are equal to (i,i).

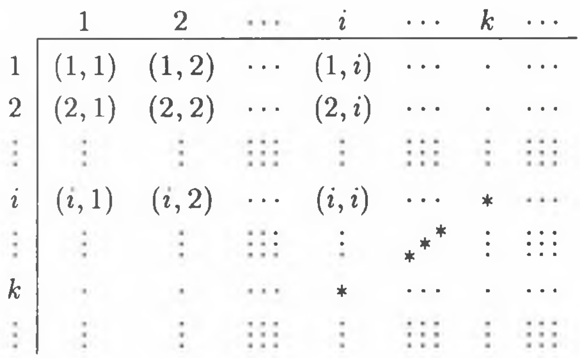


Fig. 1.2. Antidiagonal product:
$(i,k)(i+1,k-1)\dots(k-1,i+1)(k,i)=(i,i)$

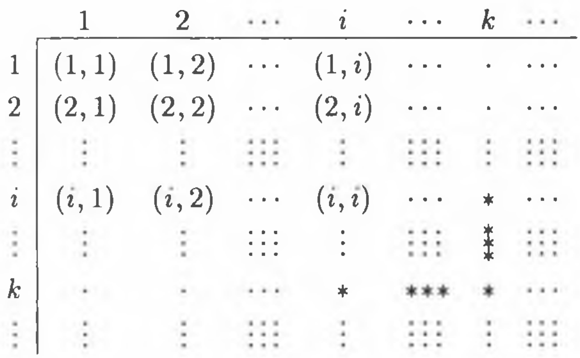


Fig. 1.3. Squared product:
$(i,k)(i+1,k)\dots(k,k)\dots(k,i+1)(k,i)=(i,i)$

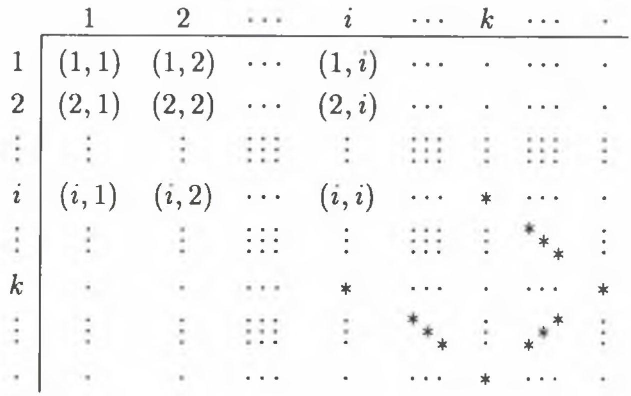


Fig. 1.4. Pincers product:
$(i,k)(i+1,k+1)\dots(i+j,k+j)(i+j-1,k+j-1)\dots$
$\dots(k+j,i+j)(k+j-1,i+j-1)\dots(k+1,i+1)(k,i)=(i,i)$

2. Symmetries in the bicyclic semigroup

Consider the elements $a_1=(i_1,j_1)$, $a_2=(i_2,j_2)$ and recall the definition of product in $C(p,q)$:

$$a_1a_2=(i_1-j_1+\max(j_1,j_2),j_2-i_2+\max(j_1,i_2)).$$

In the following, we often denote the pair (i_k,j_k) by a_k and the product $a_1a_2\dots a_n$ by P_n; we say that n is the *length* of P_n.

LEMMA 2.1. *Let* $(i,j)=P_n=a_1a_2\dots a_n$ *be a product in* $C(p,q)$. *Then* $i-j=\sum_{h=1}^{n}(i_h-j_h)$. *Also* $i\geqq i_1$ *and* $j\geqq j_n$.

PROOF. Consider a product of length equal to 2. Note that $(m,n)(p,q)$ is equal to $(m+p-n,q)$ (if $p\geqq n$) or to $(m,q+n-p)$ (if $p\leqq n$). In both cases the difference is equal to

$$m+p-n-q=(m-n)+(p-q);$$

the other part of the assertion is trivial.

This matter can be easily used to get a full proof, by induction on the length n of the product. □

DEFINITION 2.2. The product P_n is *continuous* if

$$|i_k-i_{k+1}|\leqq 1 \quad \text{and} \quad |j_k-j_{k+1}|\leqq 1$$

for all k, $1\leqq k\leqq n$.

DEFINITION 2.3. The product P_n is *symmetric* if

$$i_k = j_{n-k+1}; \qquad j_k = i_{n-k+1}$$

for all k, $1 \leqq k \leqq n$.

DEFINITION 2.4. We call an *E-intersection* in the product P_n:
(i) every element a_k with $i_k = j_k$;
(ii) every pair $[a_k, a_{k+1}]$ with the following properties:

$$\text{(a)}\ \ i_k = j_{k+1}, \quad i_{k+1} = j_k,$$
$$\text{(b)}\ \ i_k > j_k, \quad i_{k+1} < j_{k+1}, \qquad \text{or} \qquad i_k < j_k, \quad i_{k+1} > j_{k+1}.$$

We speak of E-intersection of type 2.4(i) or type 2.4(ii), respectively.

PROPOSITION 2.5. *If the product $(i,j) = P_n = a_1 a_2 \dots a_n$ is continuous and has no E-intersection, then:*
if $i_1 < j_1$, then $i = i_1$, $j = i_1 + \sum_{h=1}^{n} (j_h - i_h)$;
if $i_1 > j_1$, then $j = j_n$, $i = j_n + \sum_{h=1}^{n} (i_h - j_h)$.

PROOF. Let $(i,j) = P_n = a_1 a_2 \dots a_n$ be a continuous product with no E-intersection. Let us prove that if $i_1 < j_1$ then $i_h < j_h$ for every h, $1 \leqq h \leqq n$.

Suppose this is not true. Then $n > 1$ and let k be the least integer $1 \leqq k \leqq$ $\leqq n-1$ such that $i_k < j_k$ and $i_{k+1} > j_{k+1}$ (trivially, it cannot be $i_{k+1} = j_{k+1}$, or a_{k+1} would be an E-intersection). Then

$$i_{k+1} - 1 \leqq i_k < j_k \leqq j_{k+1} + 1 < i_{k+1} + 1.$$

Hence

$$j_k = i_{k+1} = j_{k+1} + 1; \qquad i_{k+1} - 1 = i_k = j_{k+1}.$$

The pair $[a_k, a_{k+1}]$ is an E-intersection of type 2.4(ii). Contradiction.

Now, if $n = 1$ then it is trivially true that $i = i_1$. Suppose that this has been proved up to

$$a_1 a_2 \dots a_{n-1} = (i_1, i_1 + \sum_{h=1}^{n-1} (j_h - i_h)).$$

Then

$$a_1 a_2 \dots a_n = (i_1, i_1 + \sum_{h=1}^{n-1} (j_h - i_h))(i_n, j_n).$$

Since

$$i_1 + \sum_{h=1}^{n-1} (j_h - i_h) \geqq j_{n-1} \geqq j_n - 1 > i_n - 1,$$

then

$$i_1 + \sum_{h=1}^{n-1} (j_h - i_h) \geqq i_n.$$

Whence

$$a_1 a_2 \dots a_n = (i_1, i_1 + \sum_{h=1}^{n-1} (j_h - i_h) + j_n - i_n) = (i_1, i_1 + \sum_{h=1}^{n} (j_h - i_h)).$$

Similarly, if $i_1 > j_1$ then $i_h > j_h$ for every h, $1 \leqq h \leqq n$; thus $j = j_n$ and, by Lemma 2.1

$$i = j_n + \sum_{h=1}^{n} (i_h - j_h). \qquad \square$$

A symmetric product always contains an E-intersection.

LEMMA 2.6. *Let $a_1 a_2 \dots a_n$ be a symmetric product with length $n > 1$; then P_n has an E-intersection. Namely:*

(i) *if $n = 2r + 1$, then a_{r+1} is an E-intersection of type* 2.4(i);

(ii) *if $n = 2r$, then $[a_r, a_{r+1}]$ is an E-intersection of type* 2.4(ii).

PROOF. (i) If we replace k by $r + 1$ and n by $2r + 1$ in Definition 2.4, then $i_{r+1} = j_{2r+1-(r+1)+1} = j_{r+1}$ and we get an E-intersection of type 2.4(i).

(ii) Again, if we replace n by $2r$ and k by r in Definition 2.4, then we have $i_r = j_{2r-r+1} = j_{r+1}$; likewise $j_r = i_{r+1}$ and we get an intersection of type 2.4(ii). $\square$

Note that, when the intersection is of type 2.4(ii), then it must be $|i_r - j_r| = |i_{r+1} - j_{r+1}| = 1$. In fact, these two differences must be equal to each other and less than or equal to 1, since the pair $[a_r, a_{r+1}]$ is an E-intersection and they cannot be 0, since a_r and a_{r+1} are not idempotents.

We say that two products of elements in $C(p, q)$ are *equivalent* if they are the same element.

LEMMA 2.7. *Every (symmetric) product is equivalent to a (symmetric) product which has only E-intersections of type* 2.4(i).

PROOF. Let $P_n = a_1 a_2 \dots a_n$ be a (symmetric) product. If all the E-intersections in P_n are of type we wish, then nothing has to be proved.

Suppose that there exists an E-intersection of type 2.4(ii), namely the pair $[a_r, a_{r+1}]$, where $i_{r+1} = j_r$ and $j_{r+1} = i_r$. Then P_n is equivalent to the product:

$$P'_n = a_1 a_2 \dots a_r (i_{r+1}, j_r) a_{r+1} \dots a_n.$$

If P_n is symmetric, then the product P'_n, too, is trivially symmetric, its length is equal to $n + 1$ and the $(r + 1)$-th element in P'_n is an intersection of type 2.4(i).

By iterating this procedure as many times as necessary, then we get a (symmetric) product, which has only E-intersections of type 2.4(i) and is equivalent to P. □

Now we are going to prove the main result of this section. We always suppose that all the E-intersections in a symmetric product are idempotent elements of type 2.4(i).

PROPOSITION 2.8. *Every symmetric product is idempotent.*

PROOF. By Lemma 2.6, there exists an E-intersection in the symmetric product $P_n = a_1 a_2 \dots a_n$. We can also suppose, by Lemmas 2.6 and 2.7, that the length n of the product P_n is an odd number: $n = 2r + 1$, and that the E-intersection is of type 2.4(i), namely the element (i_{r+1}, j_{r+1}), with $i_{r+1} = j_{r+1}$.

Now, consider the last r elements of the product P_n:

$$\begin{aligned} a_{r+2} \dots a_{n-1} a_n &= (j_r, i_r) \dots (j_2, i_2)(j_1, i_1) = \\ &= ((i_1, j_1)(i_2, j_2) \dots (i_r, j_r))^{-1} = \\ &= (a_1 a_2 \dots a_r)^{-1}. \end{aligned}$$

Thus, since a_{r+1} is an idempotent element,

$$\begin{aligned} P_n = a_1 a_2 \dots a_n &= a_1 a_2 \dots a_r a_{r+1} a_{r+2} \dots a_{n-1} a_n = \\ &= a_1 a_2 \dots a_r a_{r+1} (a_1 a_2 \dots a_r)^{-1} \end{aligned}$$

is idempotent. □

THEOREM 2.9. *Let $P_n = a_1 a_2 \dots a_n$ be a continuous and symmetric product with exactly one E-intersection, the idempotent element a_{r+1}. Then*
if $i_1 < j_1$, then $P_n = (i_1, i_1)$,
if $i_1 > j_1$, then $P_n = (j_{r+1} + \sum_{h=1}^{r} (i_h - j_h), j_{r+1} + \sum_{h=1}^{r} (i_h - j_h))$.

PROOF. Suppose that $r > 1$ (or the Proposition would be trivial) and that $i_1 < j_1$. As $a_{r+1} = (i_{r+1}, j_{r+1})$ is the only one idempotent element in $\{a_1, a_2, \dots, a_n\}$, then $i_h < j_h$ should be for every $1 \leqq h \leqq r$.

Since $n = 2r + 1$ and a_{r+1} is idempotent, we may consider

$$\begin{aligned} P_n = a_1 a_2 \dots a_n &= a_1 a_2 \dots a_r a_{r+1} a_{r+2} \dots a_{n-1} a_n = \\ &= a_1 a_2 \dots a_r a_{r+1} (a_1 a_2 \dots a_r)^{-1}. \end{aligned}$$

By Proposition 2.5, $a_1 a_2 \dots a_r = (i_1, j)$; also, $j \geqq j_r > i_r \geqq i_{r+1} - 1$, that is $j \geqq i_{r+1} = j_{r+1}$. Whence

$$P_n = (i_1, j)(i_{r+1}, j_{r+1})(j, i_1) = (i_1, i_1).$$

By similar arguments, if $i_1 > j_1$, then

$$P_n = (j_r + \sum(i_h - j_h), j_r)(i_{r+1}, j_{r+1})(j_r, j_r + \sum(i_h - j_h)).$$

Note that $j_r \leqq j_{r+1}$; in fact $j_r < i_r \leqq i_{r+1} + 1 = j_{r+1} + 1$.

If $j_r = j_{r+1} (= i_{r+1})$, the proof is readily finished. Suppose $j_r < j_{r+1}$, that is $j_r + 1 = j_{r+1}$ (since P_n is a continuous product). Then

$$\begin{aligned} P_n &= (j_r + \sum(i_h - j_h) + 1, j_{r+1})(j_r, j_r + \sum(i_h - j_h)) = \\ &= (j_r + 1 + \sum(i_h - j_h), j_r + 1 + \sum(i_h - j_h)) = \\ &= (j_{r+1} + \sum(i_h - j_h), j_{r+1} + \sum(i_h - j_h)). \qquad \square \end{aligned}$$

In Figure 2.10 we can see an application of Theorem 2.9: the symmetric product $P_n = a_1 a_2 \ldots a_{21}$ is equal to the idempotent $(2, 2)$.

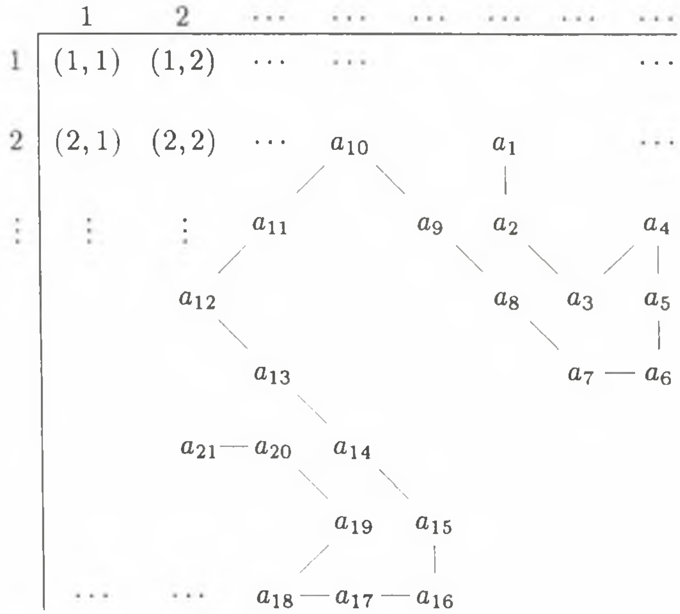


Fig. 2.10. An example of a continuous symmetric product

COROLLARY 2.11. *Let $P_n = a_1 a_2 \ldots a_n$ be a continuous and symmetric product with exactly three non-adjacent idempotent E-intersections:*

$$a_{k+1} = (i_{k+1}, j_{k+1}), \qquad a_{r+1} = (i_{r+1}, j_{r+1}),$$
$$a_{t+1} = (i_{t+1}, j_{t+1}) = (i_{k+1}, j_{k+1}).$$

Let i_1 be greater than j_1 and i_{k+2} be less than j_{k+2}. Then the product is equal to the idempotent element

$$P_n = (j_{k+1} + \sum_{h=1}^{k} (i_h - j_h), j_{k+1} + \sum_{h=1}^{k} (i_h - j_h)).$$

PROOF. Consider the restricted product

$$P'_n = a_{k+2} a_{k+3} \ldots a_{t-1} a_t.$$

By Theorem 2.9, this is equal to (i_{k+2}, i_{k+2}). Also $i_{k+2} < j_{k+2} \leqq j_{k+1} + 1$; whence $i_{k+2} \leqq j_{k+1}$. Thus

$$P_n = a_1 a_2 \ldots a_{k+1} (i_{k+2}, i_{k+2}) a_{k+1} a_k \ldots a_1 =$$
$$= \Big(j_{k+1} + \sum_{h=1}^{k} (i_h - j_h), j_{k+1}\Big) (i_{k+2}, i_{k+2}) (j_{k+1}, j_{k+1} + \sum_{h=1}^{k} (i_h - j_h)) =$$
$$= \Big(j_{k+1} + \sum_{h=1}^{k} (i_h - j_h), j_{k+1} + \sum_{h=1}^{k} (i_h - j_h)\Big). \qquad \square$$

Remark that all the hypotheses of Corollary 2.11 cannot be removed, or the result will not be predictable. For instance, if $i_1 < j_1$, one can easily give examples to show that the product depends on the length of its parts, whose elements have $i_h < j_h$ or $i_h > j_h$.

3. Special symmetric products

In this section we want to generalize the definitions of special products which were considered in [5] and to give general formulas to calculate their values. Of course, there can be many variations: antidiagonal and pincers products can have E-intersections of type 2.4(i) or 2.4(ii); squared and pincers products can have the E-intersection at the "bottom" or at the "top" of design; at last it can be $i_1 < j_1$ or $i_1 > j_1$.

We give the definitions only for the case $i_1 > j_1$: one can easily get the converse, by exchanging the i_h's with j_h's in every element of the product.

DEFINITION 3.1. Let $i \geqq 1$ and r, $0 \leqq r \leqq i - 1$, be integers. The following are *antidiagonal products*:

$$\prod_{h=1}^{2r+1}(i+r-(h-1),i-r+(h-1))=(i+r,i-r)(i+r-1,i-r+1)\dots$$
$$\dots(i,i)\dots(i-r+1,i+r-1)(i-r,i+r),$$
$$\prod_{h=1}^{2r}(i+r-(h-1),i-r+h)=(i+r,i-r+1)(i+r-1,i-r+2)\dots$$
$$\dots(i+1,i)(i,i+1)\dots(i-r+2,i+r-1)(i-r+1,i+r). \qquad \square$$

DEFINITION 3.2. Let $i \geqq 1$ and r, $0 \leqq r \leqq i-1$, be integers. The following are *squared products*:

$$\prod_{h=1}^{2r+1}(i+\min(0,r+1-h),i+\min(0,h-(r+1)))=$$
$$=(i,i-r)(i,i-r+1)\dots(i,i)\dots(i-r+1,i)(i-r,i),$$
$$\prod_{h=1}^{2r+1}(i+\max(0,r+1-h),i+\max(0,h-(r+1)))=$$
$$=(i+r,i)(i+r-1,i)\dots(i,i)\dots(i,i+r-1)(i,i+r). \qquad \square$$

DEFINITION 3.3. Let $i > 1$ and r, $0 \leqq r \leqq i-1$, be integers. The following are *pincers products*:

with one E-intersection of type 2.4(i)

$$\prod_{h=1}^{2r+1}(i+\min(h-(r+1)+2,(r+1)-h),i+\min(h-(r+1),(r+1)-h+2))=$$
$$=(i-r+2,i-r)(i-r+3,i-r+1)\dots(i,i-2)(i+1,i-1)(i,i)\cdot$$
$$\cdot(i-1,i+1)(i-2,i)\dots(i-r+1,i-r+3)(i-r,i-r+2),$$
$$\prod_{h=1}^{2r+1}(i+\max(h-(r+1)-2,(r+1)-h),i+\max(h-(r+1),(r+1)-h-2))=$$
$$=(i+r,i+r-2)(i+r-1,i+r-3)\dots(i+2,i)(i+1,i-1)(i,i)\cdot$$
$$\cdot(i-1,i+1)(i,i+2)\dots(i+r-3,i+r-1)(i+r-2,i+r);$$

with one E-intersection of type 2.4(ii)

$$\prod_{h=1}^{2r}(i+\min(h-(r+1)+2,(r+1)-h),i+\min(h-r,r-h+2))=$$
$$=(i-r+2,i-r+1)(i-r+3,i-r+2)\dots(i,i-1)(i+1,i)\cdot$$
$$\cdot(i,i+1)(i-1,i)\dots(i-r+2,i-r+3)(i-r+1,i-r+2),$$
$$\prod_{h=1}^{2r}(i+\max(h-r-2,r-h),i+\max(h-(r+1),(r+1)-h-2))=$$
$$=(i+r-1,i+r-2)(i+r-2,i+r-3)\dots(i+1,i)(i,i-1)\cdot$$
$$\cdot(i-1,i)(i,i+1)\dots(i+r-3,i+r-2)(i+r-2,i+r-1). \qquad \square$$

It is immediately seen that antidiagonal, squared and pincers products fulfil the hypotheses of Theorem 2.9. By applying that Theorem, one gets all the following results.

The first one generalizes Corollary 1 of [5].

PROPOSITION 3.4. *Every antidiagonal, squared and pincers product with* $i_1 < j_1$ *is equal to the idempotent* (i_1, i_1). $\square$

PROPOSITION 3.5. *Let* $P_n = a_1a_2\dots a_n$ *be an antidiagonal product with* $i_1 > j_1$.

If $n = 2r+1$ *and* $a_{r+1} = (i,i)$, *then the product is equal to the idempotent* $(i+r(r+1), i+r(r+1))$.

If $n = 2r$ *and* $a_r = (i+1,i)$, *then the product is equal to the idempotent* $(i+r^2, i+r^2)$.

PROOF. Suppose that the E-intersection of P_n is of type 2.4(i). For every h, $1 \leqq h \leqq r$, one has $a_h = (i+r-(h-1), i-r+(h-1))$. Hence $i_h - j_h = 2r - 2(h-1)$. Thus the product P_n is equal to the idempotent

$$(j_{r+1}+\sum_{h=1}^{r}(i_h-j_h), j_{r+1}+\sum_{h=1}^{r}(i_h-j_h))=$$
$$=(i+\sum_{h=1}^{r}(2r+2-2h), i+\sum_{h=1}^{r}(2r+2-2h))=$$
$$=(i+r(2r+2)-2r(r+1)/2, i+r(2r+2)-2r(r+1)/2)=$$
$$=(i+r^2+r, i+r^2+r)=$$
$$=(i+r(r+1), i+r(r+1)).$$

In the other case, for every h, $1 \leqq h \leqq r$, one has $a_h = (i+r-(h-1), i-r+h)$. Hence $i_h - j_h = 2r - 2h + 1 = 2r - 2(h-1) - 1$.

An easy comparison between the former case and the present one shows that now the idempotent P_n should be equal to

$$(i+r^2, i+r^2). \qquad \square$$

PROPOSITION 3.6. *Let $P_n = a_1 a_2 \dots a_n$ be a squared product with $i_1 > j_1$. If $n = 2r+1$ and $a_{r+1} = (i,i)$, then the product is equal to the idempotent $(i + r(r+1)/2, i + r(r+1)/2)$.*

PROOF. For every h, $1 \leqq h \leqq r$, one has

$$a_h = (i + \min(0, r+1-h), i + \min(0, h-(r+1))) = (i, i+h-(r+1))$$

or

$$a_h = (i + \max(0, r+1-h), i + \max(0, h-(r+1))) = (i + (r+1) - h, i).$$

In both cases $i_h - j_h = r + 1 - h$. Thus the product P_n is equal to the idempotent:

$$\Big(j_{r+1} + \sum_{h=1}^{r}(i_h - j_h), j_{r+1} + \sum_{h=1}^{r}(i_h - j_h)\Big) =$$
$$= \Big(i + \sum_{h=1}^{r}(r-h+1), i + \sum_{h=1}^{r}(r-h+1)\Big) =$$
$$= (i + r(r+1) - r(r+1)/2, i + r(r+1) - r(r+1)/2) =$$
$$= (i + r(r+1)/2, i + r(r+1)/2). \qquad \square$$

PROPOSITION 3.7. *Let $P_n = a_1 a_2 \dots a_n$ be a pincers product with $i_1 > j_1$.*
If $n = 2r+1$ and $a_{r+1} = (i,i)$, then the product is equal to the idempotent $(i+2r, i+2r)$.
If $n = 2r$ and $a_r = (i+1, i)$, then the product is equal to the idempotent $(i+r, i+r)$.

PROOF. As in the proof of Proposition 3.6, if $n = 2r+1$, then for every h, $1 \leqq h \leqq r$, one has

$$a_h = (i + \min(h-(r+1)+2, (r+1)-h), i + \min(h-(r+1), (r+1)-h+2)) =$$
$$= (i+h-(r+1)+2, i+h-(r+1))$$

or

$$a_h = (i + \max(h-(r+1)-2, (r+1)-h), i + \max(h-(r+1), (r+1)-h-2)) =$$
$$= (i + (r+1) - h, i + (r+1) - h - 2).$$

In both cases $i_h - j_h = 2$. Thus the product P_n is equal to the idempotent

$$(j_{r+1} + \sum_{h=1}^{r}(i_h - j_h), j_{r+1} + \sum_{h=1}^{r}(i_h - j_h)) = (i + 2r, i + 2r).$$

If $n = 2r$, then $i_h - j_h = 1$ and the product is equal to

$$(i + r, i + r). \qquad \square$$

REFERENCES

[1] DESQ, R., Relations d'équivalence principales en théorie des demi-groupes, *Ann. Fac. Sci. Univ. Toulouse* (4) **27** (1963), 1–149. *MR* **31** #4844

[2] GERENTE, A., Éléments inversibles et croissants dans un demi-groupe, *C.R. Acad. Sci. Paris Sér. A-B* **274** (1972), A1775–A1778. *MR* **46** #5514

[3] LJAPIN, E. S., *Semigroups*, Gosudarstv. Izdat. Fiz.-Mat. Lit., Moscow, 1960 (in Russian). *MR* **22** #11054. Translations of Mathematical Monographs Vol. 3, American Mathematical Society, Providence, RI, 1974. *MR* **50** #4789

[4] MIGLIORINI, F., Some researches on semigroups with magnifying elements, *Period. Math. Hungar.* **1** (1971), 279–286. *MR* **45** #423

[5] MIGLIORINI, F., Magnifying elements and minimal subsemigroups in semigroups, *Period. Math. Hungar.* **5** (1974), 279-288. *MR* **51** #762

[6] SHLEIFER, F. G., Looking for identities on a bicyclic semigroup with computer assistance, *Semigroup Forum* **41** (1990), 173–179. *MR* **91d**:20073

(Received February 28, 1991)

DIPARTIMENTO DI MATEMATICA
UNIVERSITÀ DI SIENA
VIA DEL CAPITANO, 15
I–53100 SIENA
ITALY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 279–294

ARITHMETICS OF AGING DISTRIBUTIONS: CONVOLUTION

T. F. MÓRI[1]

Abstract

Arithmetical properties of the convolution semigroup structure of certain aging classes of probability distributions on the nonnegative real halfline are discussed. The results include theorems on decompositions, on the density of the set of irreducible distributions and the non-existence of primes.

1. Introduction

1.1. Objective. Many fundamental theorems in classical probability theory are connected with composition or decomposition-like problems, starting with the first steps taken by de Moivre, Laplace, Poisson, Gauss and Cauchy towards a systematic study of summing independent random variables, and flourishing with the celebrated results of Kolmogorov, Lévy and Khinchin on infinitely divisible distributions as limit distributions for triangular arrays of small independent random variables.

This classical theory can be generalized in two possible directions. Firstly, one can consider probability distributions on more general structures, which, however, are special enough to assure that a convolution type operation can be defined; then an attempt can be made to extend classical results (e.g., the central limit theory for probability distributions on the Borel sets of locally compact Abelian groups, see [5]). Secondly, the convolution semigroup of probability distributions can be replaced with more general semigroups; then an appealing objective is to investigate which notions can still be defined, which results can still be proved in this general setting, mainly by algebraic tools. The successful and popular theory of Delphic semigroups, initiated by Kendall [6], can serve as a well-known example. A recent monograph of Ruzsa and Székely [11] is devoted to such an algebraic theory of

[1]This paper was written while the author was visiting the Mathematical Institute of the Hungarian Academy of Sciences.

1991 *Mathematics Subject Classification.* Primary 60E05; Secondary 62E10, 20M14.
Key words and phrases. Hun semigroups, decompositions, irreducible elements, prime elements, infinite divisibility, reliability, aging, IFR, IFRA, NBU, NBUE, HNBUE, L.

0081–6906/94/$ 4.00

probability. It turns out that many purely stochastic results are just special cases of general algebraic-arithmetical theorems valid for a large class of topological semigroups.

Of course, there always are special cases where general results appear too general to be efficient and sometimes ad hoc methods are simpler and still fruitful. Such an example is set by the aging classes of distributions, and we aim at studying their arithmetical structure under certain reliability operations.

1.2. Aging distributions. Aging distributions play a central role in reliability theory. Several different concepts of aging are defined and widely used. Let us recollect some of these aging classes of distributions with finite expectation (see [2]). In order to do this we need the following notations.

Since reliability theory deals, roughly speaking, with random lifetimes, we shall concentrate on D^+, the set of probability distributions on the nonnegative real halfline. Distributions will be identified with their cumulative distribution functions defined right continuous. For arbitrary distribution $F \in D^+$ let us introduce

- $\overline{F} = 1 - F$, the corresponding *survival function*,
- $\mathsf{E}(F) = \int_0^\infty \overline{F}(t)\,dt$, the *expectation* of F,
- $\mathsf{Var}(F) = \int_0^\infty 2t\overline{F}(t)\,dt - \mathsf{E}^2(F)$, the *variance* of F,
- $\varphi_F(t) = 1 - t\int_0^\infty \overline{F}(x)e^{-tx}\,dx$, $t \geqq 0$, the *Laplace transform* of F.

Separate notations are introduced for two scale-parameter families of distributions playing distinguished roles in reliability theory. One of them is the family of exponential distributions of the form $\mathcal{E}_\mu(t) = 1 - \exp(-t/\mu)$, $t \geqq 0$, $\mu > 0$. This distribution has expectation μ, variance μ^2 and Laplace transform $\dfrac{1}{1+\mu t}$, $t \geqq 0$. The other one is the family of degenerate distributions δ_μ, $\mu \geqq 0$. Clearly , $\delta_\mu(t) = 0$ or 1 according that $t < \mu$ or $t \geqq \mu$, resp., δ_μ has expectation μ, variance 0 and Laplace transform $\exp(-\mu t)$, $t \geqq 0$.

The most frequently used classes of aging distributions are as follows.

$f \in$ **IFR** iff for every $s > 0$ the function $t \mapsto \overline{F}(t+s)/\overline{F}(t)$, $t \geqq 0$ is decreasing.

$F \in$ **IFRA** iff the function $t \mapsto \overline{F}(t)^{1/t}$, $t > 0$ is decreasing.

$F \in$ **NBU** iff $\overline{F}(t+s) \leqq \overline{F}(t)\overline{F}(s)$ for every nonnegative t and s.

$F \in$ **NBUE** iff $\mathsf{E}(F) = \mu$ is finite and $\int_t^\infty \overline{F}(u)\,du \leqq \mu\overline{F}(t)$ for $t \geqq 0$.

$F \in$ **HNBUE** iff $\mathsf{E}(F) = \mu$ is finite and $\int_t^\infty \overline{F}(u)\,du \leqq \mu\exp(-t/\mu)$ for $t \geqq 0$.

$F \in \mathbf{L}$ iff $\mathbf{E}(F) = \mu$ is finite and $\varphi_F(t) \leqq \dfrac{1}{1+\mu t}$, $t \geqq 0$.

It is well-known that these classes form an increasing sequence in the order of definition: **IFR** $\subset$ **IFRA** $\subset$ **NBU** $\subset$ **NBUE** $\subset$ **HNBUE** $\subset$ **L**. Properties of the first four classes are found in [2]. Classes **HNBUE** and **L** were first introduced and studied by Rolski [10] and Klefsjö [8], resp. Both exponential and degenerate distributions belong to the smallest class **IFR**. In addition, exponential distributions lie on the common boundary of the above aging classes, since they satisfy each inequality-type definition with equality and show constant rate where monotonicity is required.

Furnishing D^+ with the usual topology of convergence in distribution (i.e., pointwise convergence at the continuity points of the limit distribution function) we find the above classes topologically closed.

1.3. Arithmetical definitions. The next step is to define a binary operation on D^+, which, in addition to being meaningful in reliability theory, makes some of these classes semigroups. We have three candidates, each of them is commutative, associative and continuous with respect to the weak topology. This makes it possible to extend the operation to an infinite sequence of distributions as the weak limit of the finite sections.

(a) Taking the *minimum* of two independent random variables. This means the pointwise multiplication of the corresponding survival functions: $F \wedge G = 1 - \overline{F}\,\overline{G}$. Then classes **IFR**, **IFRA** and **NBU** are algebraically closed, i.e., they are subsemigroups of D^+, while the other three aging properties are not preserved under this operation. Unfortunately, the min-structure of these subsemigroups is not too interesting, the arithmetic properties of distributions are rather trivial, e.g. every element is infinitely divisible.

(b) Taking the *maximum* of two independent random variables. This means the pointwise multiplication of the corresponding distribution functions: $F \vee G = FG$. Then classes **IFRA**, **NBU** and **NBUE** become subsemigroups of D^+, thus they serve as subject for further investigations. Counterexamples show that neither **IFR**, nor **HNBUE** are closed under the operation $\vee$. So far I have been unable to decide if **L** is closed or to find any reference on the subject. The max-arithmetical structure of the whole D^+ is not very exciting, since every nonnegative probability distribution appears max-infinitely divisible. The situation is quite different in higher dimensions (see [1] or [12]) or in subsemigroups of D^+. Arithmetical properties of the above three semigroups are planned to be studied in a forthcoming paper.

(c) Summation of independent random variables, that is, *convolution* of the corresponding distribution functions $F * G(t) = \int_0^t F(t-u)\,G(du)$, where the domain of integration is closed. This choice is very attractive for the following reasons. This seems to be the closest to the classical theory. All mentioned classes are closed under convolution. In addition, convolution is *cancellative* in D^+, i.e., $F*H=G*H$ implies $F=G$, because $*$ corresponds to the pointwise multiplication of (non-zero) Laplace transforms. This makes arithmetical investigations easier. Therefore, the objective of the present paper is to study the convolution structure of the above six reliability semigroups.

In the arithmetic of D^+ the unity is δ_0, the point mass at 0. Now, let S be an arbitrary subsemigroup of D^+ containing δ_0. For F and G belonging to S let us introduce the following arithmetical notions.

G is a *divisor* (or a *factor*) of F if there exists an H in S such that $F=G*H$. We use the notation $G\,|\,F$. A pair of elements in an arbitrary semigroup mutually dividing each other are called *associates*. Since $G\,|\,F$ implies $F\leqq G$ pointwise, it follows that D^+ is associate-free.

F is *irreducible* if $F\neq\delta_0$ and it has no divisor but the unity and itself.

F is *anti-irreducible* if it has no irreducible divisors.

F is *infinitesimally divisible* if it can be decomposed into a convolution of distributions all coming from an arbitrarily preassigned neighbourhood of δ_0. F is *infinitely divisible* if for every positive integer n there exists an $F_n\in S$ such that $F=F_n^{*n}$ (F_n is called the nth *root* of F).

F is *prime* if it is different from δ_0 and $F\,|\,G*H$ implies $F\,|\,G$ or $F\,|\,H$.

1.4. Acknowledgement. This research work was done in 1986–87, partly with G. J. Székely, especially in the beginning. Some of the statements below are to be considered joint results, namely, those of Section 2.1. These are also included in [11, Section 5.9] in a slightly different form. Although further results (not joint), which constitute the major part of the present paper, were also reported there in Remark 5.9.9, their proofs have never been published (apart from talks at conferences and other meetings, cf. [9]), so I finally decided to write this paper. Hereby I wish to thank G. J. Székely for all fruitful discussions on the topic.

2. Main results

2.1. Decompositions. In this section three lemmas are first presented which will be needed for the main results. Proofs are omitted since they can be found in [11, Section 5.9]. In order that the general theory be applicable to our reliability semigroups, we only need the following simple observation.

LEMMA 1. **IFR**, **IFRA**, **NBU**, **NBUE**, **HNBUE** *and* **L** *are stable normable Hun semigroups* (see Definitions 2.2.2, 2.10.6 and 2.15.2 in [11]).

PROOF. They all are closed subsemigroups of D^+, which is known to have these properties.

LEMMA 2. $\mathsf{Var}(F)\leqq\mathsf{E}(F)^2$ *for every* $F\in\mathbf{L}$.

PROOF. With $\mathsf{E}(F)=\mu$ we have

$$\mathsf{Var}(F) + \mu^2 = \int_0^\infty t^2\,dF(t) = \lim_{t\downarrow 0} 2t^{-2}(\varphi_F(T) - 1 + \mu t) \leqq \lim_{t\downarrow 0} \frac{2\mu^2}{1+\mu t} = 2\mu^2.$$

LEMMA 3. *Expectation is a continuous operator on* **L**.

PROOF. See [11], Remark 5.9.8.

THEOREM 1. *In the above six semigroups for an arbitrary distribution F the following three properties are equivalent.*
F is anti-irreducible,
F is infinitely divisible,
F is degenerate.

PROOF. The proof of Theorem 5.9.1 of [11] can be repeated without any changes. Here the general theory works well, the only thing to be added is that every infinitesimally divisible distribution F is degenerate, since for arbitrary decomposition $F = F_1 * \cdots * F_n$ Lemma 2 implies

$$\mathsf{Var}(F) = \sum_{i=1}^n \mathsf{Var}(F_i) \leqq \max \mathsf{E}(F_i) \sum_{i=1}^n \mathsf{E}(F_i) = \mathsf{E}(F) \max \mathsf{E}(F_i),$$

and here the right-hand side can be arbitrarily small by Lemma 3.

THEOREM 2. *In the above six semigroups every distribution can be decomposed into the convolution of at most countable many irreducible distributions and a degenerate one.*

PROOF. This is a simple corollary of our Theorem 1 and Theorem 2.23.3 of [11].

2.2. Irreducible distributions. Though it seems hopeless to characterize irreducible distributions in either of our semigroups, there exists a simple sufficient condition for a distribution to be irreducible even in the largest class **L**, which enables us to show that irreducible distributions are dense in each of the above classes.

LEMMA 4. *Let $F \in$ **L** and suppose that* $\limsup_{x\downarrow 0} x^{-2}F(x) = +\infty$. *Then F is irreducible.*

PROOF. Since

$$\varphi_F(t) = \int_0^\infty e^{-tx}\,dF(x) \geqq e^{-1}F(t^{-1}),$$

we obtain

$$\limsup_{t\to\infty} t^2\varphi_F(t) = +\infty.$$

On the other hand, if both F_1 and F_2 belong to **L**, $\varphi_{F_1*F_2}(t) = \varphi_{F_1}(T)\varphi_{F_2}(t) = \mathcal{O}(t^{-2})$.

THEOREM 3. *The set of irreducible elements is dense in* **IFR**, **IFRA**, **NBU**, **NBUE**, **HNBUE** *and* **L**.

PROOF. First, for **IFR**, **IFRA**, **NBU**. These classes are closed under the minimum operation $\wedge$. Hence, if $F \in S$ where S is any of these classes, then $F \wedge \varepsilon_\mu \in S$ for every $\mu > 0$, it is irreducible by Lemma 4, and $\lim_{\mu\to\infty} F \wedge \wedge \varepsilon_\mu = F$.

Now, let S be any of **NBUE**, **HNBUE** and **L**. Then mixtures of distributions from S also belong to S provided the elements to be mixed have the same expectation. Let $F \in S$ be arbitrary with $\mathbf{E}(F) = \mu$, then for every p, $0 < p < 1$, the mixture $pF + (1-p)\varepsilon_\mu$ belongs to S, it is irreducible by Lemma 4, and finally, it converges weakly to F as $p \to 1$.

REMARK 1. The set of irreducible elements is G_δ in each class, by Theorem 2.19.2 of [11].

REMARK 2. The above proof of our Theorem 3 shows that the subset of elements irreducible even in **L** is dense in each class. Since in the chain **IFR** $\subset$ **IFRA** $\subset$ **NBU** $\subset$ **NBUE** $\subset$ **HNBUE** $\subset$ **L** each inclusion is strict, and these classes are closed, it follows that each of the last five classes contain irreducible elements not belonging to any of the smaller classes. On the other hand, in each of the first five classes there exist elements that are irreducible there but not in the subsequent classes (see Remark 5 in Section 2.4).

2.3. Prime distributions. Perhaps the most interesting problem in the arithmetics of a given semigroup is the existence or non-existence of primes. Ruzsa and Székely proved the non-existence of primes in D, the convolution semigroup of probability distributions on the real line [11, Section 4.4]. They also gave the complete description of prime elements in $D(G)$, the convolution semigroup of probability distributions defined on the Borel field of the locally compact Hausdorff group G. It turned out that, apart from a finite number of exceptions, all on small cyclic groups, there were no primes in $D(G)$ [11, Section 4.7]. Similar results are known for some other specific semigroups. This raise the hope of the existence of a general theorem stating the non-existence of primes in a large class of semigroups having "sufficiently rich" structure, but there is not even a reasonable conjecture on what is to be meant under that.

We are able to prove the lack of primes in **NBU** and the larger classes, but not in **IFR** and **IFRA**.

THEOREM 4. *There are no primes in* **NBU**, **NBUE**, **HNBUE** *and* **L**.

The proofs follow a common pattern in all cases. Let S be an arbitrary closed subsemigroup of D^+ containing all degenerate distributions. Consider the set $T(S)$ of all distributions $F \in D^+$ that can be translated into S, that is, which can be convolved with a degenerate distribution so that the result belong to S. $T(S)$ is called the *translation hull* of S. In formulas,

$$T(S) = \{F \in D^+ : F * \delta_c \in S \text{ for some } c \geqq 0\}.$$

It is easy to see that $S \subset T(S)$ and $T(S)$ is a subsemigroup of D^+. For arbitrary $F \in T(S)$ let us define

$$c_S(F) = \inf\{c \geqq 0 : F * \delta_c \in S\},$$

the *translation distance* of F from S (subscript $_S$ will be supressed when possible). Since S is closed, $F * \delta_{c(F)} \in S$, thus infimum can be replaced with minimum. Clearly, $c(F) = 0$ iff $F \in S$ and c is subadditive on $T(S)$: $c(F * G) \leqq c(F) + c(G)$. Here equality holds if at least one of F, G is equal to δ_0 or both F and G are in S; these cases will be referred to as *trivial*.

Consider now an arbitrary $F \in S$, $F \neq \delta_0$. We wish to show that F cannot be a prime. This is immediate if F is degenerate, for degenerate distributions are infinitely divisible. If F is not degenerate, the following simple lemma applies.

LEMMA 5. *Suppose there exist a $G \in T(S)$, $G \notin S$ such that $c(F * G) < c(G)$. Then F is not prime.*

PROOF. With the notation $a = c(F * G)$ and $b = c(G)$ we have $F \mid F * (G * \delta_b)$. Here $F * (G * \delta_b) = (F * G * \delta_a) * \delta_{b-a}$, another S-decomposition, but neither $F * G * \delta_a$ nor δ_{b-a} is divisible by F.

The proof of Theorem 4 will be performed for the four reliability semigroups separately. Since the description of the translation hulls may be of independent interest, we present it whenever we can, even if it is not always necessary for the proof.

LEMMA 6. *The translation hull of* **L** *is characterized as follows.*

$$T(\mathbf{L}) = \{F \in D^+ : \mathsf{Var}(F) < \infty\}, \tag{1}$$

and c is strictly subadditive on it, i.e., equality cannot hold apart from trivial cases.

PROOF. Let us first deal with (1). Since elements of **L** are of finite variance, the same holds for the translation hull. On the other hand, let $F \in D^+$ arbitrary with finite variance σ^2 and expectation μ. We have to find a positive c for which $\delta_c * F \in \mathbf{L}$, that is,

$$\varphi_F(t) \leqq \frac{1}{1 + (\mu + c)t} e^{ct} \tag{2}$$

for every $t>0$. By Lemma 2, $\sigma^2 \leqq (\mu+c)^2$, that is, $c \geqq \sigma-\mu$. Let c be a fixed positive number, such that $c>\sigma-\mu$. Then

$$\varphi_F(t)=1-\mu t+\frac{1}{2}(\sigma^2+\mu^2)t^2+o(t^2),$$

while

$$e^{ct}(1+(\mu+c)t)^{-1}=1-\mu t+\frac{1}{2}((\mu+c)^2+\mu^2)t^2+o(t^2)$$

as $t\to 0$. Hence (2) holds for $t \leqq t_0$. On the other hand,

$$\lim_{t\to 0} e^{ct}(1+(\mu+c)t)^{-1}=+\infty,$$

thus (2) is satisfied for $t \geqq t_1$, too. For fixed t the right-hand side of (2) tends increasingly to infinity as $c\to\infty$, hence (2) holds for every positive t provided c is large enough, completing the proof of (1). Moreover, it has turned out that if $c(F)>\sigma-\mu$ and $c(F)>0$, then there exists a positive t at which

$$\varphi_F(t)=\frac{1}{1+(\mu+c(F))t}e^{c(F)t}. \tag{3}$$

Let us turn to the proof of strict subadditivity of c. Suppose the contrary, that is, let F and G be distributions with finite variances for which $c(F*G)=$ $=c(F)+c(G)$ holds non-trivially. We can suppose that neither of them is degenerate, for if $G=\delta_a$, $a>0$ and $F\notin \mathbf{L}$, then $c(F*G)=(c(F)-a)^+<c(F)$. Clearly,

$$\sqrt{\mathbf{Var}(F*G)}-\mathbf{E}(F*G)<\sqrt{\mathbf{Var}(F)}-\mathbf{E}(F)+\sqrt{\mathbf{Var}(G)}-\mathbf{E}(G)\leqq c(F)+c(G),$$

hence by (3) there exists a positive t such that

$$\varphi_F(t)\varphi_G(t)=\frac{\exp((c(F)+c(G))t)}{1+(\mathbf{E}(F)+\mathbf{E}(G)+c(F)+c(G))t}>$$
$$>\frac{\exp(c(F)t)}{1+(\mathbf{E}(F)+c(F))t}\,\frac{\exp(c(G)t)}{1+(\mathbf{E}(G)+c(G))t},$$

contradicting the definition of $c(F)$ and $c(G)$.

LEMMA 7. *The translation hull of* **HNBUE** *is characterized as follows:*

$$T(\mathbf{HNBUE})=\{F\in D^+ : \exists C_1,C_2>0, \overline{F}(x)\leqq C_1\exp(-C_2x)\}. \tag{4}$$

Besides, c is strictly subadditive on T(**HNBUE**), *apart from trivial cases.*

PROOF. Let us first observe that $F*\delta_c\in\mathbf{HNBUE}$, with $\mathbf{E}(F)=\mu$ and $c\geqq 0$, is equivalent to the inequality

$$\int_x^\infty \overline{F}(t)\,dt=\int_{x+c}^\infty \overline{F*\delta_c}(t)\,dt\leqq(\mu+c)\exp\left(-\frac{x+c}{\mu+c}\right) \tag{5}$$

to be satisfied for every positive x. In fact, (5) is needed for every $x \geqq -c$ but for negative values of x this is obvious, since then the left-hand side equals $\mu - x$, and

$$\mu - x = (\mu + c)\left(1 - \frac{x+c}{\mu+c}\right) \leqq (\mu + c)\exp\left(-\frac{x+c}{\mu+c}\right).$$

Let ξ be a random variable with distribution F. Then the left-hand side of (5) is just $\mathbf{E}((\xi - x)^+)$. This shows that in (5) equality can only hold if either $x = 0$ and $c = 0$ or x lies in the interior of the range of F: $0 < F(x-0) \leqq \leqq F(x) < 1$.

After these preliminaries, let first $F \in T(\mathbf{HNBUE})$ with $\mathbf{E}(F) = \mu$ and $c(F) = c$, then we have

$$(6) \quad \overline{F}(x) \leqq (\mu + c)^{-1} \int_{x-\mu-c}^{\infty} \overline{F}(t)\,dt \leqq \exp\left(-\frac{x-\mu}{\mu+c}\right) = C_1 \exp(-C_2 x).$$

On the other hand, assuming $\overline{F}(x) \leqq C_1 \exp(-C_2 x)$ we obtain

$$\int_x^{\infty} \overline{F}(t)\,dt \leqq \frac{C_1}{C_2} \exp(-C_2 x);$$

from which $\mathbf{E}(F) = \mu < \infty$, and (5) is satisfied for every positive x, provided c is large enough ($c = (1 + eC_1)/C_2$ will do). Consequently, $F * \delta_c \in \mathbf{HNBUE}$.

In order to prove the strict subadditivity of $c(\cdot)$, suppose in contrary that F and G satisfy $c(F * G) = c(F) + c(G)$ non-trivially. Then by (6),

$$\overline{F}(x) = \overline{F * \delta_{c(F)}}(x + c(F)) \leqq C_1 \overline{\varepsilon}_{\mathbf{E}(F)+c(F)}(x),$$
$$\overline{G}(x) \leqq C_2 \overline{\varepsilon}_{\mathbf{E}(G)+c(G)}(x),$$

where $C_1, C_2 > 1$, hence by the bilinearity and monotonicity of convolution we have

$$\overline{F * G}(x) \leqq C_1 C_2 \overline{\varepsilon_{\mathbf{E}(F)+c(F)} * \varepsilon_{\mathbf{E}(G)+c(G)}}(x) = \exp(-C_3^{-1}(1 + o(1))x),$$

with $C_3 = \max\{\mathbf{E}(F) + c(F), \mathbf{E}(G) + c(G)\} < \mathbf{E}(F) + \mathbf{E}(G) + c(F) + c(G)$. Consequently,

$$\int_x^{\infty} \overline{F * G}(t)\,dt \leqq \exp(-C_3^{-1}(1 + o(1))x),$$

which shows that it is not the tail of $F*G$ that does not allow $c(F*G)$ to be decreased. Thus, there exists a positive x for which

$$\int_x^\infty \overline{F*G}(t)\,dt=$$
$$=(\mathbf{E}(F)+\mathbf{E}(G)+c(F)+c(G))\exp\left(-\frac{x+c(F)+c(G)}{\mathbf{E}(F)+\mathbf{E}(G)+c(F)+c(G)}\right).$$

Let us decompose x in the form $x=x_1+x_2$, where

$$x_1=\frac{(\mathbf{E}(F)+c(F))x+\kappa}{\mathbf{E}(F)+\mathbf{E}(G)+c(F)+c(G)},\qquad x_2=\frac{(\mathbf{E}(G)+c(G))x-\kappa}{\mathbf{E}(F)+\mathbf{E}(G)+c(F)+c(G)},$$
$$\kappa=\mathbf{E}(F)c(G)-\mathbf{E}(G)c(F).$$

Let ξ and η be independent random variables with distribution F and G, resp. Then

$$\int_x^\infty \overline{F*G}(t)\,dt=\mathbf{E}((\xi+\eta-x_1-x_2)^+)\leqq\mathbf{E}((\xi-x_1)^+)+\mathbf{E}((\eta-x_2)^+)\leqq$$
$$\leqq(\mathbf{E}(F)+c(F))\exp\left(-\frac{x_1+c(F)}{\mathbf{E}(F)+c(F)}\right)+$$
$$+(\mathbf{E}(G)+c(G))\exp\left(-\frac{x_2+c(G)}{\mathbf{E}(G)+c(G)}\right)=$$
$$=(\mathbf{E}(F)+\mathbf{E}(G)+c(F)+c(G))\exp\left(-\frac{x+c(F)+c(G)}{\mathbf{E}(F)+\mathbf{E}(G)+c(F)+c(G)}\right).$$

Here both inequalities must hold with equality, but this leads to contradiction, since equality in the first inequality requires either $\xi\geqq x_1$, $\eta\geqq x_2$ or $\xi\leqq x_1$, $\eta\leqq x_2$ with probability 1, then the second inequality could only be satisfied when $c(F)=c(G)=x_1=x_2=0$. Hence the indirect hypothesis is disproved.

LEMMA 8. *The following four assertions are equivalent.*
$F\in T(\mathbf{NBUE})$,

$\limsup_{x\to\infty}\overline{F}(x)^{-1}\int_x^\infty\overline{F}(t)\,dt<+\infty$ *(here* $0/0$ *meant* 0*)*,

there exist constants $C_1>0$, $0<C_2<1$ *such that* $\overline{F}(x+C_1)\leqq C_2\overline{F}(x)$, $x\geqq 0$,

there exist constants $C_1,C_2>0$, *such that* $\overline{F}(x+y)\leqq C_1\exp(-C_2y)\overline{F}(x)$, $x,y\geqq 0$.

Besides, $c(F) = \sup_{x \geqq 0} \overline{F}(x)^{-1} \int_x^\infty \overline{F}(t)\,dt - \mathsf{E}(F)$, *and this function is strictly subadditive on* $T(\mathbf{NBUE})$, *apart from trivial cases.*

PROOF. Indeed, the above formula for $c(F)$ is obvious, together with the equivalence of the first two assertions. Now suppose that for every positive x we have

$$\int_x^\infty \overline{F}(t)\,dt \leqq C\overline{F}(x).$$

Then

$$\overline{F}(x+2C) \leqq \frac{1}{2C} \int_x^\infty \overline{F}(t)\,dt \leqq 0.5\overline{F}(x).$$

Suppose $\overline{F}(x+a) \leqq b\overline{F}(x)$, then

$$\overline{F}(x+y) \leqq \overline{F}(x+[y/a]a) \leqq b^{[y/a]}\overline{F}(x) \leqq C_1 \exp(-C_2 y)\overline{F}(x),$$

where $C_1 = 1/b$, $C_2 = -a^{-1}\log b$.

Finally, if $\overline{F}(x+y) \leqq C_1 \exp(-C_2 y)\overline{F}(x)$, $x, y \geqq 0$, then

$$\int_x^\infty \overline{F}(t)\,dt = \int_0^\infty \overline{F}(x+t)\,dt \leqq \frac{C_1}{C_2}\overline{F}(x),$$

thus $F \in T(\mathbf{NBUE})$. The proof of the stated equivalence is complete.

Let us turn to the proof of strict subadditivity. We show that

(7) $$c(F*G) \leqq \min\{\max\{c(F), c(G) - \mathsf{E}(F)\}, \max\{c(G), c(F) - \mathsf{E}(G)\}\},$$

which is clearly less than $c(F) + c(G)$ in the non-trivial cases.

Since

$$\overline{F*G}(x) = 1 - \int_0^x F(x-u)\,dG(u) = \overline{G}(x) + \int_0^x \overline{F}(x-u)\,dG(u),$$

integrating this we obtain

$$\int_x^\infty \overline{F*G}(t)\,dt = \int_x^\infty \overline{G}(t)\,dt + \int_x^\infty \int_0^t \overline{F}(t-u)\,dG(u)\,dt.$$

The first term in the right-hand side can be estimated by $(\mathbf{E}(G)+c(G))\overline{G}(x)$. The second term can be treated as follows.

$$\int_x^\infty\int_0^t \overline{F}(t-u)\,dG(u)\,dt=\int_x^\infty\int_0^x \overline{F}(t-u)\,dG(u)\,dt+\int_x^\infty\int_x^t \overline{F}(t-u)\,dG(u)\,dt=$$
$$=\int_0^x\int_x^\infty \overline{F}(t-u)\,dt\,dG(u)+\int_x^\infty\int_u^\infty \overline{F}(t-u)\,dt\,dG(u)\leqq$$
$$\leqq\int_0^x(\mathbf{E}(F)+c(F))\overline{F}(x-u)\,dG(u)+\mathbf{E}(F)\overline{G}(x).$$

Thus

$$\int_x^\infty \overline{F*G}(t)\,dt\leqq(\mathbf{E}(F)+\mathbf{E}(G)+c)\left(\overline{G}(x)+\int_0^x\overline{F}(x-u)\,dG(u)\right)=$$
$$=(\mathbf{E}(F)+\mathbf{E}(G)+c)\overline{F*G}(x),$$

where $c=\max\{c(G),c(F)-\mathbf{E}(G)\}$. The estimation now follows by symmetry.

PROOF OF THEOREM 4. Lemmas 6,7 and 8 combined with Lemma 5 immediately imply the lack of primes in **L**, **HNBUE** and **NBUE**, resp. The case of **NBU** can be treated as follows.

Let $F\in\mathbf{NBU}$ arbitrary non-degenerate distribution. Introduce $G=\sum_{i=0}^{\infty}p(1-p)^iF^{*i}$ and $H=F*G$, where $0<p<1$ and F^{*i} denotes the ith convolution power of F $(F^{*0}=\delta_0)$. H is the distribution of the *geometric convolution* $\eta=\sum_{i=1}^{N}\xi_i$, where $\xi_1,\xi_2,\dots$ are i.i.d. random variables with distribution F, N is a random variable, independent of the summands and geometrically distributed with parameter p. The geometric convolution is known to preserve several aging properties of F, e.g., $F\in\mathbf{NBU}$ implies that $H\in\mathbf{NBU}$ [4]. Clearly, $G\notin\mathbf{NBU}$, since G puts positive weight on 0 (such a distribution cannot even belong to **L**). Now we show that $G\in T(\mathbf{NBU})$. By Lemma 8, there exists a positive constant c such that $\overline{H}(x)<(1-p)\overline{H}(x-c)$ for every $x>c$. Since $\overline{G*\delta_c}(x)=\overline{G}(x-c)=(1-p)\overline{H}(x-c)$ for $x>c$, we have

$$\overline{G*\delta_c}(x+y)=(1-p)\overline{H}(x+y-c)\leqq(1-p)^2\overline{H}(x+y-2c)\leqq$$
$$\leqq(1-p)\overline{H}(x-c)(1-p)\overline{H}(y-c)=\overline{G*\delta_c}(x)\overline{G*\delta_c}(y),\qquad x,y>c;$$

and this trivially extends to arbitrary positive x and y. Thus, $F \mid F*(\delta_c*G) = H*\delta_c$, but F does not divide either H or δ_c.

2.4. Final remarks.

REMARK 3. In the case of **NBU** the following results may be of some interest.

$$\{F \in D^+ : \exists C \geqq 1, \overline{F}(x+y) \leqq C\overline{F}(x)\overline{F}(y),\ \ x,y \geqq 0\} \subset T(\mathbf{NBU}) \tag{8}$$

$$c(F*G) \leqq \max\{c(F), c(G)\}. \tag{9}$$

In order to prove this assertion, let F satisfy $\overline{F}(x+y) \leqq C\overline{F}(x)\overline{F}(y)$, $x,y \geqq 0$, with some positive constant C. Similarly to the proof of Lemma 8 it can be shown that

$$\overline{F}(x+y+z) \leqq C_1 \exp(-C_2 z)\overline{F}(x)\overline{F}(y), \quad x,y,z \geqq 0,$$

with positive constants C_1, C_2. This immediately implies that

$$\overline{F}(x+y+C) \leqq \overline{F}(x)\overline{F}(y)$$

for every $x,y \geqq 0$, if C is large enough, thus (8) is proved. Note that bounded distributions obviously satisfy the condition in (8) hence they belong to $T(\mathbf{NBU})$ (actually, $c(F) \leqq k$ if $F(k)=1$).

Let us turn to the proof of (9). Let us denote $\max\{c(F), c(G)\}$ by c and $F*G$ by H. Then

$$\overline{F}(x+y+c) \leqq \overline{F}(x)\overline{F}(y), \qquad \overline{G}(x+y+c) \leqq \overline{G}(x)\overline{G}(y). \tag{10}$$

Consequently,

$$\overline{H}(x+y+c) = \int_{0-}^{\infty} \overline{F}(x+y+c-t)\,dG(t) =$$

$$= \int_{0-}^{y+} \overline{F}(x+y-t+c)\,dG(t) + \int_{y+}^{x+y+c+} \overline{F}(x+y+c-t)\,dG(t) + \overline{G}(x+y+c) \leqq$$

$$\leqq \overline{F}(x)\int_{0-}^{y+} \overline{F}(y-t)\,dG(t) + \int_{0-}^{x+c-} \overline{F}(u)\,d_u\overline{G}(x+y+c-u) + \overline{G}(x+y+c).$$

Integrating by parts in the second term and using (10) we obtain

$$\overline{H}(x+y+c) \leqq \overline{F}(x)(\overline{H}(y)-\overline{G}(y)) + \left[\overline{F}(u)\overline{G}(x+y+c-u)\right]_{u=0-}^{x+c-} + \\ + \int_{0-}^{x+c-} \overline{G}(x+y+c-u)\,dF(u) + \overline{G}(x+y+c) \leqq$$

$$\leqq \overline{F}(x)(\overline{H}(y)-\overline{G}(y)) + \overline{F}(x+c-)\overline{G}(y) + \overline{G}(y) \int_{0-}^{x+c-} \overline{G}(x-u)\,dF(u) =$$

$$= \overline{F}(x)(\overline{H}(y)-\overline{G}(y)) + \overline{F}(x+c-)\overline{G}(y) + \overline{G}(y)(\overline{H}(x)-\overline{F}(x+c-)) = \\ = \overline{H}(x)\overline{H}(y) - (\overline{H}(x)-\overline{F}(x))(\overline{H}(y)-\overline{G}(y)) \leqq \\ \leqq \overline{H}(x)\overline{H}(y).$$

Thus $c(H) \leqq c$, as stated.

Unfortunately, (10) does not imply the strict subadditivity of c in the case we need it, that is, when $F \in \mathbf{NBU}$.

REMARK 4. It is no surprise that the cases of **IFR** and **IFRA** appear much harder to solve. This may be in connection with the experience that even the preservation of these aging properties under convolution is more difficult to prove; actually, for the class **IFRA** this was a long-standing conjecture till the 1976 paper [3]. Though there is still hope that our method of translation hulls can be adapted to the case of **IFRA**, it is sure to fail for the **IFR** class, since trivially $T(\mathbf{IFR}) = \mathbf{IFR}$. Below we give an alternative proof in the case of **L**, which might also be applied to **IFR**.

Let $\Gamma_{\mu,\alpha}$ denote the gamma distribution of order $\alpha \geqq 0$ and scale parameter $\mu > 0$, that is, with Laplace transform $\varphi(t) = (1+\mu t)^{-\alpha}$. Clearly, $\Gamma_{\mu,\alpha} \in \mathbf{L}$ iff $\alpha \geqq 1$. It is easy to see that the convolution of an arbitrary distribution $F \in \mathbf{L}$ with a gamma distribution of scale parameter $\mu = \mathbf{E}(F)$ still belongs to **L**, even if $\alpha < 1$. Indeed, $\mathbf{E}(F * \Gamma_{\mu,\alpha}) = (\alpha+1)\mu$, and

$$\varphi_{F*\Gamma}(t) = \varphi_F(t)(1+\mu t)^{-\alpha} \leqq (1+\mu t)^{-(\alpha+1)} \leqq (1+(\alpha+1)\mu t)^{-1}.$$

Now, let $F \in \mathbf{L}$, different from δ_0. Then F is not prime, since $F \mid F * \Gamma_{\mu,1.5}$; here $F * \Gamma_{\mu,1.5} = (F * \Gamma_{\mu,0.5}) * \Gamma_{\mu,1}$, but F does not divide $F * \Gamma_{\mu,0.5}$, and F cannot divide $\Gamma_{\mu,1}$ (which is just the irreducible ε_μ, cf. Lemma 4), unless $F = \Gamma_{\mu,1}$. In that case F divides $\Gamma_{\mu,3} = F * \Gamma_{\mu,2} = \Gamma_{\mu,1.5} * \Gamma_{\mu,1.5}$, but F does not divide $\Gamma_{\mu,1.5}$.

REMARK 5. Using the notion of translation hull one can easily find irreducible distributions in each of **IFRA**, **NBU**, **NBUE**, **HNBUE**, which are reducible in the subsequent larger classes. Apart from the **IFR** case, it is sufficient to find a distribution $F \in D^+$, which is effectively irreducible in D

(here *effectively* means that only decompositions of the form $F=(F*\delta_{-a})*\delta_a$ exist), and for which

$$c_{\mathbf{L}}(F)<c_{\mathbf{HNBUE}}(F)<c_{\mathbf{NBUE}}(F)<c_{\mathbf{NBU}}(F)<c_{\mathbf{IFRA}}(F)<\infty.$$

Namely, let S and S' be subsemigroups of D^+, $S\subset S'$, and suppose that the corresponding translation distances are not equal: $c'(F)<c(F)$. Then $F*\delta_{c(F)}$ is irreducible in S but not in S', for it can be decomposed as $(F*\delta_{c'(F)})*\delta_{c(F)-c'(F)}$. In our case let F be the mixture $F=\frac{1}{3}(\delta_0+\delta_1+U)$, where U is the uniform distribution on the interval $(0,1)$. F is irreducible in D for if F admits a non-trivial decomposition, then each component has to put positive weights on the infimum and supremum of its support, which is a contradiction, since F has exactly two atoms. (An interesting related result is due to L. S. Kudina, who proved that any closed subset of $\mathbf{R}$ can serve as the support of an irreducible distribution [11, p.161].) Now,

$$\begin{aligned} c_{\mathbf{IFRA}}(F)&=2\log 1.5\approx 0.811,\\ c_{\mathbf{NBU}}(F)&=2/3\approx 0.667,\\ c_{\mathbf{NBUE}}(F)&=1/4=0.25,\\ c_{\mathbf{HNBUE}}(F)&\approx 0.182,\\ c_{\mathbf{L}}(F)&\approx 0.097.\end{aligned}$$

Details of calculation are left to the reader.

For sake of completeness we show an example of a distribution irreducible in **IFR** but reducible in **IFRA**. Let $F=\frac{1}{2}(\delta_0+\delta_{1/2})$, then $F*U\in\mathbf{IFRA}$, $F*U\notin\mathbf{IFR}$ and $F*U*U\in\mathbf{IFR}$ (these are not so hard to see), hence $F*U*U$ is reducible in **IFRA**. At the same time $F*U*U$ is irreducible in **IFR** (this is more difficult).

REFERENCES

[1] Balkema, A. A. and Resnick, S. I., Max-infinite divisibility, *J. Appl. Probability* **14** (1977), 309–319. *MR* **55** #11338

[2] Barlow, R. E. and Proschan, F., *Statistical theory of reliability and life testing*, Holt, Rinehart & Winston, New York, 1975. *MR* **55** #11534

[3] Block, H. W. and Savits, T. H., The IFRA closure problem, *Ann. Probability* **4** (1976), 1030–1032. *MR* **54** #6431

[4] Block, H. W. and Savits, T. H., Shock models with NBUE survival, *J. Appl. Probab.* **15** (1978), 621–628. *MR* **58** #24808

[5] Heyer, H., *Probability measures on locally compact groups*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 94, Springer-Verlag, Berlin, 1977. *MR* **58** #18648

[6] Kendall, D. G., Delphic semigroups, infinite divisible regenerative phenomena, and the arithmetic of p-functions, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **9** (1968), 163–195. *MR* **37** #5320

[7] Klefsjö, B., The distribution class harmonic new better that used in expectation — some properties and tests, Research Report 1979–9, Department of Mathematical Statistics, University of Umeå, 1979.

[8] KLEFSJÖ, B., A useful ageing property based on the Laplace transform, *J. Appl. Probab.* **20** (1983), 615–626. *MR* **85a**:62021

[9] MÓRI, T. F., Max-arithmetic of aging distributions, *Fifth European Young Statisticians Meeting* (Århus, 1987), ed. by J. L. Jensen et al., Dept. Theor. Statist. Inst. Math. Århus Univ., 1987, 98–100.

[10] ROLSKI, T., Mean residual life, Proceedings of the 40th Session of the International Statistical Institute (Warsaw, 1975), Vol. 4, Contributed papers, *Bull. Inst. Internat. Statist.* **46** (1975), 266–270. *MR* **57** #1684

[11] RUZSA, I. Z. and SZÉKELY, G. J., *Algebraic probability theory*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, Wiley, New York, 1988. *MR* **89j**:60006

[12] ZEMPLÉNI, A., The description of the class I_0 in the multiplicative structure of distribution functions, *Mathematical statistics and probability theory* (Proc. of the 6th Pannonian Symposium on Math. Statist., Bad Tatzmannsdorf, 1986), M. L. Puri et al. (eds.), Vol. A, Reidel, Dordrecht, 1987, 291–303. *MR* **89b**: 60046

(Received March 4, 1991)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR
VALÓSZÍNŰSÉGELMÉLETI ÉS STATISZTIKAI TANSZÉK
MÚZEUM KRT. 6–8
H–1088 BUDAPEST
HUNGARY

e-mail: moritamas @ ludens.elte.hu

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 295–305

A THEOREM ON PERTURBED NEWTON-LIKE METHODS IN BANACH SPACES

K. ARGYROS

Abstract

The recent elegant error bounds for Newton-like methods in Banach spaces obtained by Yamamoto [14] have great theoretical value but little practical value since the iterates can rarely be composed exactly. Here we extend his results by considering perturbed iterative procedures to find error bounds on the distances between the computed and the exact iterates as well as the distances between the computed iterates and the exact solution.

§ 1. Introduction

Let F be a nonlinear operator mapping some subset D of a real Banach space E into a subset of a real Banach space $\hat{E}$. The most popular methods for approximating solutions x^* of the equation

$$F(x) = 0 \tag{1}$$

are the so-called Newton-like methods of the form

$$x_{n+1} = x_n - A(x_n)^{-1} F(x_n), \qquad n = 0, 1, 2, \ldots . \tag{2}$$

Here $x_0 \in D$ is given and $[A(x_n)]$, $n = 0, 1, 2, \ldots$ denotes a sequence of linear operators. In practice $A(x_n)$ should be a conscious approximation to the Fréchet-derivative $F'(x_n)$, since when $A(x_n) = F'(x_n)$, the iterative procedure (2) reduces to the Newton–Kantorovich method.

Yamamoto [14] has unified the study of finding sharp error bounds for Newton-like methods of the form (2) under Kantorovich type assumptions. He obtains results that improve error bounds obtained before by Rheinboldt [1], Dennis [2], Miel [7], Moret [8], Potra [9], et al.

The results obtained by the above authors, however, have great theoretical but little practical value, since the sequence generated by (2) can rarely be computed exactly.

In this paper we find it useful to consider that the iterative procedure (2) is perturbed. We suppose that all the elements contained in the construction

---

1980 *Mathematics Subject Classification* (1985 Revision). Primary 65J15, 65L50, 65M50; Secondary 47D15, 47H17.

Key words and phrases. Banach space, Newton-like method.

0081–6906/94/$ 4.00

of these procedures are known only approximately. Moreover, we suppose that at each step the solution of the respective linear system is also performed approximately.

In particular, we consider the iterative procedures corresponding to (2) to be of the form

$$(3)\quad \tilde{x}_{n+1} = \tilde{x}_n - (A(\tilde{x}_n) + L_n)^{-1}(F(\tilde{x}_n) + y_n) + z_n, \quad \tilde{x}_0 = x_0, \ n = 1, 2, \ldots$$

$$(4)\quad \tilde{x}_{n+1} = \tilde{x}_n - (F'(x_0) + L_0)^{-1}(F(\tilde{x}_n) + y_n) + z_n, \quad \tilde{x}_0 = x_0, \ n = 1, 2, \ldots$$

where

$$L_n \in L(E, \hat{E}), \quad y_n \in \hat{E} \text{ and } z_n \in E.$$

We provide upper bounds on the distances $\|x_n - \tilde{x}_n\|$ and $\|\tilde{x}_n - x^*\|$.

Finally, our results are applied to an "ill conditioned" scalar equation considered also in [6], [12].

Main results

To make the paper self-contained we will reproduce some of the results obtained in [14] to fit our purposes.

Let F, D and x_0 be defined as in § 1 and consider iterative procedure (2). According to Dinner [2], Schmidt [11] and Yamamoto [14], we assume the following

$$(5)\quad \|A(x_0)^{-1}(F'(x) - F'(y))\| \leqq K\|x - y\|, \quad x, y \in D, \ K > 0,$$

$$(6)\quad \|A(x_0)^{-1}(A(x) - A(x_0))\| \leqq L\|x - x_0\| + l, \quad x \in D, \ L \geqq 0, \ l \geqq 0,$$

$$(7)\quad \|A(x_0)^{-1}(F'(x) - A(x))\| \leqq M\|x - x_0\| + m, \quad x \in D, \ M \geqq 0, \ m > 0,$$

$$(8)\quad l + m < 1, \ \sigma = \max\left(1, \frac{L + M}{K}\right), F(x_0) \neq 0,$$

$$(9)\quad \eta = \|A(x_0)^{-1}F(x_0)\|, \quad h = \sigma K\eta/(1 - l - m)^2 \leqq \frac{1}{2},$$

$$(10)\quad t^* = (1 - l - m)(1 - \sqrt{1 - 2h})/(\sigma K),$$

$$(11)\quad t^{**} = (1 - m + \sqrt{(1 - m)^2 - 2k\eta})/K,$$

(12) $$\overline{S} = \overline{S}(x_1, t^* - \eta) = \{x \in E \mid \|x - x_1\| \leqq t^* \eta\} \subseteqq D.$$

Under these assumptions, define the sequence $\{t_n\}$ by

(13) $$t_0 = 0, \quad t_{n+1} = t_n + f(t_n)/g(t_n), \quad n = 0, 1, 2, \dots$$

where

(14) $$f(t) = \frac{1}{2}\sigma K t^2 - (1 - l - m)T + \eta$$

and

(15) $$g(t) = 1 - l - Lt.$$

We can now state the following result [14, Theorem 4.1].

THEOREM 1. *With the above notation and assumptions, we have the following:*

(a) *The iterative procedure* (2) *is well defined for every* $n \geqq 0$, $x_n \in S$ *for* $n \geqq 1$ *and* $\{x_n\}$ *converges to a solution* $x^* \in \overline{S}$ *of the equation* (1).

(b) *The solution* x^* *is unique in*

(16) $$\overline{S} = \begin{cases} S(x_0, t^{**}) \cap D & (\text{if } \ 2K\eta < (1 - m)^2) \\ \overline{S}(x_0, t^{**}) \cap D & (\text{if } \ 2K\eta < (1 - m)^2). \end{cases}$$

Moreover, the following estimates are true:

(17) $$\|x_n - x^*\| \leqq t^* - t_n, \quad n = 0, 1, 2, \dots$$

where the nonnegative sequence $\{t_n\}$, $n = 0, 1, 2, \dots$ *is increasingly converging to* t^*.

We will finally need the result [14, Corollary 4.1.1].

THEOREM 2. *Consider the modified Newton method*

(18) $$x_{n+1} = x_n = F'(x_0)^{-1} F(x_n), \quad n = 0, 1, 2 \dots$$

where we assume the following:

(19) $$x_0 \in D, \ F'(x_0)^{-1} \text{ exists},$$

(20) $$\|F'(x_0)^{1-}(F'(x) - F'(x_0))\| \leqq K\|x - x_0\|, \ x \in D,$$

(21) $$\eta = \|F'(x_0)^{-1} F(x_0)\| > 0, \ h = K\eta \leqq \frac{1}{2},$$

$$\bar{t}^* = \left(1 - \sqrt{1-2h}\right)\Big/ K, \quad \bar{t}^{**} = \left(1 + \sqrt{1-2h}\right)\Big/ K, \tag{22}$$

$$\overline{S}_1 = \overline{S}_1(x_1, \bar{x}^* - \eta) \subseteqq D. \tag{23}$$

Then:

(a) *the iterative procedure* (18) *is well defined for every* $n \geqq 0$, $x_n \in S_1$ *for* $n \geqq 1$ *and* $\{x_n\}$ *converges to a solution* x^* *of equation* (1).

(b) *The solution* x^* *is unique in*

$$\bar{S} = \begin{cases} S_1(x_0, \bar{t}^{**}) \cap D & (\text{if } \ 2h < 1) \\ \overline{S}_1(x_0, \bar{t}^{**}) \cap D & (\text{if } \ 2h = 1). \end{cases} \tag{24}$$

Moreover, the following estimates are true

$$\|x_n - x^*\| \leqq \bar{t}^* - \bar{t}_n, \quad n = 0, 1, 2, \ldots \tag{25}$$

where the nonnegative sequence $\{\bar{t}_n\}$, $n = 0, 1, 2, \ldots$ *is given by*

$$\bar{t}_0 = 0, \quad \bar{t}_{n+1} = \frac{1}{2} K \bar{t}_n^2 + \eta, \quad n = 0, 1, 2, \ldots \tag{26}$$

and is increasingly converging to $\bar{t}^*$.

In what follows we shall suppose that there exist three possible numbers ε_1, ε_2, and ε_3 such that

$$\|y_n\| \leqq \varepsilon_1, \quad \|L_n\| \leqq \varepsilon_2, \quad \|z_n\| \leqq \varepsilon_3 \text{ for all } \ n \in N. \tag{27}$$

In Theorems 1 and 2 we have seen that the sequences produced by the iterative procedures (2) and (18) remain in the open balls S and S_1, respectively, and consequently in D. However, in the perturbed case we have to suppose that F is defined on the balls $S^* = S^*(x_0, r)$ and $S_1^* = S^*(x_0, r_1)$, respectively, with $r > t^{**}$, $r_1 > \bar{t}^{**}$ and $S^*, S_1^* \in D$. Set $S_2 = S^* \cup S_1^*$.

In the perturbed case it is more convenient to suppose that the following conditions are satisfied for $x, y \in S_2$:

$$\|A(x)^{-1}(F'(x) - F'(y))\| < K\|x - y\|, \tag{28}$$

$$\|A(x)^{-1}(A(x) - A(x_0))\| \leqq L\|x - x_0\| + \varrho, \tag{29}$$

and

$$\|A(x)^{-1}(F'(y) - A(y))\| \leqq M\|y - x_0\| + m. \tag{30}$$

These conditions are more restrictive than conditions (5), (6) and (7), respectively, but they are satisfied by the usual examples of approximation [4], [12].

In order to assure the invertibility of $A(\tilde{x}_n) + E_n$ for all $n = 0, 1, 2, \dots$ we shall suppose that $A(x)$ is invertible for all $x \in S_2$ and that the norms $\|A(x)^{-1}\|$ are bounded. More precisely, in the perturbed case we shall impose one, or both of the following conditions:

(C_1) The open ball S_2 is included into the domain of definition of F and conditions (28)–(30) hold for all $x, y \in S_2$.

(C_2) The linear operator $A(x)$ is invertible for all $x \in S_2$ and there exists a positive number a such that

$$a^{-1} \geqq \sup \left\{ \|A(x)^{-1}\|;\ x \in S_2 \right\}. \tag{31}$$

We can now prove the following theorem concerning the iterative procedure (3).

THEOREM 3. *Assume:*

(a) *The hypotheses of Theorem 1 are satisfied and let $\{x_n\}$, $n = 0, 1, 2, \dots$ be the sequence generated by (2);*

(b) *the conditions (C_1) and (C_2) are satisfied;*

and

(c) *the inequalities*

$$(1-\beta)^2 - 4\alpha\gamma \geqq 0, \quad a > \varepsilon_2, \tag{32}$$

$$\beta \leqq 1, \tag{33}$$

$$8 - \frac{1 - \beta - \sqrt{(1-\beta)^2 - 4\alpha\gamma}}{2\alpha} < 1 - t^* + \eta \tag{34}$$

are satisfied, where

$$\alpha = \frac{a(K+2L)}{2(a-\varepsilon_2)}, \tag{35}$$

$$\beta = \{(K+L)\varepsilon_0 + m + l + \varepsilon_1^* + \varepsilon_2^* + L\varepsilon_0\} \frac{a}{a-\varepsilon_2}, \quad \varepsilon_0 = t^* - t_0, \quad \varepsilon_1^* = \frac{\varepsilon_1}{a}, \quad \varepsilon_2^* = \frac{\varepsilon_2}{a}, \tag{36}$$

and

$$\gamma = \frac{a}{a-\varepsilon_2}(2L + \varepsilon_0 + l + \varepsilon_2^*)\varepsilon_0 + \varepsilon_3. \tag{37}$$

Then the iterative procedure (3) is well defined and for each $n \in N$ we shall have the estimates

$$\|x_n - \tilde{x}_n\| \leqq w_n \leqq \delta, \tag{38}$$

where the real sequence $\{w_n\}$ *is given by*

$$w_{n+1}=\alpha_{n+1}w_n^2+\beta_{n+1}w_n+\gamma_{n+1},\qquad w_0=0,\quad n=0,1,2,\dots \tag{39}$$

with

$$\alpha_n=\frac{a(2L+K)}{2(a-\varepsilon_2)}, \tag{40}$$

$$\beta_n=[(K+L)(t_n-t_0)+m+l+\varepsilon_2^*+L(t_{n+1}-t_n)]\,\frac{a}{a-\varepsilon_2}, \tag{41}$$

and

$$\gamma_n=(2L(t_n-t_0)+\varepsilon_1^*+\varepsilon_2^*+l)(t_{n+1}-t_n)\frac{a}{a-\varepsilon_2}+\varepsilon_3,\ \ n=1,2,\dots. \tag{42}$$

PROOF. For $n=0$ the inequalities (38) are trivially satisfied as equalities. Suppose they are satisfied for $n=0,1,\dots,k$, $k\geqq 0$. From (38) it follows that $\bar{x}_k\in S_2$. In this case, condition (C_2) implies, according to the Banach lemma on invertible operators that the linear operator $A(\bar{x}_k)+E_k$ is invertible and

$$\begin{aligned}\|(A(\bar{x}_k)+E_k)^{-1}\|&=\|(I+A(\bar{x}_k)^{-1}L_k)^{-1}A(\bar{x}_k)^{-1}\|\leqq\\ &\leqq\|(I+A(\bar{x}_k)^{-1}L_k)^{-1}\|\,\|A(\bar{x}_k)^{-1}\|\leqq(a-\varepsilon_2)^{-1}.\end{aligned} \tag{43}$$

From (2) and (3) we obtain the identity

$$\begin{aligned}&\bar{x}_{k+1}-x_{k+1}=\\ &=(I+A(\bar{x}_k)^{1-}L_k)^{-1}A(\bar{x}_k)^{-1}\{[F(x_k)-F(\bar{x}_k)-f'(x_k)(x_k-\bar{x}_k)]+\\ &+[(F'(x_k)-A(\bar{x}_k))(x_k-\bar{x}_k)]+y_k+[L_k(\bar{x}_k-x_k)]+\\ &+[(A(\bar{x}_k)-A(x_k))A(x_k)^{-1}F(x_k)]+[L_kA(x_k)^{-1}F(x_k)]\}+z_k.\end{aligned} \tag{44}$$

Using (28), (38) and (43) we obtain

$$\|A(\bar{x}_k)^{-1}[F(x_k)-F(\bar{x}_k)-f'(x_k)(x_k-\bar{x}_k)]\|\leqq\frac{1}{2}K\|x_k-\bar{x}_k\|\leqq\frac{1}{2}Kw_k^2. \tag{45}$$

By (28)–(30) and (38) we get

$$\begin{aligned}\|A(\bar{x}_k)^{-1}(F'(x_k)-A(\bar{x}_k))\|&=\|[(F'(x_k)-F'(x_0))+(F'(x_0)-A(x_0))\\ &+(A(x_0)-A(\bar{x}_k))](x_k-\bar{x}_k)\|\leqq\\ &\leqq[K\|x_k-x_0\|+L\|\bar{x}_k-x_0\|+l+m]\|x_k-\bar{x}_k\|\leqq\\ &\leqq[K(t_k-t_0)+Lw_k+L(t_k-t_0)+m+l]w_k\leqq\\ &\leqq(K\varepsilon_0+Lw_k+L\varepsilon_0+m+l)w_k.\end{aligned} \tag{46}$$

From (27) and (38) we get

$$\|A(\bar{x}_k)^{-1}L_k(\bar{x}_k - x_k)\| \leqq \varepsilon_2^*\|\bar{x}_k - x_k\| \leqq \varepsilon_2^* w_k. \tag{47}$$

Using (2), (29) we obtain

$$\begin{aligned} &\|A(\tilde{x}_k)^{-1}(A(\tilde{x}_k) - A(x_k))\| = \\ &= \|A(\tilde{x}_k)^{-1}[(A(\tilde{x}_k) - A(x_0)) + (A(x_0) - A(x_k))](x_k - x_{k+1})\| \leqq \\ &\leqq (L(\|\tilde{x}_k - x_0\| + \|x_k - x_0\|) + 2l)\|x_k - x_{k+1}\| \leqq \\ &\leqq (Lw_k + 2L(t_k - t_0) + 2l)(t_{k+1} - t_k) \end{aligned} \tag{48}$$

(since, by Theorem 1, $\|x_k - x_{k+1}\| \leqq t_{k+1} - t_k \leqq t^* - t_0 = \varepsilon_0$)

$$\leqq (Lw_k + 2L(t^* - t_0) + 2l)(t^* - t_k).$$

Finally, by (2) and (27)

$$\|A(\bar{x}_k)^{-1}L_kA(x_k)^{-1}F(x_k)\| \leqq \varepsilon_2^*\|x_k - x_{k+1}\| \leqq \varepsilon_2^*(t_{k+1} - t_k) \leqq \varepsilon_2^*\varepsilon_0. \tag{49}$$

With these majorization in (44), using (43), (27), (32)–(37), and (39) we can easily obtain that

$$\|\tilde{x}_{k+1} - x_{k+1}\| \leqq w_{k+1} \leqq \alpha\delta^2 + \beta\delta + \gamma = \delta. \tag{50}$$

That completes the induction and the proof of the theorem.

Concerning the perturbed iteration (4) we have:

THEOREM 4. *Assume:*
(a) *The hypotheses of Theorem 2 are satisfied and let $\{x_n\}$, $n = 0, 1, 2, \ldots$ be the sequence generated by* (18)*;*
(b) *the condition* (C_1) *is satisfied and* (28) *with $A(x) = F'(x)$ and $y = x_0$.*
(c) *The inequalities*

$$d > e_2, \ \textit{with} \ 0 < d \leqq \|F'(x_0)^{-1}\|^{-1}, \tag{51}$$

$$(1 - \beta_1)^2 - 4\alpha_1\gamma_1 \geqq 0, \tag{52}$$

$$\beta_1 \leqq 1, \tag{53}$$

$$\delta_1 = \frac{1 - \beta_1 - \sqrt{(1 - \beta_1)^2 - 4\alpha_1\gamma_1}}{2\alpha_1} \leqq r_1 - t^* + \eta \tag{54}$$

are satisfied, where

$$\alpha_1 = \frac{dK}{2(d-\varepsilon_2)}, \tag{55}$$

$$\beta_1 = \frac{\varepsilon_0 dK + \varepsilon_2}{d-\varepsilon_2}, \tag{56}$$

and

$$\gamma_1 = \frac{\varepsilon_2\varepsilon_0 + \varepsilon_1 + \varepsilon_3(d-\varepsilon_2)}{d-\varepsilon_2}. \tag{57}$$

Then the iterative procedure (18) *is well defined and for each* $n \in N$ *we shall have the estimates*

$$\|x_n - \bar{x}_n\| \leqq S_n \leqq \delta_1 \tag{58}$$

where the real sequence $\{S_n\}$ *is given by*

$$S_{n+1} = \alpha^*_{n+1} S_n^2 + \beta^*{}_{n+1} S_n + \gamma^*{}_{n+1}, \qquad S_0 = 0, \quad n = 0, 1, 2 \dots \tag{59}$$

with

$$\alpha^*_n = \frac{dK}{2(d-\varepsilon_2)}, \tag{60}$$

$$\beta^*_n = \frac{dK(t_n - t_0) + \varepsilon_2}{d-\varepsilon_2}, \tag{61}$$

and

$$\gamma^*_n = \frac{\varepsilon^*_2(t_{n+1} - t_n) + \varepsilon_1}{d-\varepsilon_2} + \varepsilon_3, \quad n = 1, 2, \dots. \tag{62}$$

Proof. By the Banach lemma on invertible operators it follows that the linear operator $F'(x_0) + L_0$ is invertible and

$$\begin{aligned} &\|(F'(x_0) + L_0)^{-1}\| = \|(I + F'(x_0)^{-1})^{-1} F'(x_0)\| \\ &\leqq \|(I + F'(x_0)^{-1} L_0)^{-1}\| \, \|F'(x_0)^{-1}\| \leqq (d-\varepsilon_2)^{-1}. \end{aligned} \tag{63}$$

This fact, together with the remark that (54) and (58) imply $\bar{x}_n \in S_2$, shows us that if (58) is satisfied, then the iterative procedure (4) makes sense.

We will show that (58) holds for all $n = 0, 1, 2, \dots$. For $n = 0$ the inequalities (58) are trivially satisfied as equalities. Suppose they are satisfied for $n = 0, 1, \dots, k$. We shall prove that they hold for $n = k + 1$, too.

From (4) and (18) we obtain the identity

$$x_{k+1}-\bar{x}_{k+1}=(I+F'(x_0)^{-1}L_0)^{-1}F'(x_0)\{[F(\bar{x}_k)-F(x_k)-F'(x_0)(\bar{x}_k-x_k)]+ +y_k+L_0(x_n-\bar{x}_n)-L_0F'(x_0)^{-1}F(x_n)]\}-z_n. \tag{64}$$

By taking norms in the above inequality and using (28), (64), (27), (18), (25), (26), (51)–(62), we obtain as in (50) that

$$\begin{aligned}\|x_{k+1}-\bar{x}_{k+1}\| &\leqq \frac{dK}{2(d-\varepsilon_2)}(\|x_0-x_k\|+\|x_0-\bar{x}_k\|)\|x_k-\bar{x}_k\| \\ &+\frac{1}{d-\varepsilon_2}[\varepsilon_2\|x_k-\bar{x}_k\|+\varepsilon_2\|x_k-x_{k+1}\|+\varepsilon_1]+\varepsilon_3 \\ &\leqq \frac{dK}{2(d-\varepsilon_2)}(2\|x_0-x_k\|+\|x_k-\bar{x}_k\|)\|x_k-\bar{x}_k\| \\ &+\frac{1}{s-\varepsilon_2}[\varepsilon_2\|x-\bar{x}_k\|+\varepsilon_2\|x_k-x_{k+1}\|+\varepsilon_1]+\varepsilon_3 \\ &\leqq \frac{dK}{2(d-\varepsilon_2)}[2(\bar{x}_k-\bar{t}_0)+S_k]S_k+\frac{1}{d-\varepsilon_2}[S_k+(\bar{t}_{k+1}-\bar{t}_k)+\varepsilon_1]+\varepsilon_3 \\ &=S_{k+1}\leqq \frac{dK}{2(d-\varepsilon_2)}[2(\bar{t}^*-\bar{t}_0)+\delta_1]\delta_1 \\ &+\frac{1}{d-\varepsilon_2}(\delta_1+\bar{t}^*-\bar{t}_0+\varepsilon_1)+\varepsilon_3=\delta_1.\end{aligned} \tag{65}$$

That completes the induction and the proof of the theorem.

The following result follows immediately from Theorems 1–4.

COROLLARY. *Under the hypotheses of Theorem 3 and 4 the following estimates hold for all* $n=0,1,2,\ldots$*:*

$$\|\bar{x}_n-x^*\|\leqq w_n+t^*-t_n \tag{66}$$

and

$$\|\bar{x}_n-x^*\|\leqq S_n+t^*-t_n \tag{67}$$

for iteration (3) *and* (4)*, respectively.*

Finally, we remark that the approach employed here applies for the rest of the error bounds obtained in [14, pp. 550, 555].

Applications

We shall apply Theorem 4 to an „ ill-conditioned" example proposed by Wilkinson [12] and considered also by Lancaster [6].

EXAMPLE. Consider solving iteratively the quadratic equation

$$x^2 - 2.028888800x + 1.02876900 = 0 \tag{68}$$

using a computer characterized by the accuracy $c_1 = c_2 = c_3 = .5 \times 10^{-9}$.

Starting with $z_0 = 1.032567321$ and using (18) we get

$z_1 = 1.032567323$

$z_2 = 1.032567326$

$z_3 = 1.032567329$

$z_n = z_3,\ n \geqq 3$.

If we take $x_0 = z_2$ and $\max(r, r_1) = 0.03624585$, then we can easily obtain from (51)–(62) $\bar{K} = 55.17872776$, $\eta = 2.75893638 \cdot 10^{-9}$, $h = 1.52234599 \cdot 10^{-7}$, $\bar{t}^* = 2.75903425 \cdot 10^{-9}$, $\bar{t}^{**} = .0362445849$, $\varepsilon_0 = \bar{t}^*$, $d = .036245852$, $\alpha_1 = 27.5893654$, $\beta_1 = 1.660346909 \cdot 10^{-7}$, $\gamma_1 = 1.42946827 \cdot 10^{-8}$ and $\delta_1 = 11.3630739 \cdot 10^{-9}$.

We want to find and estimate for the distance $|z_3 - x^*|$. The hypoteses of Theorem 4 being satisfied we can use the Corollary.

From (67) we get $|z_3 - x^*| \leqq 21 \times 10^{-9}$.

Taking advantage of the fact that we know that the sequence $\{z_n\}$ becomes constant beginning with $n = 3$, we easily obtain that $|z_3 - x^*| \leqq 11.3630739 \cdot 10^{-9}$. This is very close to reality because $x^* = 1.032567332$ is the solution of equation (68).

REFERENCES

[1] ARGYROS, I. K., On Newton's method and nondiscrete mathematical induction, *Bull. Austral. Math. Soc.* **38** (1988), 131–140. *MR* **90a:**65136

[2] DENNIS, J. E., Toward a unified convergence theory for Newton-like methods, *Nonlinear Functional Anal. and Appl.* (Proc. Advanced Sem., Math. Res. Center, Univ. of Wisconsin, Madison, Wis., 1970), Academic Press, New York, 1971, 425–472. *MR* **43**#4286

[3] GRAGG, W. B. and TAPIA R. A., Optimal error bounds for the Newton–Kantorovich theorem, *SIAM J. Numer. Anal.* **11** (1974), 10–13. *MR* **49**#8334

[4] KANTOROVICH, L. V. and AKILOV, G. P., *Functional analysis in normed spaces*, International series of monographs in pure and applied mathematics, Vol. 46, Macmillan, New York; Pergamon Press, Oxford, 1964. *MR* **35**#4699

[5] KORNSTAEDT, H.-J., Funktionalungleichungen und Iterationsverfahren, *Aequationes Math.* **13** (1975), 21–45. *MR* **52**#9597

[6] LANCASTER, P., Error analysis for the Newton–Raphson method, *Numer. Math.* **9** (1966), 55–68. *MR* **35**#1208

[7] MIEL, G. J., Majorizing sequences and errors bounds for iterative methods, *Math. Comp.* **34** (1980), 185–202. *MR* **81h**:65056

[8] MORET, I., A note on Newton-type iterative methods, *Computing* **33** (1984), 65–73, *MR* **86a**:65052

[9] POTRA, F.-A. and PTAK, V., Sharp error bounds for Newton's process, *Numer. Math.* **34** (1980), 63–72. *MR* **81c**:65027

[10] RHEINBOLDT, W. C., A unified convergence theory for a class of iterative processes, *SIAM J. Numer. Anal.* **5** (1968), 42–63. *MR* **37**#1061

[11] SCHMIDT, J. W., Untere Fehlerschranken für Regula-falsi-Verfahren, *Period. Math. Hungar.* **9** (1978), 241–247. *MR* **58**#13676

[12] WILKINSON, J. H., *The algebraic eigenvalue problem*, Clarendon Press, Oxford, 1965. *MR* **32**#1894

[13] YAMAMOTO, T., A method for finding sharp error bounds for Newton's method under the Kantorovich assumptions, *Numer. Math.* **49** (1986), 203–220. *MR* **87i**:65096

[14] YAMAMOTO, T., A convergence theorem for Newton-like methods in Banach spaces, *Numer. Math.* **51** (1987), 545–557. *MR* **88i**:65081

(Received March 7, 1991)

DEPARTMENT OF MATHEMATICAL SCIENCES
CAMERON UNIVERSITY
LAWTON, OK 73505–6377
U. S. A.

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 307–313

A NOTE ON THE OPERATIONAL SOLUTION OF AN INTEGRAL EQUATION

T. FÉNYES

In the papers [1], [2], [3], we have used Mikusiński's operational calculus to obtain the solution of the integral equation

$$(1) \qquad (t+a)f(t) + \int_0^t f(\tau)g(t-\tau)d\tau = h(t), \quad t \geqq 0,$$

where g, h are locally integrable functions ($g, h \in L_{\text{loc}}$), a is a real number. For $a \leqq 0$ (1) is a convolution type integral equation of the third kind. In the following we assume that Mikusiński's operational calculus and the operational notations are familiar to the reader. In some places we denote the convolution of functions also with a star. By introducing the algebraic derivative D, (1) can be reduced to the following inhomogeneous algebraic differential equation in the operator field M

$$(2) \qquad D(f) - (a+g)f = -h, \quad f \in M$$

being more general than (2), since (2) can have not only locally integrable ($f \in L_{\text{loc}}$) solutions but solutions contained only in the field M. We have proved the following

THEOREM. *Let* $\lim_{t\to+0} g(t) = \lambda$ *exist such that* $\frac{g(t)-\lambda}{t} \in L_{\text{loc}}$. *Then the general solution of the homogeneous equation*

$$(3) \qquad D(f) - (a+g)f = 0$$

is of the form

$$f = Ce^{as}s^{\lambda} \exp\left\{\frac{g(t)-\lambda}{-t}\right\},$$

where C *is an arbitrary number and*

$$\exp\left\{\frac{g(t)-\lambda}{-t}\right\} = \sum_{k=0}^{\infty}\left\{\frac{g(t)-\lambda}{-t}\right\}^k \Big/ k!$$

1991 *Mathematics Subject Classification.* Primary 44A40; Secondary 45E10.
Key words and phrases. Operational calculus of Mikusiński, integral equations.

Research partially supported by the Hungarian National Foundation for Scientific Research Grant No. 6032/6319.

in the sense of the operational convergence.

For $C \neq 0$, $f \in L_{\mathrm{loc}}$ if and only if $\lambda < 0$ and $a \leqq 0$. The inhomogeneous equation (2) *has for $a > 0$ exactly one $f \in L_{\mathrm{loc}}$ solution of the form*

$$f = \exp\left\{\frac{g(t)-\lambda}{-t}\right\} \times \tag{5}$$
$$\times\left\{\frac{h + h * G_0}{t+a} - \lambda(t+a)^{-\lambda-1}\int_0^t \frac{[h + h * G_0](\tau)}{(\tau+a)^{1-\lambda}}d\tau\right\},$$

where

$$G_0(t) = \sum_{k=1}^{\infty}\left\{\frac{g(t)-\lambda}{t}\right\}^k \Big/ k!.$$

For $a = 0$, (2) *has a particular solution $f \in M$ of the form*

$$f = \exp\left\{\frac{g(t)-\lambda}{-t}\right\} s^2\left\{t^{-\lambda+1}\int_0^t \frac{[lh + lh * G_0](\tau)}{\tau^{2-\lambda}}d\tau\right\}, \quad \textit{for } \lambda > 1, \tag{6}$$
$$f = \exp\left\{\frac{g(t)-\lambda}{-t}\right\} s^2\left\{t^{-\lambda+1}\int_\varepsilon^t \frac{[lh + lh * G_0](\tau)}{\tau^{2-\lambda}}d\tau\right\}, \quad \textit{for } \lambda \leqq 1,$$

where $\varepsilon > 0$ is arbitrary, l is the integral operator.

In general (2) *has no locally integrable solution. If $\frac{h(t)}{t} \in L_{\mathrm{loc}}$, then* (2) *has a particular ($f \in L_{\mathrm{loc}}$) solution of the form*

$$f = \exp\left\{\frac{g(t)-\lambda}{-t}\right\} \times \tag{7}$$
$$\times\left\{\frac{h(t) + h(t) * G_0(t)}{t} - \lambda t^{-\lambda-1}\int_0^t \frac{[h + h * G_0](\tau)}{\tau^{1-\lambda}}d\tau\right\} \quad \textit{for } \lambda \geqq 0,$$
$$f = \exp\left\{\frac{g(t)-\lambda}{-t}\right\} \times$$
$$\times\left\{\frac{h(t) + h(t) * G_0(t)}{t} - \lambda t^{-\lambda-1}\int_\varepsilon^t \frac{[h + h * G_0](\tau)}{\tau^{1-\lambda}}d\tau\right\} \quad \textit{for } \lambda < 0.$$

By the result of paper [4] the condition referring to the existence of $g(+0)$ in the above Theorem may by replaced by the following

CONDITION 1. There exists a real number λ such that $\frac{g(t)-\lambda}{t} \in L_{\mathrm{loc}}$.

In this paper we shall discuss the integral equation

$$(8) \qquad (t+a)f(t)+(t+b)\int_0^t f(\tau)g(t-\tau)d\tau=h(t)$$

by the operational calculus, where a, g, h are the same as above, $b \neq a$ is another arbitrary real number. In the inhomogeneous case we restrict ourselves to $a \geqq 0$.

Introducing the algebraic derivative D we obtain

$$(8) \qquad -D(f)+af-D(fg)+bfg=h$$

and

$$D(f)-af+D(f)g+fD(g)-bfg=-h,$$

so we have

$$D(f)(1+g)+(-a+D(g)-bg)f=-h,$$

$$D(f)+\frac{-a+D(g)-bg}{1+g}f=-\frac{h}{1+g},$$

which can be written in the form

$$(9) \qquad D(f)+\left(-a+\frac{D(g)+(a-b)g}{1+g}\right)f=-\frac{h}{1+g}.$$

So we have reduced (8) to the algebraic differential equation (9).

Let us substitute

$$(10) \qquad f=\frac{u}{1+g}, \qquad u\in M$$

in (9), then an easy calculation gives that

$$(11) \qquad D(u)+\left[-a+\frac{(a-b)g}{1+g}\right]u=-h$$

holds. Since

$$(11') \qquad \frac{1}{1+g}=\sum_{\nu=0}^{\infty}(-1)^{\nu}g^{\nu}$$

in the sense of the operational convergence it is easily seen by (10) that

$$u\in L_{\text{loc}} \longleftrightarrow f\in L_{\text{loc}}$$

hold. Moreover, by introducing the function ζ by the definition

$$(12) \qquad \zeta=\frac{g}{1+g}=\sum_{\nu=0}^{\infty}(-1)^{\nu}g^{\nu+1}\in L_{\text{loc}},$$

(11) can be written as

$$(13) \qquad D(u)-[a+(b-a)\zeta]u=-h$$

being of type (2).

In the following it will be assumed that Condition 1 holds. We need here the following

LEMMA. *Let $\overline{f}(z)$ be any function of the complex variable z, analytic in a circle such that*

$$\overline{f}(z)=\sum_{\nu=0}^{\infty}a_\nu z^\nu,$$

and let

$$F=g\overline{f}(g)=\sum_{\nu=0}^{\infty}a_\nu g^{\nu+1}, \tag{14}$$

in the sense of the operational convergence, then

$$\frac{F(t)-\lambda a_0}{t}\in L_{\rm loc}$$

holds.

PROOF. By a result of Mikusiński [5], $F\in L_{\rm loc}$. In [1] we have shown the statement that for arbitrary k_1, $k_2\in L_{\rm loc}$, for which $\frac{k_1(t)}{t}\in L_{\rm loc}$,

$$\frac{k_1*k_2}{t}\in L_{\rm loc}.$$

For $\lambda=0$ the Lemma trivially holds. Let now $\lambda\neq 0$. Using operational notation we have

$$F-\frac{\lambda a_0}{s}=a_0\left(g-\frac{\lambda}{s}\right)+\left(g-\frac{\lambda}{s}\right)\sum_{\nu=1}^{\infty}a_\nu g^\nu+\frac{\lambda}{s}\sum_{\nu=1}^{\infty}a_\nu g^\nu. \tag{15}$$

Taking into account the above statement it is only necessary to show that for

$$Q=\frac{\lambda}{s}\sum_{\nu=1}^{\infty}a_\nu g^\nu$$

$\frac{Q(t)}{t}\in L_{\rm loc}$ holds. Since

$$Q=\frac{\lambda}{s}\left(g-\frac{\lambda}{s}\right)\sum_{\nu=1}^{\infty}a_\nu g^{\nu-1}+\frac{\lambda^2}{s^2}\sum_{\nu=1}^{\infty}a_\nu g^{\nu-1}= \qquad (g^0=1)$$

$$=\frac{\lambda}{s}\left(g-\frac{\lambda}{s}\right)a_1+\frac{\lambda}{s}\left(g-\frac{\lambda}{s}\right)\sum_{\nu=2}^{\infty}a_\nu g^{\nu-1}+\frac{\lambda^2}{s^2}a_1+ \tag{16}$$

$$+\frac{\lambda^2}{s^2}\sum_{\nu=2}^{\infty}a_\nu g^{\nu-1},$$

and

$$\sum_{\nu=2}^{\infty} a_\nu g^{\nu-1} \in L_{\text{loc}}, \quad \frac{\lambda}{s} = \{\lambda\}, \quad \frac{\lambda^2}{s^2} = \{\lambda^2 t\},$$

so, applying again the above statement, it is easily seen that $\frac{Q(t)}{t} \in L_{\text{loc}}$ and the Lemma holds.

So we have $\frac{\zeta(t)-\lambda}{t} \in L_{\text{loc}}$. Applying the Theorem, Lemma, (13) and (10) we obtain the following Theorems.

THEOREM 1. *Let us assume that Condition 1 holds. The homogeneous algebraic differential equation corresponding to* (9) *has the general solution of the form*

$$f = \frac{Ce^{as}s^{(b-a)\lambda}}{1+g} \exp\left\{\frac{(b-a)(\zeta(t)-\lambda)}{-t}\right\}, \tag{17}$$

where C is an arbitrary number and

$$\zeta = \frac{g}{1+g} = \sum_{\nu=0}^{\infty}(-1)^\nu g^{\nu+1},$$

$$\exp\left\{\frac{(b-a)(\zeta(t)-\lambda)}{-t}\right\} = \sum_{k=0}^{\infty} \frac{(b-a)^k}{k!}\left\{\frac{\zeta(t)-\lambda}{-t}\right\}^k.$$

For $C \neq 0$, $f \in L_{\text{loc}}$ if and only if $(b-a)\lambda < 0$ and $a \leqq 0$.

THEOREM 2. *The inhomogeneous equation* (9) *has for $a > 0$ exactly one $f \in L_{\text{loc}}$ solution of the form*

$$\begin{aligned} f = \frac{\exp\left\{\frac{(b-a)(\zeta(t)-\lambda)}{-t}\right\}}{1+g}&\left\{\frac{h+h*\overline{G}_0}{t+a} - (b-a)\lambda(t+a)^{-(b-a)\lambda-1}\times\right. \\ &\left.\times\int_0^t \frac{[h+h*\overline{G}_0](\tau)}{(\tau+a)^{1-(b-a)\lambda}}d\tau\right\}, \end{aligned} \tag{18}$$

where

$$\overline{G}_0(t) = \sum_{k=0}^{\infty}\frac{(b-a)^k}{k!}\left\{\frac{\zeta(t)-\lambda}{t}\right\}^k.$$

For $a=0$, (9) has a particular solution $f\in M$ of the form

(19)
$$f=\frac{\exp\left\{\frac{(b-a)(\zeta(t)-\lambda)}{-t}\right\}}{1+g}s^2\left\{t^{-(b-a)\lambda+1}\times\right.$$
$$\left.\times\int_0^t\frac{[lh+lh*\overline{G}_0](\tau)}{\tau^{-(b-a)\lambda+2}}d\tau\right\}\quad \textit{for}\quad (b-a)\lambda>1,$$
$$f=\frac{\exp\left\{\frac{(b-a)(\zeta(t)-\lambda)}{-t}\right\}}{1+g}s^2\left\{t^{-(b-a)\lambda+1}\times\right.$$
$$\left.\times\int_\varepsilon^t\frac{[lh+lh*\overline{G}_0](\tau)}{\tau^{-(b-a)\lambda+2}}d\tau\right\}\quad \textit{for}\quad (b-a)\lambda\leqq 1,$$

where $\varepsilon>0$ is arbitrary and $l=\frac{1}{s}$ is the integral operator. In general (9) has no locally integrable solution.

If $\frac{h(t)}{t}\in L_{\mathrm{loc}}$, then (9) has a particular ($f\in L_{\mathrm{loc}}$) solution of the form

(20)
$$f=\frac{\exp\left\{\frac{(b-a)(\zeta(t)-\lambda)}{-t}\right\}}{1+g}\left\{\frac{h+h*\overline{G}_0}{t}-\right.$$
$$\left.-(b-a)\lambda t^{-(b-a)\lambda-1}\int_0^t\frac{h+h*\overline{G}_0]}{\tau^{1-(b-a)\lambda}}d\tau\right\}\quad \textit{for}\quad (b-a)\lambda\geqq 0,$$
$$f=\frac{\exp\left\{\frac{(b-a)(\zeta(t)-\lambda)}{-t}\right\}}{1+g}\left\{\frac{h+h*\overline{G}_0}{t}-\right.$$
$$\left.-(b-a)\lambda t^{-(b-a)\lambda-1}\int_\varepsilon^t\frac{h+h*\overline{G}_0]}{\tau^{1-(b-a)\lambda}}d\tau\right\}\quad \textit{for}\quad (b-a)\lambda<0.$$

REMARK 1. Obviously, the general solution of (9) is the sum of the general solution of the corresponding homogeneous equation and of a particular solution of the inhomogeneous equation.

REMARK 2. The condition $\frac{h(t)}{t}\in L_{\mathrm{loc}}$ is not necessary. The integral equation

(21)
$$tf(t)+(t+1)\int_0^t f(\tau)d\tau=h(t)$$

where

$$h(t)=\begin{cases}\dfrac{1}{\log^2 t}-\dfrac{t}{\log t}-\dfrac{1}{\log t}, & 0\leqq t\leqq \delta<1\\ -\dfrac{t}{\log\delta}-\dfrac{1}{\log\delta}, & t>\delta\end{cases}$$

has the general (operational) solution

$$f=C\left(s-1+\{e^{-t}\}\right)\exp\left\{\frac{e^{-t}-1}{-t}\right\}+\{\overline{f}(t)\} \tag{22}$$

where

$$\overline{f}(t)=\begin{cases}\dfrac{1}{t\log^2 t}, & 0\leqq t\leqq\delta<1,\\ 0, & t>\delta.\end{cases}$$

For $C=0$ we get a locally integrable solution. Since $\frac{1}{t\log t}\notin L_{\text{loc}}$, so is $\frac{h(t)}{t}\notin L_{\text{loc}}$.

REFERENCES

[1] Fényes, T., Anwendung der Mikusińskischen Operatorenrechnung zur Lösung von Integralgleichungen dritter Art vom Faltungstypus, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **9** (1965), 365–399. *MR* **32**#8071

[2] Fényes, T., A note on the solution of integral equations of convolution type of the third kind by application of the operational calculus of Mikusiński, *Studia Sci. Math. Hungar.* **2** (1967), 81–89. *MR* **35**#2097

[3] Fényes, T., On the operational solution of a convolution type integral equation of the third kind, *Studia Sci. Math. Hungar.* **12** (1977), 65–75. *MR* **82a**:44005

[4] Fényes, T., On an integral equation of the third kind, *Mat. Lapok* **32** (1981–1985), 247–248 (in Hungarian). *MR* **87k**:45010

[5] Mikusiński, J. and Boehme, T. K., *Operational Calculus*, Vol. II, Second edition, International Series of Monographs in Pure and Applied Mathematics, **110**, Pergamon Press, Oxford – Elmsford, N. Y., PWN – Polish Scientific Publishers, Warszawa, 1987. *MR* **88k**:44010

(Received March 1, 1991)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 315–322

SEQUENCES AND THEIR TRANSFORMS WITH IDENTICAL ASYMPTOTIC DISTRIBUTION FUNCTION MODULO 1

S. H.-MOLNÁR

Let $\omega=(x_n)_{n=1}^{\infty}$ be a given sequence of real numbers. For a positive integer N and a subset E of the interval $I=[0,1)$ let the counting function $A(E;N;\omega)$ be defined as the number of terms x_n, $1\leqq n\leqq N$, for which $\{x_n\}\in E$, where $\{t\}$ denotes the fractional part of the real number t. In the following we shall use the terminology and notations of Kuipers and Niederreiter [2] adapted to suit our purposes.

We need the following known definitions.

DEFINITION 1. The sequence $\omega=(x_n)_{n=1}^{\infty}$ of real numbers is said to be uniformly distributed modulo 1 (abbreviated by u.d. mod 1) if for every pair (a,b) of real numbers with $0\leqq a<b\leqq 1$ we have

$$\lim_{N\to\infty}\frac{A([a,b);N;\omega)}{N}=b-a.$$

DEFINITION 2. The sequence $\omega=(x_n)_{n=1}^{\infty}$ is said to have asymptotic distribution function modulo 1 (abbreviated by a.d.f. mod 1) g if

$$\lim_{N\to\infty}\frac{A([0,x);N;\omega)}{N}=g(x)\quad\text{for } 0\leqq x\leqq 1.$$

Suppose that a sequence ω has a.d.f. mod 1 which is F. If f is a real valued function, defined at the terms of ω, and the sequence $\xi=(f(x_n))_{n=1}^{\infty}$ also has a.d.f. mod 1, then, in general, the two distribution functions are distinct. E.g. if $\omega=(n\Theta)_{n=1}^{\infty}$, where Θ is an irrational number, and $f(x)=\cos x$ or $f(x)=\frac{1}{x}$, then the a.d.f. mod 1 of ω is different from the a.d.f. mod 1 of the sequence $\xi=(\cos n\Theta)_{n=1}^{\infty}$. Another most striking example is the following one: The sequence $\omega=(F_n)_{n=1}^{\infty}$ of the Fibonacci numbers has a non-continuous a.d.f. mod 1 but the a.d.f. mod 1 of the sequence $\xi=(\log F_n)_{n=1}^{\infty}$ is continuous (see Kuipers [1]). However, there are sequences

---

1991 *Mathematics Subject Classification.* Primary 11K06.

Key words and phrases. Sequences of real numbers, distribution mod 1, distribution function, uniform distribution preserve function, f invariant distributed sequence mod 1.

Research (partially) supported by Hungarian National Foundation for Scientific Research Grant 273.

ω and functions f such that the sequences $\omega = (x_n)_{n=1}^{\infty}$ and $\xi = (f(x_n))_{n=1}^{\infty}$ have the same a.d.f. mod 1.

We introduce a definition.

DEFINITION 3. Let $\omega = (x_n)_{n=1}^{\infty}$ and $\xi = (f(x_n))_{n=1}^{\infty}$ be sequences of real numbers, where f is a real-valued function. If the sequences ω and ξ have a.d.f. mod 1 and these functions are identical, then we say ω is f invariant distributed sequence modulo 1 (abbreviated by i.d. mod 1 to f).

EXAMPLES. 1. If Θ is a positive irrational number, then the sequences $\omega = (n\Theta)_{n=1}^{\infty}$ and $\xi = (\sqrt{n\Theta})_{n=1}^{\infty}$ are u.d. mod 1 and so they have common asymptotic distribution function mod 1 ($F(x) = x$). Thus ω is a square root invariant distributed sequence mod 1.

2. We know sequences also having a.d.f. mod 1 which are not uniformly distributed mod 1 but they are invariant distributed to a function. E.g. in [3] conditions are given for second-order linear recurrences $(G_n)_{n=1}^{\infty}$ such that the sequence $\omega = \left(\frac{G_n}{G_{n+k}}\right)_{n=1}^{\infty}$, with a fixed positive integer k, to be invariant distributed to the function $f(x) = \frac{1}{x}$ (we say reciprocal invariant).

3. There are functions, e.g. $f_1(x) = \{x\}$ or $f_2(x) = x + k$, where k is an integer, such that any sequence of real numbers, which has a.d.f. mod 1, is invariant distributed mod 1 to these functions. But if a real number c is not integer, then not every sequence is invariant distributed mod 1 to the function $f_3(x) = x + c$, but each uniformly distributed sequence mod 1 is i.d. mod 1 to f_3.

The functions f that map the interval $[0,1]$ into itself and any uniformly distributed sequence mod 1 is invariant distributed mod 1 to f are called uniform distribution preserve function (u.d.p.). Some results for u.d.p. functions can be found in [4].

In the following we prove two theorems for mod 1 invariant distributed sequences. We need some notations.

For a real-valued function f we denote its domain of definition by D_f, furthermore the set graf f is defined by

$$\operatorname{graf} f = \{(x,y) \mid x \in D_f,\ y = f(x)\}.$$

We write graf f mod 1 instead of graf f if we reduce the coordinates x, y mod 1:

$$\operatorname{graf} f \bmod 1 = \{(x,y) \mid \exists k, i \in \mathbf{Z},\ x + k \in D_f,\ y + i = f(x+k) \text{ and } 0 \leqq x, y < 1\}.$$

We shall prove the following results.

THEOREM 1. *Let f be a real-valued function with $D_f \subseteqq \mathbf{R}$ and let F and G be non-decreasing functions defined on the interval $[0,1]$ such that $F(0) = G(0) = 0$ and $F(1) = G(1) = 1$. For any functions F and G there exists a* *sequence* $\omega = (a_n)_{n=1}^{\infty}$ *of real numbers, such that the asymptotic distribution function* mod 1 *of* ω *is* F *and the a.d.f.* mod 1 *of the sequence* $\varphi = (f(a_n))_{n=1}^{\infty}$ *is* G *if and only if* graf f mod 1 *is everywhere dense on the unit square* $0 \leqq x < 1$, $0 \leqq y < 1$.

From this theorem, with $F = G$, we obtain

COROLLARY. *If* f *is a real-valued function such that* graf f mod 1 *is everywhere dense on the unit square, then for every nondecreasing function* F, *for which* $F(0) = 0$ *and* $F(1) = 1$, *there exist a sequence* ω *of real numbers such that* F *is the asymptotic distribution function* modulo 1 *of* ω *and* ω *is* mod 1 f *invariant distributed.*

THEOREM 2. *Let* f *be a real-valued function. There exists a sequence* $\omega = (u_n)_{n=1}^{\infty}$ *of real numbers, which is uniformly distributed* mod 1 *and* f *invariant distributed* mod 1 *if and only if for any pairwise disjunct finitely many subintervals* $[a_1, b_1)$, $[a_2, b_2)$, $\ldots$, $[a_s, b_s)$ *of* $[0, 1)$ *we have*

$$(1)\quad m^*(\{y \mid \exists k, \exists x \in [a_k, b_k) \Rightarrow (x, y) \in (x, y) \in \text{graf}\, f \text{mod}\, 1\}) \geqq \sum_{i=1}^{s} (b_i - a_i)$$

and

$$(2)\quad m^*(\{x \mid \exists k, \exists y \in [a_k, b_k) \Rightarrow (x, y) \in (x, y) \in \text{graf}\, f \text{mod}\, 1\}) \geqq \sum_{i=1}^{s} (b_i - a_i)$$

where $m^*(H)$ *denotes the Jordan outer measure of the set* H.

NOTES. It is easy to check that the conditions of Theorem 1 for the function f are satisfied by the functions sin, cos, tg, ctg, and by any continuous periodic function for which the period length is irrational and the range of this function is an interval of length at least one. These functions satisfy also the conditions of Theorem 2, moreover, the function $\frac{1}{x}$ also satisfies the conditions.

PROOF OF THEOREM 1. Suppose that graf f mod 1 is everywhere dense on the unit square $0 \leqq x < 1$, $0 \leqq y < 1$ and let F and G be non-decreasing functions defined in the interval $[0, 1]$ with conditions $F(0) = G(0) = 0$ and $F(1) = G(1) = 1$. It is known (see Kuipers and Niederreiter [2], pp. 138–139), that there exist sequences $\sigma = (c_n)_{n=1}^{\infty}$ and $\varrho = (d_n)_{n=1}^{\infty}$ of real numbers with distinct terms in the interval $[0, 1)$ such that their asymptotic distribution functions mod 1 are F and G, respectively. Let us cut the sequence σ into blocks such that the k^{th} block contains terms c_n for which $\frac{(k-1)k}{2} < n \leqq \frac{k(k+1)}{2}$. The k^{th} block has elements of number k and we arrange them according to their magnitude: $\alpha_1^{(k)} < \alpha_2^{(k)} < \ldots < \alpha_k^{(k)}$. Let us cut similarly the sequence ϱ into blocks such that the k^{th} block contain terms d_n for which

$\frac{(k-1)k}{2} < n \leqq \frac{k(k+1)}{2}$ and let us denote the terms of the k^{th} block by $\beta_1^{(k)} < < \beta_2^{(k)} < \ldots < \beta_k^{(k)}$. Since graf f mod 1 is everywhere dense in the unit square $0 \leqq x < 1$, $0 \leqq y < 1$, there is a point $P_j^{(k)}(x_j^{(k)}, y_j^{(k)})$ of graf f mod 1 such that $\alpha_j^{(k)} < x_j^{(k)} < \alpha_{j+1}^{(k)}$ and $\beta_j^{(k)} < y_j^{(k)} < \beta_{j+1}^{(k)}$ for any j with $1 \leqq j \leqq k-1$.

Let $\xi = (v_n)_{n=1}^{\infty}$ and $\nu = (z_n)_{n=1}^{\infty}$ be sequences of real numbers defined by

$$\xi = \{x_1^{(1)}, x_1^{(2)}, x_2^{(2)}, x_1^{(3)}, x_2^{(3)}, x_3^{(3)}, \ldots, x_1^{(k)}, x_2^{(k)}, \ldots, x_k^{(k)}, \ldots\}$$

and

$$\nu = \{y_1^{(1)}, y_1^{(2)}, y_2^{(2)}, y_1^{(3)}, y_2^{(3)}, y_3^{(3)}, \ldots, y_1^{(k)}, y_2^{(k)}, \ldots, y_k^{(k)}, \ldots\},$$

respectively. We shall prove that the a.d.f. mod 1 of sequence ξ is the function F.

First we prove that

$$\lim_{k\to\infty} \frac{2}{k(k+1)} A\Big([0,\alpha), \frac{k(k+1)}{2}, \xi\Big) = F(\alpha)$$

for any $0 \leqq \alpha < 1$.

Let us compare the k^{th} blocks of the sequences σ and ξ. The k^{th} block $\alpha_1^{(k)}, \alpha_2^{(k)}, \ldots, \alpha_k^{(k)}$ of σ contains of the same or one more as the k^{th} block $x_1^{(k)}, x_2^{(k)}, \ldots, x_k^{(k)}$ of ξ in the interval $[0,\alpha)$. So

$$0 \leqq A\Big([0,\alpha), \frac{k(k+1)}{2}, \sigma\Big) - A\Big([0,\alpha), \frac{k(k+1)}{2}, \xi\Big) \leqq k,$$

from which

$$\lim_{k\to\infty} \frac{2}{k(k+1)} A\Big([0,\alpha), \frac{k(k+1)}{2}, \sigma\Big) =$$
$$= \lim_{k\to\infty} \frac{2}{k(k+1)} A\Big([0,\alpha), \frac{k(k+1)}{2}, \xi\Big) = F(\alpha)$$

follows. Let N be a positive integer and k be an integer defined by $\frac{k(k+1)}{2} \leqq \leqq N < \frac{(k+1)(k+2)}{2}$. Then, by

$$0 \leqq A([0,\alpha), N, \xi) - A\Big([0,\alpha), \frac{k(k+1)}{2}, \xi\Big) \leqq k+1,$$

we have

$$\lim_{N\to\infty} \frac{1}{N} A([0,\alpha), N, \xi) =$$
$$= \lim_{k\to\infty} \frac{2}{k(k+1)} A\Big([0,\alpha), \frac{k(k+1)}{2}, \xi\Big) = F(\alpha)$$

and so F really is the a.d.f. mod 1 of the sequence $\xi = (v_n)_{n=1}^{\infty}$.

It can similarly be proved that G is the a.d.f. mod 1 of the sequence ν.

Since $P_n(v_n, z_n) \in \operatorname{graf} f$ mod 1, there are integers s_n and t_n such that $f(v_n+s_n) = z_n+t_n$ for any natural number n. But then F and G are the a.d.f. mod 1 of the sequences $\omega = (a_n)_{n=1}^{\infty} = (v_n + s_n)_{n=1}^{\infty}$ and $\varphi = (f(a_n))_{n=1}^{\infty} = (z_n + t_n)_{n=1}^{\infty}$, respectively, and so we have proved that the condition of the theorem is sufficient.

Now we prove the necessity of the conditions.

Suppose that graf f mod 1 is not everywhere dense in the unit square $0 \leqq x < 1$, $0 \leqq y < 1$. Then there is a point $P_0(x_0, y_0)$, with $0 \leqq x_0, y_0 < 1$, which has a domain of radius r, with some $r > 0$, inside of the unit square not containing point of graf f mod 1. In this case for the function F and G defined by

$$F(x) = \begin{cases} 0 & \text{if } 0 \leqq x \leqq x_0 \\ 1 & \text{if } x_0 < x \leqq 1 \end{cases}$$

and

$$G(y) = \begin{cases} 0 & \text{if } 0 \leqq y \leqq y_0 \\ 1 & \text{if } y_0 < y \leqq 1, \end{cases}$$

respectively, there are no sequences $\omega = (a_n)_{n=1}^{\infty}$ and $\varphi = (f(a_n))_{n=1}^{\infty}$ such that their a.d.f. mod 1 would be F and G, respectively. Namely if F is the a.d.f. mod 1 of a sequence $\omega = (a_n)_{n=1}^{\infty}$ and φ is a sequence defined by $\varphi = (f(a_n))_{n=1}^{\infty}$ then

$$\lim_{N\to\infty} \frac{1}{N} A\left(\left[x_0, x_0 + \frac{r}{2}\right), N, \omega\right) = 1$$

and

$$\lim_{N\to\infty} \frac{1}{N} A\left(\left[y_0, y_0 + \frac{r}{2}\right), N, \varphi\right) = 0$$

and so G cannot be the a.d.f. mod 1 of the sequence of φ.

This contradiction completes the proof of the theorem.

Proof of Theorem 2. First we prove the necessity of the conditions.

Let f be a real-valued function and let $\omega = (u_n)_{n=1}^{\infty}$ be a u.d. mod 1 sequence such that $\varphi = (f(u_n))_{n=1}^{\infty}$ is also u.d. mod 1. Let $[a_1, b_1), [a_2, b_2), \ldots, [a_s, b_s)$ be subintervals of $[0, 1)$ with $0 \leqq a_1 < b_1 \leqq a_2 < b_2 \leqq \ldots < 1$ and let

$$H_1 = \{y \mid \exists i, \exists x \in [a_i, b_i); (x, y) \in \operatorname{graf} f \bmod 1\}$$

and

$$H_2 = \{x \mid \exists i, \exists y \in [a_i, b_i); (x, y) \in \operatorname{graf} f \bmod 1\}.$$

Suppose that

$$m^*(H_1) < \sum_{i=1}^{s} (b_i - a_i).$$

Then there exist finitely many subintervals $[c_1, d_1), [c_2, d_2), \dots, [c_k, d_k)$ of $[0,1)$ which cover the set H_1 and for the sum of their length we have

$$\sum_{i=1}^{k}(d_i - c_i) < \sum_{j=1}^{s}(b_j - a_j).$$

Since the sequence ω is u.d. mod 1,

$$\lim_{N\to\infty} \frac{1}{N} A([a_j, b_j), N, \omega) = b_j - a_j$$

for any $j = 1, 2, \dots, s$ and so

$$\lim_{N\to\infty} \frac{1}{N} A\Big(\bigcup_{j=1}^{s} [a_j, b_j), N, \omega\Big) = \sum_{j=1}^{s}(b_j - a_j),$$

since the intervals $[a_j, b_j)$, $j = 1, 2, \dots, s$, are pairwise disjunct.

By our conditions, using that from $\{u_n\} \in [a_i, b_i)$, $\{f(u_n)\} \in H_1$ follows, we obtain

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{k} A([c_i, d_i), N, \varphi) \geqq \sum_{j=1}^{s}(b_j - a_j) > \sum_{i=1}^{k}(d_i - c_i).$$

These imply that there is an index i_0 such that

$$\lim_{N\to\infty} \frac{1}{N} A([c_{i_0}, d_{i_0}), N, \varphi) > d_{i_0} - c_{i_0}$$

and so the sequence $\varphi = (f(u_n))_{n=1}^{\infty}$ is not u.d. mod 1, which is a contradiction. It proves the necessity of (1). The necessity of condition (2) can be proved similarly.

Let now f be a real-valued function satisfying the conditions. Let us cut the unit square $0 \leqq x < 1$, $0 \leqq y < 1$ into squares $\Gamma_{ij}^{(k)}$ of number k^2 such that $\Gamma_{ij}^{(k)}$, $1 \leqq i, j \leqq k$, contain points (x, y) with $\frac{i-1}{k} \leqq x < \frac{i}{k}$ and $\frac{j-1}{k} \leqq y < \frac{j}{k}$. This partition of the unit square and the function f define a $k \times k$ matrix $\Gamma^{(k)}$ of elements γ_{ij}, $1 \leqq i, j \leqq k$ such that $\gamma_{ij} = 1$, if there is a point $(x, y) \in \Gamma_{ij}^{(k)}$ for which $(x, y) \in \operatorname{graf} f$ mod 1, and $\gamma_{ij} = 0$ if $(x, y) \notin \operatorname{graf} f$ mod 1 for any $(x, y) \in \Gamma_{ij}^{(k)}$.

By (1) and (2), for any s columns of $\Gamma^{(k)}$ there exist s rows such that each of them has at least one non-zero common element with one of the columns of number s, furthermore for any s rows there exist s columns such that each of them has at least one non-zero common element with one of the rows of number s. Therefore we can choose elements $\gamma_{ij} = 1$ of matrix $\Gamma^{(k)}$

which number is k and which are from distinct rows and distinct columns. For each of these elements γ_{ij} we choose a point $(x_{ij}^{(k)}, y_{ij}^{(k)})$ from $\Gamma_{ij}^{(k)}$ such that $(x_{ij}^{(k)}, y_{ij}^{(k)}) \in \operatorname{graf} f \bmod 1$. The numbers $x_{ij}^{(k)}$ are obviously distinct and let

$$\omega_1^{(k)} < \omega_2^{(k)} < \ldots < \omega_k^{(k)}$$

be their arrangement.

We denote by

$$\delta_1^{(k)}, \delta_2^{(k)}, \ldots, \delta_k^{(k)}$$

the corresponding sequence of y values. By the definition of the set graf f mod 1 there are integers $p_{ij}^{(k)}$, $q_{ij}^{(k)}$ such that $(x_{ij}^{(k)} + p_{ij}^{(k)}, y_{ij}^{(k)} + q_{ij}^{(k)}) \in \operatorname{graf} f$. For any $t = 1, 2, \ldots, k$, $\omega_t^{(k)}$ is equal to some $x_{ij}^{(k)}$, so we can define the numbers $u_t^{(k)}$ by $u_t^{(k)} = x_{ij}^{(k)} + p_{ij}^{(k)}$, where $x_{ij}^{(k)} = \omega_t^{(k)}$ and $p_{ij}^{(k)}$ is defined above.

Let us consider the sequence

$$u_1^{(1)}, u_1^{(2)}, u_2^{(2)}, u_1^{(3)}, u_2^{(3)}, u_3^{(3)}, \ldots, u_1^{(k)}, u_2^{(k)}, \ldots, u_k^{(k)}, \ldots \tag{3}$$

and

$$\delta_1^{(1)}, \delta_1^{(2)}, \delta_2^{(2)}, \delta_1^{(3)}, \delta_2^{(3)}, \delta_3^{(3)}, \ldots, \delta_1^{(k)}, \delta_2^{(k)}, \ldots, \delta_k^{(k)}, \ldots . \tag{4}$$

Denote the n^{th} term of (3) and (4) by u_n and δ_n, respectively. We shall show that both of the sequences $\omega = (u_n)_{n=1}^{\infty}$ and $\varphi = (f(u_n))_{n=1}^{\infty}$ are uniformly distributed mod 1.

We shall use the sequence $\mu = (\mu_n)_{n=1}^{\infty}$ with terms

$$\frac{0}{1}, \frac{0}{2}, \frac{1}{2}, \frac{0}{3}, \frac{1}{3}, \frac{2}{3}, \frac{0}{4}, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, \ldots,$$

which is u.d. mod 1 (see [2], p. 7, Problem 1.13). Let us cut the sequences μ, ω, φ into blocks such that the m^{th} block contain elements of index n for which $\frac{(m-1)m}{2} < n < \frac{m(m+1)}{2}$. By the definition of the sequences

$$\mu_1^{(m)} \leqq \omega_1^{(m)} < \mu_2^{(m)} \leqq \omega_2^{(m)} < \ldots < \mu_m^{(m)} \leqq \omega_m^{(m)}$$

and

$$\mu_1^{(m)} \leqq \delta_{i_1}^{(m)} < \mu_2^{(m)} \leqq \delta_{i_2}^{(m)} < \ldots < \mu_m^{(m)} \leqq \delta_{i_m}^{(m)}$$

follows, where $\mu_i^{(m)}$ is the i^{th} term of the m^{th} block of the sequence μ and $(i_1, i_2, \ldots, i_m)$ is a permutation of $(1, 2, \ldots, m)$, i.e. $\delta_{i_1}^{(m)}, \delta_{i_2}^{(m)}, \ldots, \delta_{i_m}^{(m)}$ is a permutation of the terms of the m^{th} block of sequence φ reduced modulo 1.

Let N (> 2) be an integer and let k be a natural number defined by $\frac{k(k+1)}{2} \leqq N < \frac{(k+1)(k+2)}{2}$. For a real number α, $0 < \alpha \leqq 1$, denote by $A(\alpha, m, z)$

the number of terms z_n of a sequence $z=(z_n)_{n=1}^{\infty}$ which belong to the m^{th} block of z and for which $0 \leqq z_n < \alpha$. Then

$$A(\alpha, m, \mu) - A(\alpha, m, \omega) = 0 \quad \text{or} \quad 1$$

and

$$A(\alpha, m, \mu) - A(\alpha, m, \varphi) = 0 \quad \text{or} \quad 1$$

for $m = 1, 2, \dots, k$ and similar relation holds for sequences μ and ω if we consider the terms from the last term of the k^{th} block till the N^{th} term. But the terms of the sequence $(\delta_n)_{n=1}^{\infty}$ are not arranged by their magnitude, so we have

$$0 \leqq A([0,\alpha); N; \mu) - A([0,\alpha); N; \omega) \leqq k+1$$

and

$$0 \leqq A([0,\alpha); N; \mu) - A([0,\alpha); N; \varphi) \leqq k+k+1.$$

From these

$$\lim_{N\to\infty} \frac{1}{N} A([0,\alpha); N; \omega) = \lim_{N\to\infty} \frac{1}{N} A([0,\alpha); N; \mu)$$

and

$$\lim_{N\to\infty} \frac{1}{N} A([0,\alpha); N; \varphi) = \lim_{N\to\infty} \frac{1}{N} A([0,\alpha); N; \mu)$$

follow. However, the sequence μ is u.d. mod 1 and so the sequences ω and φ are also u.d. mod 1 which completes the proof of Theorem 2.

REFERENCES

[1] KUIPERS, L., Remark on a paper by R. L. Duncan concerning the uniform distribution mod 1 of the sequence of the logarithms of the Fibonacci numbers, *Fibonacci Quart.* 7 (1969), 465–466, 473. *MR* **41**#1656

[2] KUIPERS, L. and NIEDERREITER, H., *Uniform distribution of sequences*, Pure and Applied Mathematics, J. Wiley, New York, 1974. *MR* **54** #7415

[3] MOLNÁR, S. H., Reciprocal invariant distributed sequences constructed by second order linear recurrences (to appear).

[4] PORUBSKY, S., SALÁT, T. and STRAUCH, O., Transformations that preserve uniform distribution, *Acta Arith.* **49** (1988), 459–479. *MR* **89m**:11072

(Received April 4, 1991)

PÉNZÜGYI ÉS SZÁMVITELI FŐISKOLA
BUZOGÁNY U. 10
H–1149 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 323–340

DIE DÜNNSTE 8-FACHE GITTERFÖRMIGE KREISÜBERDECKUNG DER EBENE

ÁGOTA H. TEMESVÁRI

Eine Menge von abgeschlossenen Kreisen bildet eine k-fache Überdeckung, wenn jeder Punkt der Ebene zu mindestens k Kreisen gehört. Die Definition der Dichte von k-fachen Kreisüberdeckungen kann man z.B. in [2] finden. Die Grundaufgabe ist die Bestimmung der k-fachen Kreisüberdeckung mit der minimalen Dichte (wenn eine solche Überdeckung existiert), d.h., die Bestimmung der dünnsten k-fachen Kreisüberdeckung. Wenn die Kreisüberdeckung gitterförmig ist, d.h., die Kreismittelpunkte ein ebenes Punktgitter bilden, dann ist die Grundaufgabe für $k \leqq 7$ in den Artikeln [4], [1], [6], [7], [8], [3], [9] gelöst. Im nicht gitterförmigen Fall ist die dünnste k-fache Kreisüberdeckung nur für $k = 1$ [4] und mit kongruenten Kreisen bekannt.

Von J. Linhart stammt eine Methode, mit deren Hilfe er etwa mit drei- bis vierstelligen Genauigkeit mit einer Rechnenmaschine die minimale Dichte der k-fachen gitterförmigen Kreisüberdeckungen ausrechnen kann. In [5] beschreibt er gute Vermutungen über die exakten Werte der minimalen Dichte für $k \leqq 20$ mit dieser Methode.

In diesem Artikel beweisen wir, daß seine Vermutung für $k = 8$ richtig ist.

Zuerst führen wir die Bezeichnungen (s. noch in [8]) ein. Es sei Γ ein ebenes Punktgitter mit den Basisvektoren $\overrightarrow{OA}$ und $\overrightarrow{OB}$. k und m seien reelle Zahlen. Den Vektor $\overrightarrow{OX}$ bzw. $k\overrightarrow{OX} + m\overrightarrow{OY}$ und seinen Endpunkt bezeichnen wir mit X bzw. $kX + mY$. Γ ist in normaler Darstellung, wenn die Ungleichungen

$$(1) \qquad |A| \leqq |B| \leqq |B - A|, \qquad \sphericalangle(AOB) \leqq \frac{\pi}{2}$$

für seine Basisvektoren gelten. Es seien $|A| = a$, $|B| = b$, $\frac{a}{b} = x$ und $\sphericalangle(AOB) = \alpha$ (Abb. 1). Mit diesen Bezeichnungen kann man (1) folgenderweise aufschreiben:

$$(2) \qquad 0 < x \leqq 1, \qquad 0 \leqq \cos x \leqq \frac{x}{2}.$$

1991 *Mathematics Subject Classification.* Primary 52C15; Secondary 11H31.
Key words and phrases. k-fold covering, dense, thinnest lattice covering.

0081–6906/94/$ 4.00

Es sei $T(\Gamma)$ der Inhalt des Grundparallelogramms von Γ.

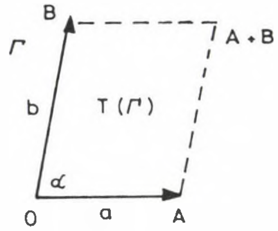


Abb. 1

Der Kreis, der durch die nicht kollinearen Punkte X, Y, Z bestimmt ist, sei $k[XYZ]$. Später sind die folgenden, durch Gitterdreiecke bestimmten Kreise nötig: $k_1 = k[O(8A)(4A+B)]$, $k_2 = k[O(7A)(4A+B)]$, $k_3 = k[O(6A)(4A+B)]$, $k_4 = k[O(5A)(4A+B)]$, $k_5 = k[O(4A)(A+2B)]$, $k_6 = k[O(4A)(2A+2B)]$, $k_7 = k[O(3A)(A+2B)]$, $k_8 = k[O(3A)(2A+2B)]$, $k_9 = k[O(2A)(2B)]$, $k_{10} = k[O(2A)(A+2B)]$, $k_{11} = k[O(2A)(3B)]$, $k_{12} = k[O(2A)(A+3B)]$, $k_{13} = k[OA(3B)]$, $k_{14} = k[OA(3B-A)]$, $k_{15} = k[O(4A-B)(4A+B)]$, $k_{16} = k[O(3A-B)(4A+B)]$, $k_{17} = k[O(2A-B)(4A+B)]$, $k_{18} = k[O(3A-B)(3A+B)]$, $k_{19} = k[O(A-B)(3A+B)]$, $k_{20} = k[O(3A-2B)(A+B)]$, $k_{21} = k[O(3A-2B)(2A+B)]$, $k_{22} = k[O(2A-B)(2A+2B)]$, $k_{23} = k[O(3A-B)(A+2B)]$.

Mit D_i $(i = 1, \dots, 23)$ bezeichnen wir das Gitterdreieck, das den Kreis k_i im vorigen bestimmt hat. $\widehat{k}[XYZ]$ bzw. $\widehat{k}_i$ sei der Rand des Kreises $k[XYZ]$ bzw. k_i. Mit $k[0,r]$ bzw. $\widehat{k}[0,r]$ bezeichnen wir den Kreis bzw. die Kreislinie mit dem Mittelpunkt O und mit dem Radius r. Der Radius des Kreises $k[XYZ]$ bzw. k_i sei $r[XYZ]$ bzw. r_i. Endlich, mit Q_i bezeichnen wir den Quotient $r_i^2/T(\Gamma)$.

Auch hier gebrauchen wir die Formel

$$r = \frac{uvw}{4T} \tag{3}$$

oft, wobei u, v, w die Seiten eines Dreiecks mit Inhalt T und r der Radius des Umkreises sind.

Das Gitter Γ_8 (mit den Basisvektoren A und B) wird folgenderweise definiert: $|A| = |B| = \frac{\sqrt{6}}{4}$ und $\cos\alpha = \frac{|A|}{4|B|} = \frac{1}{4}$ (Abb. 2). Es ist leicht einzusehen, daß $r[O(3A)(2B)] = r_8 = r_{11} = r_{20} = r_{22} = 1$ gilt.

Mit $L(\Gamma, R)$ bezeichnen wir die Kreisanordnung, bei der Γ das Gitter der Kreismittelpunkte und R der Radius der Kreise sind.

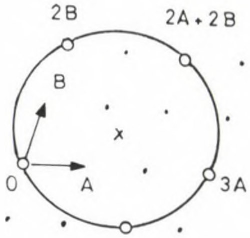


Abb. 2

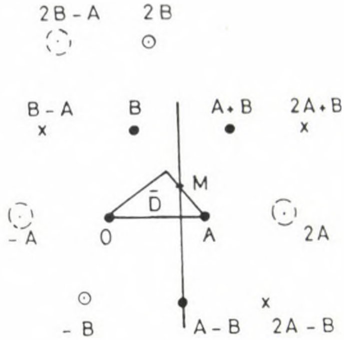


Abb. 3

HILFSSATZ 1. *Die Kreisanordnung* $L(\Gamma_8, 1)$ *ist eine 8-fache gitterförmige Überdeckung von Einheitskreisen.*

BEWEIS. Nun betrachten wir die Abbildung 3. Weil $x = 1$ für Γ gilt, d.h., das Viereck $OA(A+B)B$ ein Rhombus ist, ist es genug zu beweisen, daß das rechtwinklige Dreieck $OA(\frac{A+B}{2})$ mindestens 8-fach überdeckt ist. Dieses Dreieck wird mit $\overline{D}$ bezeichnet.

Es gilt $|A+B| = \frac{\sqrt{15}}{4} < 1$, deshalb überdecken $k[0,1]$, $k[A,1]$, $k[B,1]$ und $k[A+B,1]$ das Dreieck $\overline{D}$. Aus dem rechtwinkligen Dreieck $(A-B)O(\frac{A+B}{2})$ ergibt sich $|\frac{A+B}{2} - (A-B)| = \frac{\sqrt{51}}{8} < 1$, d.h., auch $k[A-B,1]$ überdeckt $\overline{D}$.

Es sei M der Schnittpunkt der Gerade AB und der Gerade, die durch $A-B$ verläuft und zu OA orthogonal ist. Es ist leicht zu sehen, daß $|M-(-B)| = |M-(2A-B)| = \sqrt{\frac{7}{8}} < 1$ gilt. Weil $r[(-B)(2B)(2A-B)] = 1$ ist, überdecken die Kreise $k[-B,1]$ und $k[2B,1]$ das Dreieck $\overline{D}$ gemeinsam mindestens einfach. Ähnlicherweise kann man zeigen, daß auch $k[B-A,1]$, $k[B+2A,1]$ und $k[2A-B,1]$ das Dreieck $\overline{D}$ gemeinsam einfach überdecken.

Das Dreieck $(-A)(2A)(-A+2B)$ enthält $\overline{D}$ und wegen

$$r[(-A)(2A)(-A+2B)] = 1$$

überdecken $k[-A,1]$, $k[2A,1]$ und $k[-A+2B,1]$ das Dreieck $\overline{D}$ gemeinsam einfach.

Damit haben wir den Hilfssatz bewiesen.

In dieser Arbeit sehen wir den folgenden Satz ein.

SATZ. *Die Dichte einer 8-fachen gitterförmigen Überdeckung von Einheitskreisen ist* $\geq \frac{32}{3\sqrt{15}}\pi \sim 2,7541 \cdot \pi$ $(\sim 8,6523)$ *und Gleichheit tritt nur bei der Kreisüberdeckung* $L(\Gamma_8, 1)$ *auf.*

Der Beweis des Satzes geschieht mit der Methode, die wir auch in [7], [8] und [9] verwandt haben. Auch hier betrachten wir das rechtwinklige

Koordinatensystem $x, y = \cos\alpha$. Jedem Gitter Γ in normaler Darstellung entspricht ein Punkt mit den Koordinaten $(x, \cos\alpha)$, der im rechtwinkligen Dreieck OPQ liegt, wo $OP = 1$, $PQ = \frac{1}{2}$ und $OP \perp PQ$ sind.

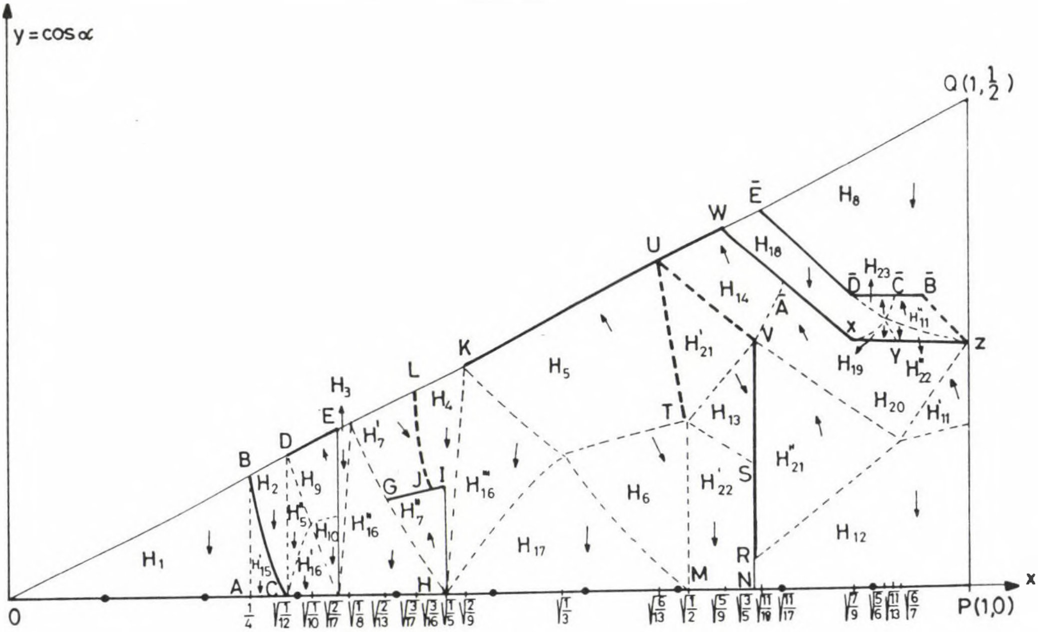


Abb. 4

Betrachten wir die folgenden Mengen (Abb. 4):

$$H_1 := \left\{ (x, \cos\alpha) \mid x \in \left(0, \frac{1}{4}\right], \ \cos\alpha \in \left[0, \frac{x}{2}\right] \right\};$$

$$H_2 := \left\{ (x, \cos\alpha) \mid x \in \left[\frac{1}{4}, \sqrt{\frac{1}{12}}\right], \ \cos\alpha \in \left[\frac{1-12x^2}{8x}, \frac{x}{2}\right] \right\};$$

$$H_3 := \left\{ (x, \cos\alpha) \mid x \in \left[\sqrt{\frac{2}{17}}, \sqrt{\frac{1}{8}}\right], \ \cos\alpha \in \left[\frac{17x^2-2}{2x}, \frac{x}{2}\right] \right\};$$

$$H_4 := \left\{ (x, \cos\alpha) \mid x \in \left[\sqrt{\frac{3}{17}}, \sqrt{\frac{3}{16}}\right], \ \cos\alpha \in \left[\frac{3-15x^2}{4x}, \frac{x}{2}\right] \right. \text{ oder}$$

$$x \in \left[\sqrt{\frac{3}{16}}, \sqrt{\frac{1}{5}}\right], \ \cos\alpha \in \left[\frac{x}{4}, \frac{x}{2}\right] \text{ oder}$$

$$\left. x \in \left[\sqrt{\frac{1}{5}}, \sqrt{\frac{2}{9}}\right], \ \cos\alpha \in \left[\frac{5x^2-1}{x}, \frac{x}{2}\right] \right\};$$

$$H_5 := \left\{ (x, \cos\alpha) \mid x \in \left[\sqrt{\frac{2}{9}}, \sqrt{\frac{1}{3}}\right], \ \cos\alpha \in \left[\frac{2-3x^2}{12x}, \frac{x}{2}\right] \right. \text{ oder}$$

$$x \in \left[\sqrt{\frac{1}{3}}, \sqrt{\frac{6}{13}}\right], \ \cos\alpha \in \left[\frac{x}{4}, \frac{x}{2}\right] \text{ oder}$$

$$x \in \left[\sqrt{\frac{6}{13}}, \sqrt{\frac{1}{2}}\right], \ \cos\alpha \in \left[\frac{x}{4}, \frac{6-11x^2}{4x}\right]\Big\};$$

$$H_6 := \Big\{(x, \cos\alpha) \mid x \in \left[\sqrt{\frac{1}{3}}, \sqrt{\frac{1}{2}}\right], \ \cos\alpha \in \left[\frac{1-2x^2}{4x}, \frac{x}{4}\right]\Big\};$$

$$H_7' := \Big\{(x, \cos\alpha) \mid x \in \left[\sqrt{\frac{1}{8}}, \sqrt{\frac{2}{13}}\right], \ \cos\alpha \in \left[\frac{1-5x^2}{6x}, \frac{x}{2}\right] \text{ oder}$$

$$x \in \left[\sqrt{\frac{2}{13}}, \sqrt{\frac{3}{17}}\right], \ \cos\alpha \in \left[\frac{x}{4}, \frac{x}{2}\right] \text{ oder}$$

$$x \in \left[\sqrt{\frac{3}{17}}, \sqrt{\frac{3}{16}}\right], \ \cos\alpha \in \left[\frac{x}{4}, \frac{3-15x^2}{4x}\right]\Big\};$$

$$H_7'' := \Big\{(x, \cos\alpha) \mid x \in \left[\sqrt{\frac{2}{13}}, \sqrt{\frac{1}{5}}\right], \ \cos\alpha \in \left[\frac{1-5x^2}{6x}, \frac{x}{4}\right]\Big\};$$

$$\overline{H}_8 := \Big\{(x, \cos\alpha) \mid x \in \left[\sqrt{\frac{11}{18}}, \sqrt{\frac{5}{6}}\right], \ \cos\alpha \in \left[\frac{11-6x^2}{24x}, \frac{x}{2}\right] \text{ oder}$$

$$x \in \left[\sqrt{\frac{5}{6}}, 1\right], \ \cos\alpha \in \left[\frac{1}{4x}, \frac{x}{2}\right]\Big\};$$

$$H_9 := \Big\{(x, \cos\alpha) \mid x \in \left[\sqrt{\frac{1}{12}}, \sqrt{\frac{1}{10}}\right], \ \cos\alpha \in \left[\frac{1-8x^2}{8x}, \frac{x}{2}\right] \text{ oder}$$

$$x \in \left[\sqrt{\frac{1}{10}}, \sqrt{\frac{2}{17}}\right], \ \cos\alpha \in \left[\frac{x}{4}, \frac{x}{2}\right]\Big\};$$

$$H_{10} := \Big\{(x, \cos\alpha) \mid x \in \left[\sqrt{\frac{1}{10}}, \sqrt{\frac{2}{17}}\right], \ \cos\alpha \in \left[\frac{2-17x^2}{12x}, \frac{x}{4}\right]\Big\};$$

$$H_{11}' := \Big\{(x, \cos\alpha) \mid x \in \left[\sqrt{\frac{6}{7}}, 1\right], \ \cos\alpha \in \left[\frac{x}{6}, \frac{3x^2-2}{4x}\right]\Big\}.$$

Es ist leicht einzusehen, daß die Gleichung

$$0,3 = \frac{8 - x^2 - \sqrt{9x^4 - 8x^2 + 24}}{8x} \quad \text{im Fall } x \in \left[\frac{0,6+\sqrt{24,36}}{6}, 1\right]$$

eine einzige Wurzel hat. Diese Wurzel wird mit x_k bezeichnet.

$$H_{11}'' := \Big\{(x, \cos\alpha) \mid x \in \left[\sqrt{\frac{5}{6}}, \frac{0,6+\sqrt{24,36}}{6}\right], \ \cos\alpha \in \left[\frac{1}{4x}, \frac{3x^2-2}{2x}\right] \text{ oder}$$

$$x \in \left[\frac{0,6+\sqrt{24,36}}{6}, x_k\right], \ \cos\alpha \in \left[\frac{1}{4x}, 0,3\right] \text{ oder}$$

$$x \in [x_k, 1], \cos\alpha \in \left[\frac{1}{4x}, \frac{8-x^2-\sqrt{9x^4-8x^2+24}}{8x}\right]\Big\};$$

$$H_{12} := \Big\{(x, \cos\alpha) \mid x \in \left[\sqrt{\frac{3}{5}}, \sqrt{\frac{6}{7}}\right], \cos\alpha \in \left[0, \frac{11x^2-6}{24x}\right] \text{ oder}$$

$$x \in \left[\sqrt{\frac{6}{7}}, 1\right], \ \cos\alpha \in \left[0, \frac{x}{6}\right]\Big\};$$

$$H_{13} := \Big\{(x, \cos\alpha) \mid x \in \left[\sqrt{\frac{1}{2}}, \sqrt{\frac{3}{5}}\right], \cos\alpha \in \left[\frac{1-x^2}{4x}, \frac{3x^2-1}{4x}\right]\Big\};$$

$$H_{14} := \Big\{(x, \cos\alpha) \mid x \in \left[\sqrt{\frac{6}{13}}, \sqrt{\frac{5}{9}}\right], \ \cos\alpha \in \left[\frac{3-2x^2}{9x}, \frac{x}{2}\right] \ \text{oder}$$
$$x \in \left[\sqrt{\frac{5}{9}}, \sqrt{\frac{3}{5}}\right], \ \cos\alpha \in \left[\frac{3-2x^2}{9x}, \frac{5-3x^2}{12x}\right] \ \text{oder}$$
$$x \in \left[\sqrt{\frac{3}{5}}, \sqrt{\frac{11}{17}}\right], \ \cos\alpha \in \left[\frac{7x^2-3}{6x}, \frac{5-3x^2}{12x}\right]\Big\};$$

$$H'_{15} := \Big\{(x, \cos\alpha) \mid x \in \left[\frac{1}{4}, \sqrt{\frac{1}{12}}\right], \ \cos\alpha \in \left[0, \frac{1-12x^2}{8x}\right]\Big\};$$

$$H''_{15} := \Big\{(x, \cos\alpha) \mid x \in \left[\sqrt{\frac{1}{12}}, \sqrt{\frac{1}{10}}\right], \ \cos\alpha \in \left[\frac{12x^2-1}{8x}, \frac{1-8x^2}{8x}\right]\Big\};$$

$$H'_{16} := \Big\{(x, \cos\alpha) \mid x \in \left[\sqrt{\frac{1}{12}}, \sqrt{\frac{1}{10}}\right], \ \cos\alpha \in \left[0, \frac{12x^2-1}{8x}\right] \ \text{oder}$$
$$x \in \left[\sqrt{\frac{1}{10}}, \sqrt{\frac{2}{17}}\right], \ \cos\alpha \in \left[0, \frac{2-17x^2}{12x}\right]\Big\};$$

$$H''_{16} := \Big\{(x, \cos\alpha) \mid x \in \left[\sqrt{\frac{2}{17}}, \sqrt{\frac{1}{8}}\right], \ \cos\alpha \in \left[0, \frac{17x^2-2}{2x}\right] \ \text{oder}$$
$$x \in \left[\sqrt{\frac{1}{8}}, \sqrt{\frac{1}{5}}\right], \ \cos\alpha \in \left[0, \frac{1-5x^2}{6x}\right]\Big\};$$

$$H'''_{16} := \Big\{(x, \cos\alpha) \mid x \in \left[\sqrt{\frac{1}{5}}, \sqrt{\frac{2}{9}}\right], \ \cos\alpha \in \left[\frac{5x^2-1}{8x}, \frac{5x^2-1}{x}\right] \ \text{oder}$$
$$x \in \Big[\sqrt{\frac{2}{9}}, \sqrt{\frac{1}{3}}, \ \cos\alpha \in \left[\frac{5x^2-1}{8x}, \frac{2-3x^2}{12x}\right]\Big\};$$

$$H_{17} := \Big\{(x, \cos\alpha) \mid x \in \left[\sqrt{\frac{1}{5}}, \sqrt{\frac{1}{3}}\right], \cos\alpha \in \left[0, \frac{5x^2-1}{8x}\right] \ \text{oder}$$
$$x \in \left[\sqrt{\frac{1}{3}}, \sqrt{\frac{1}{2}}\right], \ \cos\alpha \in \left[0, \frac{1-2x^2}{4x}\right]\Big\};$$

$$H_{18} := \Big\{(x, \cos\alpha) \mid x \in \left[\sqrt{\frac{5}{9}}, \sqrt{\frac{11}{18}}\right], \ \cos\alpha \in \left[\frac{5-3x^2}{12x}, \frac{x}{2}\right] \ \text{oder}$$
$$x \in \left[\sqrt{\frac{11}{18}}, \sqrt{\frac{7}{9}}\right], \ \cos\alpha \in \left[\frac{5-3x^2}{12x}, \frac{11-6x^2}{24x}\right] \ \text{oder}$$
$$x \in \left[\sqrt{\frac{7}{9}}, \sqrt{\frac{5}{6}}\right], \ \cos\alpha \in \left[\frac{3x^2-1}{6x}, \frac{11-6x^2}{24x}\right]\Big\};$$

$$H_{19} := \Big\{ (x, \cos\alpha) \mid x \in \Big[\sqrt{\frac{7}{9}}, \sqrt{\frac{5}{6}}\Big],\ \cos\alpha \in \Big[\frac{1+x^2}{8x}, \frac{3x^2-1}{6x}\Big] \text{ oder}$$
$$x \in \Big[\sqrt{\frac{5}{6}}, \sqrt{\frac{11}{13}}\Big],\ \cos\alpha \in \Big[\frac{1+x^2}{8x}, \frac{3-3x^2}{2x}\Big] \Big\};$$

$$H_{20} := \Big\{ (x, \cos\alpha) \mid x \in \Big[\sqrt{\frac{3}{5}}, \sqrt{\frac{11}{17}}\Big],\ \cos\alpha \in \Big[\frac{3-2x^2}{9x}, \frac{7x^2-3}{6x}\Big] \text{ oder}$$
$$x \in \Big[\sqrt{\frac{11}{17}}, \sqrt{\frac{7}{9}}\Big],\ \cos\alpha \in \Big[\frac{3-2x^2}{9x}, \frac{5-3x^2}{12x}\Big] \text{ oder}$$
$$x \in \Big[\sqrt{\frac{7}{9}}, \sqrt{\frac{6}{7}}\Big],\ \cos\alpha \in \Big[\frac{3-2x^2}{9x}, \frac{1+x^2}{8x}\Big] \text{ oder}$$
$$x \in \Big[\sqrt{\frac{6}{7}}, 1\Big],\ \cos\alpha \in \Big[\frac{3x^2-2}{4x}, \frac{1+x^2}{8x}\Big] \Big\};$$

$$H'_{21} := \Big\{ (x, \cos\alpha) \mid x \in \Big[\sqrt{\frac{6}{13}}, \sqrt{\frac{1}{2}}\Big],\ \cos\alpha \in \Big[\frac{6-11x^2}{4x}, \frac{3-2x^2}{9x}\Big] \text{ oder}$$
$$x \in \Big[\sqrt{\frac{1}{2}}, \sqrt{\frac{3}{5}}\Big],\ \cos\alpha \in \Big[\frac{3x^2-1}{4x}, \frac{3-2x^2}{9x}\Big] \Big\};$$

$$H''_{21} := \Big\{ (x, \cos\alpha) \mid x \in \Big[\sqrt{\frac{3}{5}}, \sqrt{\frac{6}{7}}\Big],\ \cos\alpha \in \Big[\frac{11x^2-6}{24x}, \frac{3-2x^2}{9x}\Big] \Big\};$$

$$H'_{22} := \Big\{ (x, \cos\alpha) \mid x \in \Big[\sqrt{\frac{1}{2}}, \sqrt{\frac{3}{5}}\Big],\ \cos\alpha \in \Big[0, \frac{1-x^2}{4x}\Big] \Big\};$$

$$H''_{22} := \Big\{ (x, \cos\alpha) \mid x \in \Big[\sqrt{\frac{5}{6}}, \sqrt{\frac{11}{13}}\Big], \cos\alpha \in \Big[\frac{3-3x^2}{2x}, \frac{1}{4x}\Big] \text{ oder}$$
$$x \in \Big[\sqrt{\frac{11}{13}}, 1\Big],\ \cos\alpha \in \Big[\frac{1+x^2}{8x}, \frac{1}{4x}\Big] \Big\};$$

$$H_{23} := \Big\{ (x, \cos\alpha) \mid x \in \Big[\frac{-1,8+\sqrt{19,74}}{3}, \sqrt{\frac{5}{6}}\Big],\ \cos\alpha \in \Big[\frac{11-6x^2}{24x}, 0,3\Big] \text{ oder}$$
$$x \in \Big[\sqrt{\frac{5}{6}}, \frac{0,6+\sqrt{24,36}}{6}\Big], \cos\alpha \in \Big[\frac{3x^2-2}{2x}, 0,3\Big] \Big\}.$$

Es seien $H_7 = H'_7 \cup H''_7$, $H_8 = \overline{H}_8 \setminus (H''_{11} \cup H_{23})$, $H_{11} = H'_{11} \cup H''_{11}$, $H_{15} = H'_{15} \cup H''_{15}$, $H_{16} = H'_{16} \cup H''_{16} \cup H'''_{16}$, $H_{21} = H'_{21} \cup H''_{21}$, $H_{22} = H'_{22} \cup H''_{22}$. Es ist nicht schwer zu beweisen, daß die Mengen H_i $(i = 1, \ldots, 23)$ eine Zerlegung des Dreiecks OPQ bedeuten.

Hilfssatz 2. *Wenn* $(x, \cos\alpha) \in H_i$ $(i = 1, \ldots, 23)$ *für das Gitter* Γ *in normaler Darstellung gilt, dann liegen höchstens 7 Gitterpunkte im Inneren von* k_i *und ist das Gitterdreieck* D_i *nicht stumpfwinklig.*

Der Beweis des Hilfssatzes geschieht ebenso wie der Beweis des entsprechenden Hilfssatzes von [7], deshalb sehen wir vom Beweis ab. Wir geben aber die Gitterpunkte, die im Fall $(x, \cos\alpha) \in H_i$ $(i = 1, \dots, 23)$ in k_i liegen können:

$(x, \cos\alpha) \in H_1$	$kA\ k = 1, \dots, 7$
$(x, \cos\alpha) \in H_2$	$kA\ k = 1, \dots 6,\ 3A + B$
$(x, \cos\alpha) \in H_3$	$kA\ k = 1, \dots, 5,\ mA + B\ m = 2, 3$
$(x, \cos\alpha) \in H_4$	$kA\ k = 1, \dots, 4,\ mA + B\ m = 1, 2, 3$
$(x, \cos\alpha) \in H_5$	$kA\ k = 1, 2, 3,\ mA + B\ m = 0, 1, 2, 3$
$(x, \cos\alpha) \in H_6$	$kA\ k = 1, 2, 3,\ mA + B\ m = 0, 1, 2, 3$
$(x, \cos\alpha) \in H_7$	$kA\ k = 1, 2,\ mA + B\ m = -1, \dots, 3$
$(x, \cos\alpha) \in H_8$	$kA\ k = 1, 2,\ mA + B\ m = 0, 1, 2,\ sA + 2B\ s = 0, 1$
$(x, \cos\alpha) \in H_9$	$A,\ kA + B\ k = -2, \dots, 3$
$(x, \cos\alpha) \in H_{10}$	$A,\ kA + B\ k = -2, \dots, 3$
$(x, \cos\alpha) \in H'_{11}$	$A,\ kA + B,\ kA + 2B\ k = 0, 1, 2$
$(x, \cos\alpha) \in H''_{11}$	$A,\ kA + B\ k = 0, 1, 2,\ mA + 2B\ m = 1, 0, 1$
$(x, \cos\alpha) \in H_{12}$	$A,\ kA + B,\ kA + 2B\ k = 0, 1, 2$
$(x, \cos\alpha) \in H_{13}$	$kA + B\ k = -1, \dots, 2,\ mA + 2B\ m = -1, 0, 1$
$(x, \cos\alpha) \in H_{14}$	$kA + B\ k = -1, 0, 1,\ mA + 2B\ m = -2, \dots, 1$
$(x, \cos\alpha) \in H'_{15}$	$kA\ k = 1, \dots, 7$
$(x, \cos\alpha) \in H''_{15}$	$kA\ k = 1, \dots, 6,\ 3A + B$
$(x, \cos\alpha) \in H'_{16}$	$kA\ k = 1, \dots, 6,\ 3A + B$
$(x, \cos\alpha) \in H''_{16}$	$kA\ k = 1, \dots, 5,\ mA + B\ m = 2, 3$
$(x, \cos\alpha) \in H'''_{16}$	$kA\ k = 1, \dots, 4,\ mA + B\ m = 1, 2, 3$
$(x, \cos\alpha) \in H_{17}$	$kA\ k = 1, \dots, 4,\ mA + B\ m = 1, 2, 3$
$(x, \cos\alpha) \in H_{18}$	$kA,\ (k-1)A + B\ k = 1, 2, 3,\ 2A - B$
$(x, \cos\alpha) \in H_{19}$	$kA,\ (k-1)A + B\ k = 1, 2, 3,\ 2A - B$
$(x, \cos\alpha) \in H_{20}$	$kA,\ kA - B\ k = 1, 2, 3,\ 2A - 2B$
$(x, \cos\alpha) \in H'_{21}$	$kA\ k = 1, 2, 3,\ mA - B\ m = 1, \dots, 4$
$(x, \cos\alpha) \in H''_{21}$	$kA,\ kA - B\ k = 1, 2, 3,\ 2A - 2B$
$(x, \cos\alpha) \in H'_{22}$	$kA\ k = 1, 2, 3,\ mA + B\ m = 0, \dots, 3$
$(x, \cos\alpha) \in H''_{22}$	$kA,\ (k-1)A + B\ k = 1, 2, 3,\ A + 2B$
$(x, \cos\alpha) \in H_{23}$	$kA,\ (k-1)A + B\ k = 1, 2, 3,\ 2A - B.$

Wir bemerken, daß das Dreieck D_i nur in einigen Fällen rechtwinklig sein kann. Jeder Fall ist auf dem Rand des entsprechenden Bereiches H_i.

Auch hier wenden wir die in [8] definierten Transformationen g, g^{-1}, g_1 an. Während der Anwendung von g ist der Basisvektor A fest und der Endpunkt des Basisvektors B bewegt sich auf der zu OA parallelen Geraden derart, daß α inzwischen abnimmt. g^{-1} war die inverse Transformation von g. Während der Anwendung von g_1 ist der Basisvektor A fix und der Endpunkt von B dreht sich um 0 so, daß α zunimmt.

Im folgenden brauchen wir noch zwei weitere Transformationen. Bei g_3 sei die Gittergerade $(2A + B)(3A - 2B)$ von Γ fest und wir bewegen 0 auf der zu $(2A + B)(3A - 2B)$ parallelen Geraden mit der Zunahme von $|2A + B|$. Im Fall der Transformation g_4 ist die Gittergerade $(A{+}2B)(3A{-}B)$ fix und

0 bewegt sich auf der zu $(A+2B)(3A-B)$ parallelen Geraden derart, daß $|A+2B|$ zunimmt.

Später wenden wir die vorigen Transformationen höchstens bis der Lage an, bei der das entstehende Gitter in normaler Darstellung ist.

Der folgende Hilfssatz handelt von der Wirkung der vorigen Transformationen auf Q_i $(i=1,\ldots,23)$.

HILFSSATZ 3. *Es sei* Γ *ein Gitter in normaler Darstellung.*

1. *Wenn* $(x,\cos\alpha)\in H_i$ $(i=5,9,11,14,20)$ *bzw.* $(x,\cos\alpha)\in H''_7$ *oder* H''_{21} *für* Γ *gilt, dann nimmt* Q_i, Q_7 *bzw.* Q_{21} *während der Anwendung von* g *ab.*

2. *Wenn* $(x,\cos\alpha)\in H_i$ $(i=6,13)$ *bzw.* $(x,\cos\alpha)\in H'_7$ *für* Γ *gilt, dann können wir* Q_i *bzw.* Q_7 *mit* g^{-1} *vermindern.*

3. *In den Fällen* $(x,\cos\alpha)\in H_i$ $(i=1,2,3,4,8,10,12,15,16,17,18,19,22)$ *nimmt* Q_i *während der Anwendung von* g_1 *ab.*

4. *Endlich, wenn* $(x,\cos\alpha)\in H'_{21}$ *bzw.* H_{23} *für* Γ *gilt, dann nimmt* Q_{21} *bzw.* Q_{23} *während der Anwendung von* g_3 *bzw.* g_4 *ab.*

BEWEIS. In den Fällen 1, 2, 3 kann man die Richtigkeit der Behauptung ebenso beweisen, wie in den entsprechenden Hilfssätzen von [7] und [8], deshalb gehen wir auf diese Fälle nicht des Näheren ein.

Es sei $(x,\cos\alpha)\in H'_{21}$. $|2A+B|<|3A-2B|$ ist, weil diese Ungleichung mit $\cos\alpha<\frac{5x^2+3}{16x}$ im Fall $x<1$ gilt. Aus $|2A+B|<|3A-2B|$ ergibt sich unmittelbar, daß r_{21} während der Anwendung von g_3 abnimmt. Der Inhalt von D_{21} ist $\frac{7}{2}T(\Gamma)$. Das ändert sich nicht, d.h., $T(\Gamma)$ ist konstant. Deshalb nimmt Q_{21} offenbar ab.

Wir betrachten den Fall $(x,\cos\alpha)\in H_{23}$. $|A+2B|<|3A-B|$ gilt dann und nur dann, wenn $\cos\alpha<\frac{8x^2-3}{10x}$ ist. Wegen $x>\frac{-1,8+\sqrt{19,74}}{3}$ ist $0,3<$ $<\frac{8x^2-3}{10x}$. In diesem Fall ist $\cos\alpha\leqq 0,3$, d.h., $|A+2B|<|3A-B|$ gilt offenbar. Daraus folgt schon, daß r_{23} während der Anwendung von g_4 abnimmt. Der Inhalt von D_{23} ist auch $\frac{7}{2}T(\Gamma)$. Das bleibt während der Anwendung von g_4 konstant, d.h., auch $T(\Gamma)$ ist konstant, deshalb nimmt Q_{23} offenbar ab.

BEMERKUNG 1. Während der Anwendung von g nimmt x ab und nimmt $\cos\alpha$ zu, bei g_1 ist x konstant und nimmt $\cos\alpha$ ab. Bei g_3 nimmt $\cos\alpha$ zu und wenn auch $\cos\alpha\leqq\frac{x}{3}$ gilt, dann nimmt x sicher ab. Während der Anwendung von g_4 nimmt $\cos\alpha$ zu und nimmt x ab.

BEMERKUNG 2. Betrachten wir die Gitter $\overline{\Gamma}_8$, für deren Basisvektoren A und B die Bedingungen $|A|=|B|$ und $\cos\alpha=\frac{|A|}{4|B|}$ gelten. Der Punkt $Z(1,\frac{1}{4})$ entspricht diesen Gittern auf der Abb. 3. Für diese Gitter gelten auch $r_8=r_{11}=r_{20}=r_{22}$. Aus dem Hilfssatz 1 folgt, daß die Kreisanordnung $L(\overline{\Gamma}_8,r_i)$ $(i=8,11,20,22)$ eine 8-fache Kreisüberdeckung ist.

BEWEIS des Satzes. Betrachten wir eine beliebige 8-fache gitterförmige Kreisüberdeckung $L(\Gamma,R)$ und nehmen wir an, daß Γ in normaler Darstellung ist. Die Dichte von $L(\Gamma, R)$ ist $\frac{R^2}{T(\Gamma)}\pi$, wo $T(\Gamma)$ der Inhalt des Grundparallelogramms von Γ ist. Die Methode des Beweises ist dieselbe wie in [7], [8] und [9]. Wenn $(x, \cos\alpha) \in H_i$ $(i = 1, \ldots, 23)$ für Γ gilt, dann ist $R \geqq r_i$. Es geben nämlich höchstens 7 Gitterpunkte im Inneren von k_i (Hilfssatz 2), deshalb wäre z.B. der Mittelpunkt von k_i im Fall $r_i > R$ höchstens 7-fach überdeckt. Es gilt also $\frac{R^2}{T(\Gamma)}\pi \geqq \frac{r_i^2}{T(\Gamma)}\pi = Q_i\pi$. Auf Grund der Formel (3) kann man Q_i mit den Variablen x und α ausdrücken.

Mit Hilfe der dem Bereich H_i entsprechenden Transformation (vgl. Hilfssatz 3) können wir Q_i weiter vermindern und inzwischen erreichen wir einen Grenzpunkt von H_i (Bemerkung 1). Wenn dieser Punkt ein Grenzpunkt irgendeines Bereiches H_j ist und $Q_i = Q_j$ (die gestrichelten Linien auf der Abb. 4) in diesem Grenzpunkt gilt, dann können wir vielleicht Q_j mit einer Transformation weiter vermindern, d.h., wir können den ursprünglichen Quotient Q_i von unten schätzen. Z.B. im Fall $(x, \cos\alpha) \in H_9$ können wir die Transformation g anwenden und endlich erreichen wir den Fall $\cos\alpha = \frac{x}{2}$ (auf der Abbildung 4 DE) oder einen Punkt der Kurve $\cos\alpha = \frac{1-8x^2}{8x}$, $x \in \left[\sqrt{\frac{1}{12}}, \sqrt{\frac{1}{10}}\right]$. Im letzteren Fall gilt $Q_9 = Q_{15}$ und wir können Q_{15} mit g_1 vermindern.

Endlich bekommen wir einen von den folgenden Fällen (die dicken Linien auf der Abb. 4). Hier sind die Quotienten Q_i Funktionen mit einer einzigen Variablen.

1. Im Fall $x \in (0, \frac{1}{4}]$, $\cos\alpha = 0$ (auf der Abb. 4 OA) müssen wir das Minimum der Funktion

$$Q_1(x) = \frac{(16x^2+1)^2}{4x}, \qquad x \in \left(0, \frac{1}{4}\right]$$

bestimmen. Z.B. mit Hilfe der ersten Ableitung bekommen wir, daß die Minimumstelle $x = \sqrt{\frac{1}{48}}$ ist, d.h.,

$$Q_1(x) \geqq Q_1\left(\sqrt{\frac{1}{48}}\right) = \frac{16\sqrt{3}}{9} > 3.$$

2. Im Fall $x \in [\frac{1}{4}, \sqrt{\frac{1}{12}}]$, $\cos\alpha = 0$ (AC) handelt es sich um die Funktion

$$Q_{15}(x) = \frac{(16x^2+1)^2}{64x^3}, \qquad x \in \left[\frac{1}{4}, \sqrt{\frac{1}{12}}\right]$$

und einfach ergibt sich

$$Q_{15}(x) \geqq Q_{15}\left(\sqrt{\frac{1}{12}}\right) = \frac{49\sqrt{3}}{24} > 3.$$

3. Unter den Bedingungen $x \in \left[\frac{1}{4}, \sqrt{\frac{1}{12}}\right]$, $\cos\alpha = \frac{1-12x^2}{8x}(BC)$ muß man das Minimum von Q_2 bestimmen. Mit Hilfe von (3) ergibt sich

$$Q_2 = \frac{(16x^2 + 1 + 8x\cos\alpha)(9x^2 + 1 - 6x\cos\alpha)}{4x\sin^3\alpha}.$$

Mit der Substituierung $\cos\alpha = \frac{1-12x^2}{8x}$ bekommen wir, daß

$$Q_2(x) = 64\frac{144x^6 + 74x^4 + x^2}{\sqrt{-144x^4 + 88x^2 - 1}^3}$$

ist. Es ist leicht einzusehen, daß $Q_2(x)$ zunimmt, d.h.,

$$Q_2(x) \geqq Q_2\left(\frac{1}{4}\right) = \frac{24,75}{\sqrt{3,9375}^3} > 3$$

gilt.

4. Im Fall $x \in \left[\sqrt{\frac{1}{12}}, \sqrt{\frac{2}{17}}\right]$, $\cos\alpha = \frac{x}{2}(DE)$ ergibt sich die Funktion

$$Q_9(x) = \frac{8}{x\sqrt{4 - x^2}^3},$$

die abnimmt, d.h.,

$$Q_9(x) \geqq Q_9\left(\sqrt{\frac{2}{17}}\right) = \frac{578}{33\sqrt{33}} > 3$$

gilt.

5. Unter den Bedingungen $x \in \left[\sqrt{\frac{1}{12}}, \sqrt{\frac{1}{5}}\right]$, $\cos\alpha = O(CH)$ muß man das Minimum von Q_{16} finden. Dann ist

$$Q_{16}(x) = \frac{(16x^2 + 1)(9x^2 + 1)(x^2 + 4)}{196x^3}$$

und z.B. mit Hilfe der ersten Ableitung ist es einzusehen, daß $Q_{16}(x)$ abnimmt, deshalb gilt

$$Q_{16}(x) \geqq Q_{16}\left(\sqrt{\frac{1}{5}}\right) = \frac{63\sqrt{5}}{50} > 2,8.$$

6. Es ist leicht einzusehen, daß $Q_4 \gtrless Q_7$ dann und nur dann gilt, wenn $\cos\alpha \gtrless \frac{3-15x^2}{4x}$ gilt.

6.1. Deshalb ist $Q_4 > Q_7$ im Fall $x \in \left[\sqrt{\frac{3}{16}}, \sqrt{\frac{1}{5}}\right]$, $\cos\alpha = \frac{x}{4}$ (JI), d.h., wir können Q_4 mit dem Minimum von Q_7 von unten schätzen.

6.2. Im Fall $x \in \left[\sqrt{\frac{2}{13}}, \sqrt{\frac{1}{5}}\right]$, $\cos\alpha = \frac{x}{4}$ (GI) kann man Q_7 folgenderweise aufschreiben:

$$Q_7(x) = 16\frac{(x^2+2)^2}{x\sqrt{16-x^2}^3}.$$

Es ist leicht einzusehen, daß $Q_7(x)$ abnimmt, so ist

$$Q_7(x) \geqq Q_7\left(\sqrt{\frac{1}{5}}\right) = \frac{1936}{79\sqrt{79}} > \frac{32}{3\sqrt{15}}.$$

6.3. Wenn $\cos\alpha = \frac{3-15x^2}{4x}$, $x \in \left[\sqrt{\frac{3}{17}}, \sqrt{\frac{3}{16}}\right]$ (JL) gilt, ergibt sich die Funktion

$$56\frac{-34x^6+19x^4-x^2}{\sqrt{-225x^4+106x^2-9}^3}. \tag{4}$$

Man kann einsehen, daß (4) abnehmend ist. Der Fall $\left(\sqrt{\frac{3}{16}}, \frac{1}{4}\sqrt{\frac{3}{16}}\right)$ gehört aber zum Fall 6.2, d.h.,

$$(4) > \frac{32}{3\sqrt{15}}$$

gilt.

7. Im Fall $x \in \left[\sqrt{\frac{1}{5}}, \sqrt{\frac{1}{2}}\right]$, $\cos\alpha = 0$ (HM) ist Q_{17} die folgende Funktion

$$Q_{17}(x) = \frac{(16x^2+1)(4x^2+1)(x^2+1)}{36x^3}, \qquad x \in \left[\sqrt{\frac{1}{5}}, \sqrt{\frac{1}{2}}\right].$$

Mit Hilfe der ersten Ableitung ergibt sich, daß $x = \frac{1}{2}$ die Minimumstelle von $Q_{17}(x)$ ist, deshalb gilt

$$Q_{17}(x) \geqq Q_{17}\left(\frac{1}{2}\right) = \frac{25}{9} > \frac{32}{3\sqrt{15}}.$$

8.1. Das Minimum von Q_5 braucht im Fall $x \in \left[\sqrt{\frac{2}{9}}, \sqrt{\frac{6}{13}}\right]$, $\cos\alpha = \frac{x}{2}$ (KU). Wir bekommen die Funktion

$$\frac{(3x^2+4)^2}{2x\sqrt{4-x^2}^3}, \tag{5}$$

für deren Minimumstelle sich $x = \sqrt{\frac{4}{13}}$ ergibt. So gilt

$$(5) \geqq \frac{16}{3\sqrt{3}} > 3.$$

8.2. Im Fall $x \in \left[\sqrt{\frac{6}{13}}, \sqrt{\frac{1}{2}}\right]$, $\cos\alpha = \frac{6-11x^2}{4x}$ (UT) ist $Q_5 = Q_{21}$. Die erste Ableitung von

$$Q_5(x) = Q_{21}(x) = 560\frac{-3x^6+4x^4-x^2}{\sqrt{-121x^4+148x^2-36}^3}, \quad x \in \left[\sqrt{\frac{6}{13}}, \sqrt{\frac{1}{2}}\right]$$

ist mit der Funktion

$$F(x) = -91x^6 + 189x^4 - 107x^2 + 18, \qquad x \in \left[\sqrt{\frac{6}{13}}, \sqrt{\frac{1}{2}}\right]$$

vom Gesichtspunkt des Vorzeichens gleich. Es ist leicht einzusehen, daß $F(x)$ zunimmt und $F(0,68) < 0$ und $F(0,69) > 0$ gelten, $\overline{x}$ sei die Nullstelle von $F(x)$ und es seien $H_1(x) = -3x^6 + 4x^4 - x^2$ und $H_2(x) = \sqrt{-121x^4+148x^2-36}^3$. $H_1(x)$ und $H_2(x)$ nehmen zu, deshalb gilt

$$Q_5(x) = Q_{21}(x) \geqq Q_5(\overline{x}) > 560\frac{H_1(0,68)}{H_2(0,69)} > 2,88.$$

9.1. Unter den Bedingungen $x \in \left[\sqrt{\frac{6}{13}}, \sqrt{\frac{3}{5}}\right]$ und $\cos\alpha = \frac{3-2x^2}{9x}$ (UV) muß man das Minimum von Q_{21} finden. In diesem Fall gilt $Q_{21} = Q_{14}$ und es handelt sich um die Funktion

$$Q_{21}(x) = Q_{14}(x) = \frac{315}{4}\frac{4x^6+15x^4+9x^2}{\sqrt{-4x^4+93x^2-9}^3}, \quad x \in \left[\sqrt{\frac{6}{13}}, \sqrt{\frac{3}{5}}\right].$$

Mit Hilfe der ersten Ableitung können wir zeigen, daß $Q_{14}(x)$ abnimmt. Folglich gilt

$$Q_{21}(x) = Q_{14}(x) \geqq Q_{21}\left(\sqrt{\frac{3}{5}}\right) = \frac{45}{4\sqrt{14}} > 3.$$

9.2. Im Fall $x = \sqrt{\frac{3}{5}}$, $\cos\alpha \in \left[\frac{1}{24}\sqrt{\frac{3}{5}}, \frac{1}{3}\sqrt{\frac{3}{5}}\right]$ (VR) ist Q_{21} die folgende Funktion

$$Q_{21}(\alpha) = \frac{720\sqrt{15}\cos^2\alpha - 5520\cos^2\alpha - 927\sqrt{15}\cos\alpha + 6392}{490\sqrt{15}\sin^3\alpha}.$$

Es ist einzusehen, daß $Q_{21}(\alpha)$ zunimmt. Deshalb ist

$$Q_{21}(\alpha) \geqq Q_{21}\left(\arccos\frac{1}{3}\sqrt{\frac{3}{5}}\right)$$

und wir haben im Fall 9.1 gesehen, daß $Q_{21}\Big(\arccos\frac{1}{3}$

10. Wenn $x = \sqrt{\frac{3}{5}}$ und $\cos\alpha \in \left[\frac{1}{6}\sqrt{\frac{3}{5}}, \frac{1}{3}\sqrt{\frac{3}{5}}\right]$ (VS) sind, dann ist

$$Q_{13}(\alpha) = \frac{3}{2\sqrt{15}} \frac{8 - \sqrt{15}\cos\alpha}{\sin^3\alpha},$$

für deren Minimumstelle $\cos\alpha_m = \frac{4\sqrt{15}-\sqrt{190}}{10}$ ist. Es ist auszurechnen, daß

$$Q_{13}(\alpha_m) > 2,8$$

gilt.

11.1. Im Fall $x \in \left[\sqrt{\frac{1}{2}}, \sqrt{\frac{3}{5}}\right]$, $\cos\alpha = 0$ (MN) ist

$$Q_{22}(x) = \frac{4x^4 + 5x^2 + 1}{4x^3},$$

die abnimmt, deshalb gilt

$$Q_{22}(x) \geqq Q_{22}\left(\sqrt{\frac{3}{5}}\right) = \frac{34}{3\sqrt{15}} > \frac{32}{3\sqrt{15}}.$$

11.2. Unter den Bedingungen $x \in \left[\sqrt{\frac{11}{13}}, 1\right]$, $\cos\alpha = \frac{1+x^2}{8x}$ (YZ) können wir Q_{22} folgenderweise aufschreiben:

$$80\frac{7x^4 + 8x^2 + 1}{\sqrt{-x^4 + 62x^2 - 1}^3}. \tag{6}$$

Z.B., mit Hilfe der ersten Ableitung können wir einsehen, daß (6) abnimmt. (6) ist also im Punkt $(1, \frac{1}{4})$, d.h. bei einer Kreisüberdeckung $L(\overline{\Gamma}_8, r_{22})$ (vgl. Bemerkung 2) minimal.

12. $x \in \left[\sqrt{\frac{3}{5}}, 1\right]$ und $\cos\alpha = 0$ (NP). Dann ist

$$Q_{12}(x) = \frac{(x^2 + 9)^2}{36x}.$$

Es ist leicht zu zeigen, daß $Q_{12}(x)$ abnehmend ist. Es gilt

$$Q_{12}(x) \geqq Q_{12}(1) = \frac{25}{9} > \frac{32}{3\sqrt{15}}.$$

13.1. Im Fall $x \in \left[\sqrt{\frac{6}{13}}, \sqrt{\frac{5}{9}}\right]$, $\cos\alpha = \frac{x}{2}$ (UW) ist

$$Q_{14}(x) = \frac{2}{9}\frac{(9 - 2x^2)^2}{x\sqrt{4 - x^2}^3}.$$

$Q_{14}(x)$ ist abnehmend, weil $x\sqrt{4-x^2}^3$ zunimmt, d.h.,

$$Q_{14}(x) \geqq Q_{14}\left(\sqrt{\frac{5}{9}}\right) > 2,9$$

gilt.

13.2. Unter den Bedingungen $x \in \left[\sqrt{\frac{5}{9}}, \sqrt{\frac{11}{17}}\right]$ und $\cos\alpha = \frac{5-3x^2}{12x}$ $(W\overline{A})$ ist Q_{14} die folgende Funktion

$$(7) \qquad 24\frac{35x^6+111x^4+52x^2}{\sqrt{-9x^4+174x^2-25}^3}, \qquad x \in \left[\sqrt{\frac{5}{9}}, \sqrt{\frac{11}{17}}\right].$$

Es ist leicht einzusehen, daß (7) abnimmt. Deshalb ist $x = \sqrt{\frac{11}{17}}$ die Minimumstelle. Hier ist aber der Funktionwert größer als 2,8.

14. Betrachten wir den Quotient Q_{18} im Fall $x \in \left[\sqrt{\frac{5}{9}}, \sqrt{\frac{7}{9}}\right]$, $\cos\alpha = \frac{5-3x^2}{12x}$ (WX):

$$Q_{18}(x) = 36\frac{105x^4+34x^2-7}{\sqrt{-9x^4+174x^2-25}^3}.$$

Mit der in 8.2 verwandten Methode ist es zu zeigen, daß

$$Q_{18}(x) > 2,7578 > \frac{32}{3\sqrt{15}}$$

gilt.

15.1. Im Fall $x \in \left[\sqrt{\frac{11}{17}}, \sqrt{\frac{7}{9}}\right]$, $\cos\alpha = \frac{5-3x^2}{12x}$ $(\overline{A}X)$ ist Q_{20} die folgende Funktion:

$$Q_{20}(x) = \frac{72}{25}\frac{252x^6+1047x^4+439x^2-44}{\sqrt{-9x^4+174x^2-25}^3}.$$

Es ist leicht einzusehen, daß $Q_{20}(x) > Q_{18}(x)$ im Fall $x \in \left[\sqrt{\frac{11}{17}}, \sqrt{\frac{7}{9}}\right]$ gilt, deshalb ist auch

$$Q_{20}(x) > \frac{32}{3\sqrt{15}}.$$

15.2. Wenn $x \in \left[\sqrt{\frac{7}{9}}, 1\right]$ und $\cos\alpha = \frac{1+x^2}{8x}$ (XZ) sind, dann kann man Q_{20} folgenderweise aufschreiben:

$$(8) \qquad 40\frac{3x^6+13x^4+13x^2+3}{\sqrt{-x^4+62x^2-1}^3}.$$

Es ist leicht einzusehen, daß (8) abnimmt. Deshalb tritt das Minimum im Punkt $(1,\frac{1}{4})$, d.h., bei der 8-fachen Kreisüberdeckung $L(\Gamma_8, r_{20})$ (vgl. Bemerkung 2) auf.

16. Im Fall $x \in \left[\sqrt{\frac{7}{9}}, \sqrt{\frac{11}{13}}\right]$, $\cos\alpha = \frac{1+x^2}{8x}$ (XY) müssen wir das Minimum der Funktion

$$Q_{19}(x) = \frac{15}{2}\frac{39x^6 + 85x^4 + 53x^2 + 7}{\sqrt{-x^4 + 62x^2 - 1}^3}$$

bestimmen. Es ist einzusehen, daß die erste Ableitung von $Q_{19}(x)$ positiv ist, d.h.

$$Q_{19}(x) \geqq Q_{19}\left(\sqrt{\frac{7}{9}}\right) > 2,78$$

gilt.

17.1. Im Fall $x \in \left[\sqrt{\frac{11}{18}}, \frac{-1,8+\sqrt{19,74}}{3}\right]$, $\cos\alpha = \frac{11-6x^2}{24x}$ $(\bar{E}\bar{D})$ ist Q_8 die folgende Funktion:

$$48\frac{72x^6 + 354x^4 + 299x^2}{\sqrt{-36x^4 + 708x^2 - 121}^3}. \tag{9}$$

Es ist einzusehen, daß (9) abnimmt und der Minimumwert $> 2,78$ ist.

17.2. Unter den Bedingungen $x \in \left[\frac{-1,8+\sqrt{19,74}}{3}, x_k\right]$ (s. x_k bei H''_{11}) und $\cos\alpha = 0,3$ $(\bar{B}\bar{C})$ ergibt sich aus Q_8 die Funktion

$$\frac{x^4 - 0,6x^3 + 4,28x^2 + 1,2x + 4}{4x\sqrt{1 - 0,3^2}^3}. \tag{10}$$

Mit Hilfe der ersten Ableitung kann man einsehen, daß (10) zunimmt. Deshalb stimmt das Minimum von (10) mit dem Minimum von (9) überein, d.h., $(10) > 2,78$ gilt.

17.3. Im Fall $x \in [x_k, 1]$ und $Q_8 = Q_{11}$ $\left(\cos\alpha = \frac{8-x^2-\sqrt{9x^4-8x^2+24}}{8x}\right)$ $(\bar{B}Z)$ muß man das Minimum von Q_8 geben. Es ist

$$Q_8(x) = Q_{11}(x) = 64\frac{11x^4 - 6x^2 + 3x^2\sqrt{9x^4 - 8x^2 + 24}}{\sqrt{-10x^4 + 88x^2 - 88 + 2(8 - x^2)\sqrt{9x^4 - 8x^2 + 24}}^3},$$

deren erste Ableitung vom Gesichtspunkt des Vorzeichens mit der Funktion

$$G(x) = 234x^8 + 680x^6 - 2392x^4 + 4248x^2 - 4320 + \\ + (82x^6 + 272x^4 - 944x^2 + 840)\sqrt{9x^4 - 8x^2 + 24}$$

gleich ist. Im Fall $x \geqq x_k$ nimmt $9x^4 - 8x^2 + 24$ zu, d.h., $9x^4 - 8x^2 + 24 \leqq 25$ gilt und ist $82x^6 + 272x^4 - 944x^2 + 840 > 0$. Deshalb gilt

$$\begin{aligned} G(x) &\leqq 234x^8+680x^6-2392x^4+4248x^2-4320+(82x^6+272x^4-944x^2+840)5= \\ &= 234x^8 + 1090x^6 - 1032x^4 - 472x^2 - 120 = \\ &= 234x^2(x^6 - 1) + (1090x^4 - 1032x^2 - 238)x^2 - 120 < 0. \end{aligned}$$

Das bedeutet aber, daß $Q_8(x)$ abnimmt und ihr Minimum im Punkt $(1, \frac{1}{4})$, d.h. bei der Kreisüberdeckung $L(\overline{\Gamma}_8, r_8)$ auftritt.

18. Im Fall $x \in \left[\frac{0,6+\sqrt{24,36}}{6}, x_k\right]$ und $\cos\alpha = 0,3$ (BC) ist Q_{11} die folgende Funktion:

$$\frac{1}{4\sqrt{1-0,3^2}^3} \frac{4x^2 - 3,6x + 9}{x}. \tag{11}$$

Es ist leicht zu zeigen, daß (11) abnimmt. Ihre Minimumstelle ist x_k. Das gehört aber zum Fall 17.3.

19. Endlich müssen wir das Minimum von Q_{23} unter den Bedingungen $x \in \left[\frac{-1,8+\sqrt{19,74}}{3}, \frac{0,6+\sqrt{24,36}}{6}\right]$ und $\cos\alpha = 0,3$ (CD) geben. Dann ist

$$Q_{23}(x) = \frac{36x^6 + 3,6x^5 + 187,96x^4 - 68,424x^3 + 351,16x^2 - 68,4x + 36}{196\sqrt{1-0,3^2}^3 x^3}.$$

Es ist einzusehen, daß $Q_{23}(x)$ abnimmt and deshalb gilt

$$Q_{23}(x) \geqq Q_{23}\left(\frac{0,6+\sqrt{24,36}}{6}\right) > 2,8.$$

Damit haben wir den Beweis des Satzes beendet.

LITERATURVERZEICHNIS

[1] BLUNDON, W. J., Multiple covering of the plane by circles, *Mathematika* **4** (1957), 7–16. *MR* **19**-877

[2] FEJES TÓTH, L., *Lagerungen in der Ebene, auf der Kugel und im Raum*, 2. verbesserte und erweiterte Auflage, Die Grundlehren der mathematischen Wissenschaften, Band 65, Springer-Verlag, Berlin, 1972. *MR* **50** #5603

[3] HAAS, A., Die dünnste siebenfache gitterförmige Überdeckung der Ebene durch kongruente Kreise, Dissertation, Wien, 1976.

[4] KERSHNER, R., The number of circles covering a set, *Amer. J. Math.* **61** (1939), 665–671. *Zbl* **21**, 114

[5] LINHART, J., Eine Methode zur Berechnung der Dichte einer dichtesten gitterförmigen k-fachen Kreispackung, Arbeitsbericht, Math. Inst. Univ. Salzburg.

[6] SUBAK, H., Mehrfache gitterförmige Überdeckungen der Ebene durch Kreise, Dissertation, Wien, 1960.

[7] TEMESVÁRI, Á. H., Die dünnste gitterförmige 5-fache Kreisüberdeckung der Ebene, *Studia Sci. Math. Hungar.* **19** (1984), 285–298. *MR* **88k**:52024

[8] TEMESVÁRI, Á. H., Die dünnste 6-fache gitterförmige Kreisüberdeckung der Ebene, *Berzsenyi Dániel Tanárképző Főiskola Tud. Közl.* **9** (1992), 93–112.

[9] TEMESVÁRI, Á. H., Die dünnste 7-fache gitterförmige Kreisüberdeckung der Ebene, *Berzsenyi Dániel Tanárképző Főiskola Tud. Közl.* **9** (1992), 113–126.

(Eingegangen am 7. September 1983)

ERDÉSZETI ÉS FAIPARI EGYETEM
ERDŐMÉRNÖKI KAR
MATEMATIKA ÉS ÁBRÁZOLÓ GEOMETRIA TANSZÉK
POSTAFIÓK 132
H–9401 SOPRON
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 341–347

ON THE RECONSTRUCTION OF COMBINATORIAL STRUCTURES FROM LINE-GRAPHS

P. L. ERDŐS

Abstract

We present some generalisations of reconstruction results of H. Whitney, C. Berge, J. C. Fournier and L. Lovász. Certain classes of hypergraphs are described which are determined (up to isomorphism) by their line-graphs. There are Hamming schemes, t-designs and finite vector spaces among the classes described.

1. Introduction

Let $G=(E_i : i\in M)$ and $G'=(F_i : i\in M)$ be two connected simple graphs with $|M|>2$ edges. H. Whitney ([Wh]) had

THEOREM 1 (H. Whitney). *Whenever the minimum valencies in the above graphs are at least* 4 *then*

$$|E_i\cap E_j|=|F_i\cap F_j| \qquad (i,j\in M)$$

implies that $G\cong G'$. □

In other words: the line-graphs with the degree condition are isomorphic if and only if the graphs themselves are isomorphic.

This paper presents some new analogous results for hypergraphs. (For the notions of hypergraph theory not defined here we follow Berge [B2].)

DEFINITION. The hypergraphs $\mathcal{H}=(E_i : i\in M)$ and $\mathcal{H}'=(F_i : i\in M)$ are *isomorphic* if there exists a bijection α between their vertex sets and a permutation π of M such that $\alpha(E_i)=F_{\pi(i)}$ for each $i\in M$. Denote the set of the *automorphisms* of the hypergraph $\mathcal{H}$ by $\mathrm{Aut}(\mathcal{H})$.

For every $I\subseteqq M$ we put

$$E_I=\bigcup_{i\in I}E_i.$$

1991 *Mathematics Subject Classification.* Primary 05D05; Secondary 05C60.

Key words and phrases. Erdős–Ko–Rado theorem, hypergraph reconstruction.

The author was partially supported by PHARE ACCORD program under contract H9112–0711.

0081–6906/94/$ 4.00

The *line-graph* $\mathcal{L}(H)$ of the hypergraph $\mathcal{H}$ is defined as follows: the underlying set of $\mathcal{L}(H)$ is the set of edges of $\mathcal{H}$ and the pair (E_i, E_j) is an edge of $\mathcal{L}(H)$ $(E_i \neq E_j,\ i,j \in M)$ iff $E_i \cap E_j \neq \emptyset$. Every automorphism $\alpha \in \mathrm{Aut}(\mathcal{H})$ *induces* an automorphism a_α of $\mathcal{L}(H)$ in a natural way, namely

$$a_\alpha(E_i) = \{\alpha(x) : x \in E_i\} \qquad (i \in M).$$

Finally let K_n^r denote the r-uniform complete hypergraph on n vertices.

The following result was proved by Lovász ([Lo]):

THEOREM 2. *When* $a \in \mathrm{Aut}(\mathcal{L}(K_n^r))$ *where* $n > 2r$ *then there exists an automorphism* $\alpha \in \mathrm{Aut}(K_n^r)$ *which induces the automorphism* a, *i.e.* $a = a_\alpha$.

Lovász strengthened an earlier theorem of C. Berge:

COROLLARY 3 ([B1]). *Let* $\mathcal{H} = (E_i : i \in M)$ *be isomorphic to* K_n^r *where* $n > 2r$. *Suppose that*

$$|E_i \cap E_j| = |F_i \cap F_j| \qquad (i,\ j \in M).$$

Then $\mathcal{H}'$ *is isomorphic to* $\mathcal{H}$. □

(In fact, Berge proved a bit more, namely that this isomorphism is *strong*, but from the Lovász theorem this can be derived easily as well.) In the same year J.C. Fournier proved the analogous result under the condition $n < 2r$ ([F1]). Two years later Fournier found a common proof for both results ([F2]).

2. A general reconstruction principle

First of all we give a very simple proof for Theorem 2. The original proof applied counting to prove the existence of the wanted permutation α. The following proof was developed by Z. Füredi and the author ([EF]) to construct the prescribed permutation. A very similar proof is contained in the paper of Poljak and Rödl ([PR]), also.

This proof is based on a well-known theorem of Paul Erdős, Chao Ko and Richard Rado ([EKR]). Let a hypergraph be called *intersecting* if the pairwise intersections of the edges are not empty.

ERDŐS–KO–RADO THEOREM. *If* $\mathcal{H} = (E_i : i \in M)$ *is an* r-*uniform intersecting hypergraph, where* $|E_M| > 2r$, *then*

$$|M| \leq \binom{|E_M| - 1}{r - 1}$$

and the equality holds iff there is a point $x \in E_M$ *for which* $\cap\mathcal{H} = \{x\}$. *In other words:*

$$\mathcal{H} = \mathcal{H}_x = \{E \subset E_M : |E| = r, x \in E\}. \qquad \square$$

PROOF OF THEOREM 2 ([EF]). Let $\mathcal{C}_x$ denote the system of edges of $\mathcal{H}$ containing the vertex x ($x \in X = \cup K_n^r$), i. e. let

$$\mathcal{C}_x = \{E \in K_n^r : x \in E\}.$$

Then $\mathcal{C}_x$ is an intersecting set system, and $|\mathcal{C}_x| = \binom{n-1}{r-1}$. Since a is an automorphism of the line-graph $\mathcal{L}(H)$, therefore $\{a(E) : E \in \mathcal{C}_x\} = \mathcal{C}'_x$ is intersecting, furthermore, due to the Erdős–Ko–Rado Theorem, there exists a point $\alpha(x) \in X$ for which $\mathcal{C}'_x = a(\mathcal{C}_x) = \mathcal{H}_{\alpha(x)}$. One can prove easily that the map $\alpha : X \longrightarrow X$ is a bijection and $a = a_\alpha$. □

The previous proof is just a special case of a rather general reconstruction principle. This is the following:

EKR RECONSTRUCTION PRINCIPLE. Let S be a finite underlying set and $\mathcal{F}$ be the system of all subsets of S of a certain kind. Let the notion of *pairwise intersection* be defined in $\mathcal{F}$ (not necessarily by intersection in S). We say, that the pair $(S, \mathcal{F})$ satisfies the *EKR-property* with the function $f(S, \mathcal{F})$ if every subfamily $\mathcal{G}$ of $\mathcal{F}$ with pairwise intersecting elements and with cardinality $|\mathcal{G}| \geqq f(S, \mathcal{F})$ satisfies the condition $|\bigcap \mathcal{G}| = 1$.

METATHEOREM. *Let $\mathcal{H}$ be a subhypergraph of $\mathcal{F}$ satisfying the following* valency condition *in every vertex $x \in S$:*

$$v_{\mathcal{H}}(x) := |\{E \in \mathcal{H} : x \in E\}| \geqq f(S, \mathcal{F}).$$

Let $a \in \mathrm{Aut}(L(\mathcal{H}))$. Then there exists an automorphism $\alpha \in \mathrm{Aut}(\mathcal{H})$ such that $a = a_\alpha$.

METAPROOF. We just have to repeat the previous proof. Let C_x denote the "star" of the vertex x in the hypergraph $\mathcal{H}$. The image $a(C_x)$ is a pairwise intersecting subhypergraph of $\mathcal{F}$. Since $\mathcal{F}$ satisfies the EKR-property due to the valency condition of $\mathcal{H}$, therefore the vertex $\alpha(x) := \bigcap a(C_x)$ is well defined, and one can easily see that the map $\alpha : X \longrightarrow X$ is a bijection and $a = a_\alpha$. □

Hereafter we examine several structures which satisfy the EKR-property and prove analogues reconstruction results for them. For any structure at first we list the known EKR-type result, and then determine the reconstruction result. The proofs will contain the required "extra" facts, only.

In the remaining part of this section we apply the EKR reconstruction principle to improve Fournier's theorem. (This improvement for Berge's theorem was done in the paper [EF].) At first we remark that, due to Hilton and Milner ([HM]), the family $\mathcal{F}$ of all r-element subsets of the n-element set X satisfies the EKR-property with the function

$$v(n, r) = \binom{n-1}{r-1} - \binom{n-r-1}{r-1} + 2.$$

THEOREM 4. *Let $\mathcal{H}=(E_i : i \in M)$ be an r-uniform hypergraph with $|E_M|=2r-k$ points ($0<k<r$). Suppose that for any point $x \in E_M$ the condition*

$$|M| - v_{\mathcal{H}}(x) > \binom{2r-k-1}{r} - \binom{r-1}{k} + 1$$

holds. Let $a \in \mathrm{Aut}(\mathcal{L}(H))$ for which

$$|E_i \cap E_j| = |a(E_i) \cap a(E_j)|$$

whenever $i,j \in M$. Then there exists an $\alpha \in \mathrm{Aut}(\mathcal{H})$ for which $a = a_\alpha$.

PROOF. Let $\bar{\mathcal{H}} = (\bar{E}_i : i \in M)$ where $\bar{E}_i = E_M \setminus E_i$ ($i \in M$). Then $\bar{\mathcal{H}}$ is an $(r-k)$-uniform hypergraph with $|\bar{E}_M| = 2r-k$ points. Let $\bar{a}$ be the permutation of $\{\bar{E}_i\}$ for which

$$\bar{a}(\bar{E}) = E_M \setminus a(E).$$

By the intersection condition of the automorphism a:

$$\begin{aligned}|\bar{E}_i \cap \bar{E}_j| &= |(E_M \setminus E_i) \cap (E_M \setminus E_j)| = |E_M \setminus (E_i \cup E_j)| = \\ &= |E_M| - 2r + |E_i \cap E_j| = |E_M| - 2r + |a(E_i) \cap a(E_j)| = \\ &= |(E_M \setminus a(E_i)) \cap (E_M \setminus a(E_j))| = |\bar{a}(\bar{E}_i) \cap \bar{a}(\bar{E}_j)|\end{aligned}$$

for any $i,j \in M$, that is $\bar{a} \in \mathrm{Aut}(\mathcal{L}(\bar{\mathcal{H}}))$. Furthermore

$$v_{\bar{\mathcal{H}}}(x) \geqq \binom{2r-k-1}{r} - \binom{r-1}{k} + 2$$

holds for each points of $\bar{\mathcal{H}}$, since $(x \in \bar{E} \in \bar{\mathcal{H}}) \Longleftrightarrow (x \notin E \in \mathcal{H})$. So we can apply Metatheorem (with the Hilton–Milner condition) for the hypergraph $\bar{\mathcal{H}}$ and the automorphism $\bar{a} \in \mathrm{Aut}(\mathcal{L}(\bar{\mathcal{H}}))$. Consequently there exists an automorphism $\alpha \in \mathrm{Aut}(\bar{\mathcal{H}})$ for which $\bar{a}(\bar{E}) = \{\alpha(x) : x \in \bar{E}\}$ (when $\bar{E} \in \bar{\mathcal{H}}$). This permutation $\alpha : E_M \longrightarrow E_M$ is also a suitable automorphism of the hypergraph $\mathcal{H}$. We must show, that the condition

$$a(E) = \{\alpha(x) : x \in E\}$$

holds for every edge $E \in \mathcal{H}$. But we know, that for every point $x \notin E$ the relation $\alpha(x) \in \bar{a}(\bar{E}) = E_M \setminus a(E)$ holds, that is $\alpha(x) \notin a(E)$. Hence $\alpha(\bar{E}) = E_M \setminus a(E)$ therefore the equality $\alpha(E) = a(E)$ holds for the bijection α. □

REMARK 5. We know, that any pair of edges of the r-uniform hypergraph $\mathcal{H}$ is automatically intersecting because $|E_M| = 2r-k$. Furthermore the proof used only the condition $\bar{E}_i \cap \bar{E}_j = \emptyset \Longleftrightarrow \bar{a}(\bar{E}_i) \cap \bar{a}(\bar{E}_j) = \emptyset$. Therefore the map a must be a permutation of $\{E_i\}$ which satisfies the condition

$$E_i \cup E_j = E_M \Longleftrightarrow a(E_i) \cup a(E_j) = E_M \qquad (i,j \in M).$$

3. Hamming schemes $H(r,q)$

Let $Q=\{1, 2, \dots, q\}$ and $Q^r=\{x=(x_1, \dots, x_r):x_i\in Q\}$. We define the *distance* $d(x,y)$ $(x,y\in Q^r)$ as follows:

$$d(x,y)=|\{i\in\{1, 2, \dots, r\}:x_i\neq y_i\}|.$$

The structure (Q^r,d) is called *Hamming scheme* $H(r,q)$. Two elements of $H(r,q)$ are called *t-intersecting* if their distance $\leqq r-t$.

For convenience we reformulate these notions.

Let $X_1, X_2, \dots, X_r$ be sets that are pairwise disjoint. Let $|X_i|=q$ $(1\leqq i\leqq r)$. Let $X=\bigcup_i X_i$. Let n denote the cardinality $|X|=rq$. Let $K_{r,q}$ be the following r-class hypergraph:

$$K_{r,q}=\{E\subset X:|E\cap X_i|=1 \text{ if } 0\leqq i\leqq r\}.$$

By now the distance of any two elements of $H(r,q)$ is at most $r-t$ iff the corresponding edges of hypergraph $K_{r,q}$ are t-intersecting. In the case $t=$ $=1$ and $q\geqq 3$ the pair $(X,K_{r,q})$ satisfies EKR-property with the function $f(X,K_{r,q})=q^{r-1}$. (See Livingston [Li], Frankl and Füredi [FF] or Moon [Mo].)

THEOREM 6. *Let $q\geqq 3$. For every automorphism $a\in \mathrm{Aut}(L(K_{r,q}))$ there exists an automorphism $\alpha\in \mathrm{Aut}(K_{r,q})$ which induces the automorphism a, i.e. $a=a_\alpha$.*

PROOF. Just apply the Metatheorem. □

We note that there is a permutation $\pi:\{1,2,\dots,r\}\longrightarrow\{1,2,\dots,r\}$ for which

$$\alpha(X_i)=X_{\pi(i)}$$

if $i=1,\dots,r$. To prove it is enough to realize that the points $x,y\in X$ are in same classes iff there is no edge in $K_{r,q}$ which contains both of them. But $\alpha(x)$ and $\alpha(y)$ also satisfy this condition.

4. t-designs

A $t-(n,r,\lambda)$ *design* is an r-uniform $(B_i:i\in M)$ hypergraph for which $|B_M|=n$ and every t-element subset of B_M is covered by exactly λ *blocks* B_i. To avoid degenerate cases it is assumed, that $0<t\leqq r\leqq n$. It is a well-known fact that every s-element $(s\leqq t)$ subset are contained by exactly b_s blocks. This number depends on the parameters of the design and the number s, only. Namely:

$$b_s=\lambda\binom{n-s}{t-s}\Big/\binom{r-s}{t-s}.$$

The t-designs also satisfy the EKR-property.

RANDS' THEOREM ([Ra]). *There exists a function $f(r,t,s)$ with the following property: Let $\mathcal{B}$ an arbitrary $t-(n,r,\lambda)$ design with $n \geqq f(r,t,s)$ points. Let $\mathcal{H}$ be a system of s-intersecting blocks of $\mathcal{B}$. Then*

$$|\mathcal{H}| \leqq b_s$$

and in the case of equality $\mathcal{H}$ can be described as follows: there exists an $X_0 \subset B_M$, $|X_0| = s$ that

$$\mathcal{H} = \{B \in \mathcal{B} : X_0 \subset B\}. \qquad \square$$

The following estimation for the function f is known.

$$f(r,t,s) \leqq \begin{cases} s + \binom{r}{s}(r-s+1)(r-s) & \text{if } s < t-1 \\ s + (r-s)\binom{r}{s}^2 & \text{if } s = t-1. \end{cases}$$

This theorem generalizes the Erdős–Ko–Rado theorem since all r-element subsets of an n-element underlying set form an $r-(n,r,1)$ design.

The following result is proven by our Metatheorem.

THEOREM 7. *Let $\mathcal{B}$ be an arbitrary $t-(n,r,\lambda)$ design. Let $n \geqq f(r,t,1)$. Furthermore let $a \in \operatorname{Aut}(\mathcal{L}(\mathcal{B}))$. Then there exists an automorphism $\alpha \in \operatorname{Aut}(\mathcal{B})$ such that $a = a_\alpha$.* $\square$

5. Finite vector spaces

Let V denote the n-dimensional vector space over the q-element finite field. Furthermore let V^r denote the set of all r-dimensional subspace of the vector space V. It is a well-known fact that the number of r-dimensional subspaces which contains a prescribed t-dimensional subspace is the Gaussian q-binomial coefficient $\begin{bmatrix} n-t \\ r-t \end{bmatrix}_q$.

Two subspace of V is called intersecting if the dimension of their intersection is at least 1. As Hsieh proved ([Hs] or [FW]) if $n \geqq 2r+2$ or $n \geqq 2r+1$ and $q \geqq 3$ then the family V^r of all r-dimensional subspaces as a hypergraph of all 1-dimensional subspaces satisfies the EKR-property with the function $\begin{bmatrix} n-t \\ r-t \end{bmatrix}_q$

THEOREM 8. *Let $n \geqq 2r+2$ or $n \geqq 2r+1$ and $q \geqq 3$. Let $a \in \operatorname{Aut}(\mathcal{L}(V^r))$. Then there exists an automorphism $\alpha \in \operatorname{Aut}(V^r)$ such that $a = a_\alpha$.*

PROOF. The application of Metatheorem shows that there exists an automorphism $\bar{\alpha} \in \operatorname{Aut}(V^1)$ which induces the automorphism a. But the map

$\bar{\alpha}$ can be extended to a bijection $\alpha : V \longrightarrow V$ easily. Namely for every $X \in \in V^1$ let α_X be any bijection $X \longrightarrow \bar{\alpha}(X)$ which preserves the origin. (This bijection exists because the cardinality of the non-zero elements in any one dimensional subspace is constant.) Finally let $\alpha(y) = \alpha_X(y)$ if $y \in X$. Easy to see, that the map α is an automorphism of the hypergraph V^r and $a = a_\alpha$. □

REMARK. As L. Babai pointed out ([Ba]) the transformation α can be defined to be linear.

REFERENCES

[Ba] BABAI, L. (personal communication).

[B1] BERGE, C., Une condition pour qu'un hypergraphe soit fortement isomorphe à un hypergraphe complet ou multiparti, *C.R. Acad. Sci. Paris Sér. A-B* **274** (1972), A1783–A1786. *MR* **47** #4857

[B2] BERGE, C., *Graphs and hypergraphs*, North-Holland Mathematical Library, Vol. 6, North-Holland Publishing Co., Amsterdam–London; American Elsevier Publishing Co., Inc., New York, 1973. *MR* **50** #9640

[EF] ERDŐS, P. L. and FÜREDI, Z., On automorphisms of line-graphs, *European J. Combin.* **1** (1980), 341–345. *MR* **82c**:05081

[EKR] ERDŐS, P., KO, CHAO and RADO, R., Intersection theorems for systems of finite sets, *Quart. J. Math. Oxford Ser.* (2) **12** (1961), 313–320. *MR* **25**#3839

[F1] FOURNIER, J. C., Sur les isomorphismes d'hypergraphes, *C. R. Acad. Sci. Paris Sér. A-B* **274** (1972), A1612–A1614. *MR* **47** #4860

[F2] FOURNIER, J. C., Une condition pour qu'un hypergraphe, ou son complémentaire, soit fortement isomorphe à un hypergraphe complet, *Hypergraph Seminar* (Proc. First Working Sem., Ohio State Univ., Columbus, Ohio, 1972; dedicated to Arnold Ross), Lecture Notes in Math., Vol. 411, Springer, Berlin, 1974, 95–98. *MR* **51** #7953

[FF] FRANKL, P. and FÜREDI, Z., The Erdős–Ko–Rado theorem for integer sequences, *SIAM J. Algebraic Discrete Methods* **1** (1980), 376–381. *MR* **83d**:05008

[Hs] HSIEH, W.N., Intersection theorems for systems of finite vector spaces, *Discrete Math.* **12** (1975), 1–16. *MR* **52** #2903

[HM] HILTON, A. J. W. and MILNER, E. C., Some intersection theorems for systems of finite sets, *Quart. J. Math. Oxford Ser.* (2) **18** (1967), 369–384. *MR* **36** #2510

[Li] LIVINGSTON, M. L., An ordered version of the Erdős–Ko–Rado theorem, *J. Combin. Theory Ser. A* **26** (1979), 162–165. *MR* **80e**:05010

[Lo] LOVÁSZ, L., *Combinatorial problems and exercises*, North-Holland, Amsterdam–New York, 1979. Problems 15.1–2, pp. 506–510. *MR* **80m**:05001

[Mo] MOON, A., An analogue of the Erdős-Ko-Rado theorem for the Hamming schemes $H(n,q)$, *J. Combin. Theory Ser. A* **32** (1982), 386–390. *MR* **84i**:05004

[PR] POLJAK, S. and RÖDL, V., On set systems determined by intersections, *Discrete Math.* **34** (1981), 173–184. *MR* **82e**:05102

[Ra] RANDS, B.M.I., An extension of the Erdős–Ko–Rado theorem to t-design, *J. Combin. Theory Ser. A* **32** (1982), 391–395. *MR* **84i**:05024

[Wh] WHITNEY, H., Congruent graphs and the connectivity of graphs, *Amer. J. Math.* **54** (1932), 150–168. *Zbl* **3**, 328

(Received March 27, 1991)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 349–362

EXTENDING AND COMPLETING QUIET QUASI-UNIFORMITIES

J. DEÁK

Abstract

Complete extensions of quiet quasi-uniformities (and also of more general ones) are considered adopting a bitopological point of view: the original space is required to be doubly dense in the extension.

Doitchinov [9, 10] introduced the quiet quasi-uniformities, and developed for them a nice theory of completion. We are going to show that his method (which is closely related to some independent results of [5b]) works under more general assumptions, although at the price of losing some good properties of the construction. Quiet complete extensions will turn out to be essentially unique, assuming that double density is required in the definition of an extension.

0. Preliminaries

0.1. Terminology. See [14] Chapter 1 for basic definitions. We shall follow the notations and terminology of [5a, 5b]; the reader familiar with [5a] can skip the next paragraph and continue with 0.2.

$\mathcal{U}^{tp}$ is the topology of the quasi-uniformity $\mathcal{U}$; $(\mathcal{U}^{-tp}, \mathcal{U}^{tp})$ is its *bitopology*, where $\mathcal{U}^{-tp} = \mathcal{U}^{-1tp} = (\mathcal{U}^{-1})^{tp}$. $\mathrm{cl}^{\imath}$ denotes the $\mathcal{U}^{\imath tp}$-closure. We write Ux and $U[A]$ instead of $U(x)$ and $U(A)$. Prescribing a system of *trace filter pairs* in the bitopological space X means that filter pairs $(\mathfrak{f}^{-1}(a), \mathfrak{f}^{1}(a))$ are assigned to each element a of a set $Y \supset X$ such that, for $x \in X$, $(\mathfrak{f}^{-1}(x), \mathfrak{f}^{1}(x))$ is the neighbourhood filter pair of x; there exist bitopological extensions inducing the prescribed trace filter pairs (i.e. $(\mathfrak{f}^{-1}(a), \mathfrak{f}^{1}(a))$ is the trace of the neighbourhood filter pair $(\mathfrak{n}^{-1}(a), \mathfrak{n}^{1}(a))$ of a in the extension) iff the trace filter pairs are open (i.e. $\mathfrak{f}^{\imath}(a)$ is $\mathcal{U}^{\imath tp}$-open); the coarsest one of these extensions is called *doubly strict.* The quasi-uniform space $(Y, \mathcal{V})$ is an *extension*

1991 *Mathematics Subject Classification.* Primary 54E15; Secondary 54E55, 54E05.
Key words and phrases. Quasi-uniformity, bitopology, quiet, complete, extension, quasi-proximity.

Research supported by the Hungarian National Foundation for Scientific Research Grant No. 2114.

of $(X,\mathcal{U})$ if the latter is a subspace of the former, and X is doubly dense ($\mathcal{V}^{-tp}$-dense as well as $\mathcal{V}^{tp}$-dense) in Y. One can consider quasi-uniform extensions inducing some trace filter pairs prescribed in a given quasi-uniform space. The filter pair $(\mathfrak{f}^{-1}, \mathfrak{f}^{1})$ in the quasi-uniform space $(X,\mathcal{U})$ is *round* if for any $S \in \mathfrak{f}^{i}$ there are $U \in \mathcal{U}$ and $T \in \mathfrak{f}^{i}$ such that $U^{i}[T] \subset S$ $(i = \pm 1)$; it is *Cauchy* if for any $U \in \mathcal{U}$ there are $S_i \in \mathfrak{f}^{i}$ with $S_{-1} \times S_1 \subset U$. A *distance* d on X is a real function defined on some subset of $X \times X$ such that if $d(x,y)$ and $d(y,z)$ are both defined then so is $d(x,z)$, and $d(x,y) + d(y,z) \geqq d(x,z)$; the entourages $U_{(\varepsilon)} = U_{(\varepsilon)}(d) = \Delta \cup \{(x,y) : d(x,y) < \varepsilon\}$ $(\varepsilon > 0)$ form a base for a quasi-uniformity denoted by $\mathcal{U}(d)$ (where Δ is the diagonal of $X \times X$). $\mathcal{U}_{so} = \mathcal{U}(d_{so})$ is the *Sorgenfrey quasi-uniformity* on $\mathbf{R}$, where $d_{so}(x,y) = y - x$ if $x \leqq y$. For $x > y$, the notation $]x,y[$ means the interval $]y,x[$.

0.2 A convention. There exist quasi-uniform extensions inducing some prescribed trace filter pairs iff they are round and Cauchy ([5b] Theorem 6.1), so let us agree that *trace filter pairs are always understood to be round and Cauchy* (but this convention does not apply to filter pairs in general).

0.3 Quietness and related notions. A filter pair $(\mathfrak{f}^{-1}, \mathfrak{f}^{1})$ in a quasi-uniform space $(X,\mathcal{U})$ is *weakly concentrated* (for Cauchy filter pairs: equivalent to the original definition given in [5b] 7.6, see [5b] Lemma 7.7) if for any $U \in \mathcal{U}$ there is a $V \in \mathcal{U}$ such that $x_{-1}\, U\, x_1$ whenever $V^{i} x_{-i} \in \mathfrak{f}^{i}$ $(i = \pm 1)$; a family Φ of filter pairs is *uniformly weakly concentrated* (cf. [5b] 7.15) if the above condition holds with the same V for each $(\mathfrak{f}^{-1}, \mathfrak{f}^{1}) \in \Phi$. $\mathcal{U}$ is *quiet* [9, 10] if the family of all the Cauchy filter pairs is uniformly weakly concentrated. Generalizing the terminology of [12], we shall say that the entourage V is *quiet for U* and Φ (or: quiet for U if Φ is clear from the context). A quasi-uniformity is quiet iff the round Cauchy filter pairs are uniformly weakly concentrated. (If V is quiet for U and these filter pairs and $W^2 \subset V$ then W is quiet for U and all the Cauchy filter pairs.)

$(\mathfrak{f}^{-1}, \mathfrak{f}^{1})$ is *concentrated* if it is weakly concentrated and minimal Cauchy (for Cauchy filter pairs: equivalent to the original definition, see [5b] 7.3 and 7.13; minimal is to be understood in the partial order $(\mathfrak{f}^{-1}, \mathfrak{f}^{1}) \leqq (\mathfrak{g}^{-1}, \mathfrak{g}^{1})$ iff $\mathfrak{f}^{i} \subset \mathfrak{g}^{i}$). The neighbourhood filter pairs are concentrated, and they are the only convergent concentrated filter pairs (*convergent* means that there is a point with $\mathfrak{f}^{i}$ $\mathcal{U}^{itp}$-converging to it). Concentrated filter pairs are round. A family of filter pairs is *uniformly concentrated* if it is uniformly weakly concentrated and each filter pair is concentrated (i.e. minimal Cauchy). Given a weakly concentrated Cauchy filter pair $(\mathfrak{f}^{-1}, \mathfrak{f}^{1})$, there exists a unique concentrated filter pair coarser than $(\mathfrak{f}^{-1}, \mathfrak{f}^{1})$; it can be described as the coarsest one among the (weakly concentrated) Cauchy filter pairs coarser than $(\mathfrak{f}^{-1}, \mathfrak{f}^{1})$; if each member of a family of weakly concentrated Cauchy filter pairs is replaced by the concentrated one described above then we obtain a uniformly concentrated family ([5b] Lemma 7.11 and Remark 8.13 b)).

Let $(\mathfrak{m}^{-1}(\mathfrak{f}^{-1}), \mathfrak{m}^{1}(\mathfrak{f}^{1}))$ denote the concentrated filter pair coarser than the given weakly concentrated Cauchy filter pair $(\mathfrak{f}^{-1}, \mathfrak{f}^1)$. According to [5b] 7.11, the sets $M_i(U)$ $(U \in \mathcal{U})$ constitute a base for $\mathfrak{m}^i(\mathfrak{f}^i)$ where

$$M_i(U) = \bigcup\{S_i \in \mathfrak{f}^i : \exists S_{-i} \in \mathfrak{f}^{-i}, S_{-1} \times S_1 \subset U\}.$$

[9] Proposition 2 suggests an alternative construction: $\mathfrak{m}^1(\mathfrak{f}^1)$ is the intersection of all the filters $\mathfrak{g}^1$ for which there exists a filter $\mathfrak{g}^{-1}$ such that $(\mathfrak{g}^{-1}, \mathfrak{g}^1)$ and $(\mathfrak{g}^{-1}, \mathfrak{f}^1)$ are both Cauchy, and the latter is weakly concentrated; a dual statement is valid for $\mathfrak{m}^{-1}(\mathfrak{f}^{-1})$.

Indeed, let $(\mathfrak{g}^{-1}, \mathfrak{g}^1)$ be as above. One can easily check that $(\mathfrak{f}^{-1} \cap \mathfrak{g}^{-1}, \mathfrak{f}^1)$ and $(\mathfrak{g}^{-1}, \mathfrak{f}^1 \cap \mathfrak{g}^1)$ are Cauchy; both are finer than $(\mathfrak{m}^{-1}(\mathfrak{g}^{-1}), \mathfrak{m}^1(\mathfrak{f}^1))$, thus $(\mathfrak{f}^{-1} \cap \mathfrak{g}^{-1}, \mathfrak{f}^1 \cap \mathfrak{g}^1)$ is Cauchy, hence so is $(\mathfrak{f}^{-1}, \mathfrak{f}^1 \cap \mathfrak{g}^1)$, implying $\mathfrak{m}^1(\mathfrak{f}^1) \subset \subset \mathfrak{f}^1 \cap \mathfrak{g}^1 \subset \mathfrak{g}^1$. Conversely, $(\mathfrak{f}^{-1}, \mathfrak{m}^1(\mathfrak{f}^1))$ is one of the filter pairs $(\mathfrak{g}^{-1}, \mathfrak{g}^1)$ considered.

0.4 Completeness. The quasi-uniformity $\mathcal{U}$ is *D-complete* ([9, 10]; terminology from [11, 15]) provided that the second element of any Cauchy filter pair (such a filter will be called *D-Cauchy*) is $\mathcal{U}^{tp}$-convergent. $\mathcal{U}$ is *C-complete* if any Cauchy filter pair is convergent (C stands for Cauchy; this notion has nothing to do with the C-completeness in the sense of [1] or [13]). Other kinds of quasi-uniform completeness defined by filter pairs can be found in [11, 6, 8].

1. Generalizations of quietness

1.1 A quasi-uniformity $\mathcal{U}$ on X is *uniformly regular* [2, 12] if for any $U \in \mathcal{U}$ there is a $V \in \mathcal{U}$ such that $\mathrm{cl}^1 Vx \subset Ux$ $(x \in X)$. Quiet quasi-uniformities are uniformly regular ([12] Proposition 1.2). The following rewording of the definition might make this connexion clearer.

PROPOSITION. *A quasi-uniformity is uniformly regular iff all the Cauchy filter pairs* $(\mathfrak{f}^{-1}, \mathfrak{f}^1)$ *for which* $\bigcap \mathfrak{f}^{-1} \neq \emptyset$ *form a uniformly weakly concentrated family.*

REMARK. When showing the sufficiency, we shall just repeat the usual proof of the fact that quiet spaces are uniformly regular.

PROOF. Let Φ denote the family of filter pairs mentioned in the proposition.

Sufficiency. Given a $U \in \mathcal{U}$, choose a $V \in \mathcal{U}$ quiet for U and Φ. Then $\mathrm{cl}^1 Vx \subset Ux$ for each x. Indeed, take a $y \in \mathrm{cl}^1 Vx$, and consider the filter pair

$$(\mathfrak{f}^{-1}, \mathfrak{f}^1) = (\mathrm{fil}\,\{\{y\}\}, \mathrm{fil}\,(\mathfrak{n}^1(y) \mid Vx)).$$

$(\mathfrak{f}^{-1}, \mathfrak{f}^1)$ is Cauchy, since it converges to y; hence $(\mathfrak{f}^{-1}, \mathfrak{f}^1) \in \Phi$. Now $V^{-1}y \in \in \mathfrak{f}^{-1}$ and $Vx \in \mathfrak{f}^1$, implying $x\, U\, y$.

Necessity. For $U \in \mathcal{U}$, take $U_0 \in \mathcal{U}$ with $U_0^2 \subset U$, then choose $V \in \mathcal{U}$ such that $\mathrm{cl}^1 Vx \subset U_0 x$ $(x \in X)$. We claim that V is quiet for U and Φ. Let $(\mathfrak{f}^{-1}, \mathfrak{f}^1) \in \Phi$, $Vx \in \mathfrak{f}^1$, $V^{-1}y \in \mathfrak{f}^{-1}$; it is to be proved that $x\,U\,y$. Pick a $z \in \bigcap \mathfrak{f}^{-1}$. The Cauchy property implies that $\mathfrak{f}^1$ $\mathcal{U}^{tp}$-converges to z, thus $z \in \mathrm{cl}^1 Vx \subset U_0 x$. On the other hand, $V^{-1}y \in \mathfrak{f}^{-1}$ implies that $z \in V^{-1}y$, so $y \in Vz \subset U_0 z$. Hence $x\,U_0\,z\,U_0\,y$, i.e. $x\,U\,y$. □

1.2 A quasi-uniformity $\mathcal{U}$ is called *doubly uniformly regular* if both $\mathcal{U}$ and $\mathcal{U}^{-1}$ are uniformly regular. We shall later need the following elementary result:

LEMMA. *A quasi-uniformity $\mathcal{U}$ is uniformly regular (doubly uniformly regular) iff for any $U \in \mathcal{U}$ there is a $V \in U$ such that $A \times B \subset V$ implies $A \times \mathrm{cl}^1 B \subset U$ $(\mathrm{cl}^{-1} A \times \mathrm{cl}^1 B \subset U)$.*

PROOF. It is enough to consider the case of uniform regularity, because it can be applied then twice to obtain the other statement.

Sufficiency. Apply the condition to $A = \{x\}$ and $B = Vx$.

Necessity. A V chosen by the definition of uniform regularity will do in the condition of the lemma, too. □

1.3 DEFINITION. A quasi-uniformity is *subquiet* if there is a uniformly concentrated family Φ of Cauchy filter pairs such that any Cauchy filter pair is finer than some element of Φ; it is *weakly quiet* if it is subquiet and all the Cauchy filter pairs are weakly concentrated. □

Φ is unique in this definition (since its elements are minimal Cauchy). A quasi-uniformity is weakly quiet iff the Cauchy filter pairs are weakly concentrated and the minimal ones are uniformly so (due to the weak concentratedness, we may indeed speak about the minimal Cauchy filter pairs). A C-complete quasi-uniformity is (i) always subquiet: Φ consists of the neighbourhood filter pairs; (ii) weakly quiet iff the Cauchy filter pairs are weakly concentrated. (The assumption that the Cauchy filter pairs are weakly concentrated seems to be of no use in itself as far as we deal with extension theory, but it can be of interest elsewhere; e.g. the proof given in [16] for the fact that totally bounded quiet quasi-uniformities are always symmetric goes through without any changement for these more general quasi-uniformities.) We have the following implications:

$$\begin{array}{ccccc} & & & & \text{C-complete} \\ & & & & \Downarrow \\ \text{quiet} & \Longrightarrow & \text{weakly quiet} & \Longrightarrow & \text{subquiet} \\ \Downarrow & & & & \\ \text{doubly uniformly regular} & & & & \end{array}$$

PROPOSITION. *Any C-complete doubly uniformly regular quasi-uniformity is quiet.*

PROOF. For $U \in \mathcal{U}$ take $V \in \mathcal{U}$ such that $\mathrm{cl}^i V^i x \subset U_0^i x$ $(x \in X,\ i = \pm 1)$ where $U_0 \in \mathcal{U}$ satisfies $U_0^2 \subset U$; then V is quiet for U. Indeed, let $(\mathfrak{f}^{-1}, \mathfrak{f}^1)$ be a Cauchy filter pair that converges to some y, and assume that $V^\iota x_{-\iota} \in \mathfrak{f}^\iota$; now $y \in \mathrm{cl}^\iota V^\iota x_{-\iota}$, thus $x_{-1}\, U_0\, y\, U_0\, x_1$. □

It is enough to know in the above proof that y is a cluster point of $(\mathfrak{f}^{-1}, \mathfrak{f}^1)$, so we have in fact proved that a doubly uniformly regular quasi-uniformity with each Cauchy filter pair having a cluster point is quiet. A quasi-uniformity with these properties is, however, C-complete: Assume only that any D-Cauchy filter has a $\mathcal{U}^{tp}$-cluster point. For $U \in \mathcal{U}$ take $S_\iota \in \mathfrak{f}^\iota$ with $S_{-1} \times S_1 \subset V$ where $\mathrm{cl}^1\, Vx \subset Ux$ $(x \in X)$. For each $x \in S_{-1}$ we have $Vx \in \mathfrak{f}^1$, thus $y \in \mathrm{cl}^1\, Vx$, implying $x\, U\, y$, i.e. $S_{-1} \subset U^{-1}y$, so $\mathfrak{f}^{-1}$ $\mathcal{U}^{-tp}$-converges to y. Now a dual argument yields that $\mathfrak{f}^1$ $\mathcal{U}^{tp}$-converges to y, too.

By the above reasoning, any uniformly regular (in particular, quiet) D-complete quasi-uniformity is C-complete (this is in fact proved in [11] 2.1, although the lemma itself says less), while C-completeness evidently implies D-completeness in any quasi-uniform space. Hence in Doitchinov's theory of completing quiet spaces we may replace D-completeness by C-completeness, which seems to be the more convenient notion when larger classes of spaces are considered.

Remark. A uniformly regular C-complete space is not necessarily quiet ([12] Example 3.2), so double uniform regularity is essential in the proposition. The one-sided analogue of the observation made just after the proof of the proposition is also false: consider the quasi-uniformity on $(\mathbf{R} \times \{0\}) \cup \cup \{(0,1)\}$ defined by the distance

$$d((x',x''),(y',y'')) = y' - x' \text{ if either } x' < 0 \leqq y' \text{ or } x' = x'' = 0 < y',\ y' \text{ is rational or } x'' = 1,\ 0 < y',\ y' \text{ is irrational}$$

is uniformly regular, each D-Cauchy filter has a cluster point, but the D-Cauchy filter $\mathrm{fil}\{]0,\varepsilon[\times\{0\} : \varepsilon > 0\}$ is not convergent. On the other hand, the completeness condition "each D-Cauchy filter has a cluster point" could be included in the theorem of [6]; it looks as if the semi-symmetry compensated for $\mathcal{U}^{-1}$ not being uniformly regular, but the spaces occurring in that theorem (semi-symmetric, uniformly regular) are in fact doubly uniformly regular: They are locally symmetric by [6] Remark b), hence point-symmetric, which means that $\mathcal{U}^{-tp}$ is finer than $\mathcal{U}^{tp}$; but if $\mathcal{U}$ is point-symmetric then $\mathcal{U}^{-1}$ is uniformly regular: $\mathrm{cl}^{-1}\, U^{-1}x \subset \mathrm{cl}^1\, U^{-1}x \subset U^{-2}x$. (This observation on point-symmetric spaces makes it also possible to deduce [12] Theorem 2.1 from the proposition.) *Added in proof.* Any semi-symmetric uniformly regular quasi-uniformity is quiet, see Theorem 4 in: Künzi, H.-P., Mršević, M., Reilly, I. L. and Vamanamurthy, M. K., Convergence, precompactness and symmetry in quasi-uniform spaces, *Math. Japon.* **38** (1993), 239–253.

Examples. No further implications are valid between the properties shown in the diagram. Notations: $\mathbf{Q}$ = the rationals, $\mathbf{D}$ = the dyadic rationals, $\mathbf{I}=\mathbf{R}\setminus\mathbf{Q}$, $\mathbf{R}_0=\mathbf{R}\setminus\{0\}$,

$$\mathfrak{f}^i_A=\operatorname{fil}_{\mathbf{R}_0}\{]0,i\varepsilon[\cap A:\varepsilon>0\}\qquad (A\subset\mathbf{R},\ i=\pm 1).$$

a) *Quiet, not C-complete.* $\mathcal{U}_{so}\,|\,\mathbf{R}_0$.

b) *Doubly uniformly regular, not subquiet.* On $\mathbf{R}_0$, let $\mathcal{U}$ be induced by the distance

$$d(x,y)=y-x \text{ if } x<0<y \text{ and either } x,y\in\mathbf{Q}$$
$$\text{or } x\in\mathbf{Q},\ y\in\mathbf{I},\ -x<y$$
$$\text{or } x\in\mathbf{I},\ y\in\mathbf{Q},\ y<-x.$$

$\mathcal{U}$ is doubly uniformly regular, since its bitopology is discrete. $\mathcal{U}$ is not subquiet: the filter pair $(\mathfrak{f}^{-1}_{\mathbf{Q}},\mathfrak{f}^{1}_{\mathbf{Q}})$ is minimal Cauchy, but not weakly concentrated.

c) *C-complete, not weakly quiet.* A non-symmetric quasi-uniformity on a two point set.

d) *Subquiet, doubly uniformly regular, not weakly quiet.* Let $\mathcal{U}$ be as in b), but with $\mathbf{D}$ substituted for $\mathbf{Q}$ in the second and third lines of the definition of d. The filter pair $(\mathfrak{f}^{-1}_{\mathbf{Q}},\mathfrak{f}^{1}_{\mathbf{Q}})$ is now concentrated and it is coarser than any non-convergent Cauchy filter pair. On the other hand, $(\mathfrak{f}^{-1}_{\mathbf{D}},\mathfrak{f}^{1}_{\mathbf{D}})$ is Cauchy, but not weakly concentrated.

e) *Weakly quiet, doubly uniformly regular, not quiet.* Let $X=\mathbf{R}_0$, $n\in\mathbf{N}$ and

$$e_n(x,y)=\begin{cases} y-x & \text{if } x<0<y,\ (x\in\mathbf{Q} \text{ or } y\in\mathbf{Q}),\\ n(y-x) & \text{if } x<0<y,\ x,y\in\mathbf{I}.\end{cases}$$

Now $\mathcal{U}_n=\mathcal{U}(d_n)$ is subquiet, since any non-convergent Cauchy filter pair is finer than $(\mathfrak{f}^{-1}_{\mathbf{R}},\mathfrak{f}^{1}_{\mathbf{R}})$, which is concentrated, with $U_{(\varepsilon/2)}$ quiet for $U_{(\varepsilon)}$. $\mathcal{U}_n$ is in fact quiet: $U_{(\varepsilon/2n)}$ is quiet for $U_{(\varepsilon)}$; but U_δ with $\delta>\varepsilon/2n$ is not quiet for $U_{(\varepsilon)}$ and $(\mathfrak{f}^{-1}_{\mathbf{Q}},\mathfrak{f}^{1}_{\mathbf{Q}})$. Taking now $X=\mathbf{R}_0\times\mathbf{N}$, and the distance d_n on $\mathbf{R}_0\times\{n\}$, we obtain a quasi-uniformity on X that is weakly quiet but not quiet; it is doubly uniformly regular, as its bitopology is discrete.

f) *C-complete, weakly quiet, not doubly uniformly regular.* For $n\geqq 2$, define a distance d_n on $\mathbf{R}$ as follows:

$$d_n(x,y)=\begin{cases} y-x & \text{if } x\leqq 0<y,\\ n(y-x) & \text{if } x<0=y.\end{cases}$$

$\mathcal{U}_n=\mathcal{U}(d_n)$ is clearly C-complete; it is weakly quiet, since $U_{(\varepsilon/n)}$ is quiet for $U_{(\varepsilon)}$. But no $U_{(\delta)}$ with $\delta>\varepsilon/n$ is quiet for $U_{(\varepsilon)}$ and $(\mathfrak{f}^{-1}_{\mathbf{R}},\mathfrak{f}^{1}_{\mathbf{R}})$. The method used at the end of the preceding example yields a C-complete, weakly quiet, not quiet quasi-uniformity; it is not doubly uniformly regular by the proposition. □

2. Complete extensions

2.1 Let trace filter pairs $(\mathfrak{f}^{-1}(a), \mathfrak{f}^{1}(a))$ $(a \in Y \supset X)$ be prescribed in the quasi-uniform space $(X, \mathcal{U})$ (assuming, of course, that $\mathfrak{f}^{\prime}(x)$ is the $\mathcal{U}^{\prime tp}$-neighbourhood filter of $x \in X$). In [5b] we defined

$$a\,{}^4U\,b \text{ iff there are } A \in \mathfrak{f}^{-1}(a) \text{ and } B \in \mathfrak{f}^{1}(b) \text{ with } A \times B \subset U$$

(for $a, b \in Y$ and U an entourage in X), and proved that $\{{}^4U : U \in \mathcal{U}\}$ is a base for a quasi-uniform extension inducing the given trace filter pairs iff they are uniformly concentrated; if so then this quasi-uniformity, denoted by ${}^4\mathcal{U}$, is the coarsest extension with these trace filter pairs, and $({}^4\mathcal{U}^{-tp}, {}^4\mathcal{U}^{tp})$ is a doubly strict extension of $(\mathcal{U}^{-tp}, \mathcal{U}^{tp})$. If $\mathcal{B}$ is a (sub)base for $\mathcal{U}$ then $\{{}^4U : U \in \mathcal{B}\}$ is a (sub)base for ${}^4\mathcal{U}$, since ${}^4(U \cap V) = {}^4U \cap {}^4V$.

Assume now that $(X, \mathcal{U})$ is a subquiet space, and let $({}^{\mathrm{C}}X, {}^{\mathrm{C}}\mathcal{U})$ denote ${}^4U$ taken with all the concentrated filter pairs, such that there is a bijection between ${}^{\mathrm{C}}X \setminus X$ and the non-convergent concentrated filter pairs (assume e.g. that these filter pairs themselves are the elements of ${}^{\mathrm{C}}X \setminus X$); we shall avoid the more precise but cumbersome notations ${}^{\mathrm{C}}(X, \mathcal{U})$ or $({}^{\mathrm{C}}X(\mathcal{U}), {}^{\mathrm{C}}\mathcal{U})$. More generally, ${}^4X$ denotes the fundamental set of ${}^4\mathcal{U}$. Doitchinov [9, 10] proved that if $\mathcal{U}$ is quiet then ${}^{\mathrm{C}}\mathcal{U}$ is quiet and D-complete (he constructed ${}^{\mathrm{C}}\mathcal{U}$ in a slightly different but equivalent way) and showed that this *completion* has several good properties. Subquietness will, however, turn out to be sufficient for the C-completeness of ${}^{\mathrm{C}}\mathcal{U}$. Moreover, if $\mathcal{U}$ is weakly quiet then so is ${}^{\mathrm{C}}\mathcal{U}$; the analogous result for quietness will be re-proved at the same time.

LEMMA. *If $(Y, \mathcal{V})$ is a doubly strict extension of $(X, \mathcal{U})$, and for any $\mathcal{U}$-Cauchy filter pair there exists a trace filter pair coarser than it then $\mathcal{V}$ is C-complete.*

PROOF. According to [3] 1.1, if $\mathcal{V}$ is a strict extension, $(\mathfrak{f}^{-1}, \mathfrak{f}^{1})$ is $\mathcal{V}$-Cauchy, $\mathfrak{f}^{1}$ is $\mathcal{V}^{tp}$-open, and $\mathfrak{f}^{1}(a) \subset \mathfrak{f}^{1} \mid X$ then $\mathfrak{f}^{1}$ $\mathcal{V}^{tp}$-converges to a.

To prove the C-completeness, it is enough to show that any round Cauchy filter pair $(\mathfrak{f}^{-1}, \mathfrak{f}^{1})$ is convergent in $(Y, \mathcal{V})$. As $(\mathfrak{f}^{-1} \mid X, \mathfrak{f}^{1} \mid X)$ is $\mathcal{U}$-Cauchy, it is coarser than the trace filter pair of some $a \in Y$, so the result cited above can be applied to $\mathcal{V}$ as well as to $\mathcal{V}^{-1}$, giving that $(\mathfrak{f}^{-1}, \mathfrak{f}^{1})$ converges to a. □

THEOREM. *If $\mathcal{U}$ is subquiet then ${}^{\mathrm{C}}\mathcal{U}$ is C-complete.*

PROOF. The above lemma, recalling the definition of subquietness and using that ${}^4\mathcal{U}$ is always a doubly strict extension. □

2.2 THEOREM. *If $\mathcal{U}$ is (weakly) quiet then so is ${}^4\mathcal{U}$ taken with arbitrary concentrated trace filter pairs. In particular, ${}^{\mathrm{C}}\mathcal{U}$ is (weakly) quiet, too.*

PROOF. Let $(\mathfrak{f}^{-1}, \mathfrak{f}^{1})$ be a round Cauchy filter pair in ${}^4X$, and $U \in \mathcal{U}$. As $(\mathfrak{f}^{-1} \mid X, \mathfrak{f}^{1} \mid X)$ is $\mathcal{U}$-Cauchy, there is a $V \in \mathcal{U}$ quiet for $\mathcal{U}$ and this filter pair. We claim that if $W \in {}^4\mathcal{U}$ satisfies $W^2 \subset {}^4V$ then W is quiet for ${}^4U$ and $(\mathfrak{f}^{-1}, \mathfrak{f}^1)$.

Indeed, assume that $Wa \in \mathfrak{f}^1$, $W^{-1}b \in \mathfrak{f}^{-1}$. Then $Wa \cap X \in \mathfrak{f}^1 \mid X$. For each $x \in W^{-1}a \cap X$ and $z \in Wa \cap X$ we have $x\, W^2\, z$, i.e. $x\, {}^4V\, z$, implying $x\, V\, z$. Therefore $Vx \in \mathfrak{f}^1 \mid X$ for each $x \in W^{-1}a \cap X = A \in \mathfrak{f}^{-1}(a)$. Analogously, $V^{-1}y \in \mathfrak{f}^{-1} \mid X$ for each $y \in Wb \cap X = B \in \mathfrak{f}^1(b)$. Now $x\, U\, y$ follows from the choice of V, thus $A \times B \subset U$, i.e. $a\, {}^4U\, b$.

If $\mathcal{U}$ is quiet then V, and therefore W, too, can be chosen independently of $(\mathfrak{f}^{-1}, \mathfrak{f}^1)$, hence ${}^4\mathcal{U}$ is quiet. If $\mathcal{U}$ is only weakly quiet then, as we have seen, the round Cauchy filter pairs, hence all the Cauchy filter pairs, are weakly concentrated in ${}^4X$. To complete the proof, it is enough to show that the trace of a minimal Cauchy filter pair is minimal Cauchy, since then, using that $\mathcal{U}$ is subquiet, we can again choose V and W independently of the given minimal Cauchy filter pair.

So assume that $(\mathfrak{f}^{-1}, \mathfrak{f}^1)$ is a minimal Cauchy filter pair in ${}^4X$, and $(\mathfrak{g}^{-1}, \mathfrak{g}^1)$ is a $\mathcal{U}$-Cauchy filter pair coarser than $(\mathfrak{f}^{-1} \mid X, \mathfrak{f}^1 \mid X)$. Now, with fil understood in ${}^4X$, $(\mathfrak{f}^{-1}, \mathfrak{f}^1)$ and $(\operatorname{fil} \mathfrak{g}^{-1}, \operatorname{fil} \mathfrak{g}^1)$ are Cauchy filter pairs coarser than $(\operatorname{fil}(\mathfrak{f}^{-1} \mid X), \operatorname{fil}(\mathfrak{f}^1 \mid X))$; but the latter is weakly concentrated, so there is a coarsest one among the Cauchy filter pairs coarser than it, hence, $(\mathfrak{f}^{-1}, \mathfrak{f}^1)$ being minimal Cauchy, we have $\mathfrak{f}^i \subset \operatorname{fil} \mathfrak{g}^i$, implying $\mathfrak{f}^i \mid X \subset \mathfrak{g}^i$. □

REMARK. The last part of the proof is superfluous in the special case of ${}^C\mathcal{U}$, since it is C-complete (and so subquiet) by Theorem 2.1.

2.3 Doitchinov proved that any quasi-uniformly continuous map from a quiet space $(X, \mathcal{U})$ into a quiet D-complete (= C-complete) space can be quasi-uniformly continuously extended onto ${}^CX$. No similar theorem holds for weakly quiet or subquiet spaces: it can occur that $\mathcal{U}$ is quiet, and it has a weakly quiet C-complete extension $(Y, \mathcal{V})$ such that $Y = {}^CX$, and $\mathcal{V}$ is strictly finer than ${}^C\mathcal{U}$:

EXAMPLE. On $X = (\mathbf{R} \setminus \{0\}) \times H$, where $H = \{1/n \colon n \in \mathbf{N}\}$, take the distance

$$d((x', x''), (y', y'')) = y' - x' + y'' - x'' \quad \text{if} \quad x' \leqq y', x'' \leqq y'',$$

and let $\mathcal{U} = \mathcal{U}(d)$ (i.e. $\mathcal{U} = \mathcal{U}^2_{\mathrm{so}} \mid X$). $\mathcal{U}$ is quiet. ${}^CX$ can be identified with $\mathbf{R} \times H$, on which a weakly quiet C-complete extension $\mathcal{V}$ strictly finer than ${}^C\mathcal{U}$ can be defined by the distance

$$e((a', a''), (b', b'')) = b' - a' + b'' - a'' \text{ if } \; a' \leqq b', a'' \leqq b'', \; (a', b') \neq (0, 0). \quad \square$$

If the point $(0,0)$ is added to X in the above example then the same construction gives a weakly quiet C-complete extension that is not doubly strict.

REMARK. In [5b] we introduced ${}^2\mathcal{U}$, which is an extension for the prescribed trace filter pairs under the same conditions as ${}^4\mathcal{U}$. The quasi-uniformity $\mathcal{V}$ in the example is in fact ${}^2\mathcal{U}$. It is, however, not true that ${}^2\mathcal{U}$ is always C-complete: Taking the same example with H replaced by $\{0\}\cup H$, the filter pair

$$(\mathrm{fil}\,\{]-\varepsilon,0[\times\{0\}:\varepsilon>0\},\mathrm{fil}\,\{\{(0,1/k):k\geqq n\}:n\in\mathbb{N}\})$$

is ${}^2\mathcal{U}$-Cauchy, but not convergent (the points $(0,1/k)$ are to be understood as in the example).

2.4 PROPOSITION. *If $(X,\mathcal{U})$ is subquiet then no proper subspace of $({}^{\mathrm{C}}X,{}^{\mathrm{C}}\mathcal{U})$ is a C-complete extension of $(X,\mathcal{U})$.*

PROOF. Assume that $X\subset Y\subsetneqq {}^{\mathrm{C}}X$, and ${}^{\mathrm{C}}\mathcal{U}\,|\,Y$ is C-complete. Pick a $p\in {}^{\mathrm{C}}X\setminus Y$. Now $(\mathfrak{f}^{-1}(p),\mathfrak{f}^{1}(p))$, which is not convergent in $(X,\mathcal{U})$, converges to some $q\in Y\setminus X$ in ${}^{\mathrm{C}}\mathcal{U}\,|\,Y$, implying that $(\mathfrak{f}^{-1}(q),\mathfrak{f}^{1}(q))$ is coarser than $(\mathfrak{f}^{-1}(p),\mathfrak{f}^{1}(p))$, a contradiction, since the trace filter pairs are minimal Cauchy, and the non-convergent ones belong to one point only. □

2.5 DEFINITION. An extension Y of a bitopological space X is *reduced* provided that the neighbourhood filter pairs of p and a are different whenever $p\in Y\setminus X$, $a\in Y$, $p\neq a$. □

Starting from a C-complete extension of a quasi-uniformity, we can obtain a reduced C-complete extension just by weeding out the superfluous points; (weak) quietness is clearly preserved in this process. ${}^{\mathrm{C}}\mathcal{U}$ is always reduced; we are going to show that it is, up to isomorphism, the only reduced quiet C-complete extension of a quiet quasi-uniformity $\mathcal{U}$. In particular, if $\mathcal{U}$ is quiet and T_0 then ${}^{\mathrm{C}}\mathcal{U}$ is essentially the only quiet C-complete T_0 extension of $\mathcal{U}$.

THEOREM. *If $(Y,\mathcal{V})$ is a reduced quiet C-complete extension of $(X,\mathcal{U})$ then there exists an isomorphism f from $(Y,\mathcal{V})$ onto $({}^{\mathrm{C}}X,{}^{\mathrm{C}}\mathcal{U})$ satisfying $f(x)=x$ $(x\in X)$.*

We begin with proving a lemma.

LEMMA. *Let $(Y,\mathcal{V})$ and $(Y,\mathcal{W})$ be extensions of $(X,\mathcal{U})$ belonging to the same trace filter pairs, and assume that $\mathcal{V}$ is doubly uniformly regular. Then $\mathcal{V}\subset\mathcal{W}$.*

PROOF. Let $V\in\mathcal{V}$ be fixed, and choose $V_0\in\mathcal{V}$ such that $\mathrm{Cl}^{-1}A\times\mathrm{Cl}^{1}B\subset V$ whenever $A\times B\subset V_0$ (Lemma 1.2; $\mathrm{Cl}^{\imath}$ is the $\mathcal{V}^{\imath tp}$-closure). Pick a $W\in\mathcal{W}$ such that $W^3|X\subset V_0|X$. It is enough to show that $W\subset V$.

Assume $a\,W\,b$ and define $A=W^{-1}a\cap X$, $B=Wb\cap X$. Now $A\times B\subset W^3\,|\,X\subset V_0$, implying $\mathrm{Cl}^{-1}A\times\mathrm{Cl}^{1}B\subset V$. As $A\in\mathfrak{f}^{-1}(a)$, we have $a\in\mathrm{Cl}^{-1}A$, and analogously $b\in\mathrm{Cl}^{1}B$, i.e. $a\,V\,b$. □

PROOF OF THE THEOREM. Let the notation $(\mathfrak{f}^{-1}(a), \mathfrak{f}^1(a))$ be used for the $\mathcal{V}$-trace filter pairs. Any non-convergent concentrated filter pair $(\mathfrak{f}^{-1}, \mathfrak{f}^1)$ in $(X, \mathcal{U})$ $\mathcal{V}$-converges to some $a \in Y$; $(\mathfrak{f}^{-1}, \mathfrak{f}^1)$ being minimal Cauchy, we have $\mathfrak{f}^i(a) = \mathfrak{f}^i$. Choose such a point for each non-convergent concentrated filter pair, and let Y_0 consist of these points, and of the points of X. It is now enough to prove that $Y_0 = Y$, because then Y can be identified with ${}^{\mathrm{C}}X$, and the lemma applies ($\mathcal{V}$ and ${}^{\mathrm{C}}\mathcal{U}$ are both doubly uniformly regular).

Pick a point $p \in Y \setminus X$. $(\mathfrak{f}^{-1}(p), \mathfrak{f}^1(p))$ is Cauchy, so it is finer than a concentrated filter pair, which is of the form $(\mathfrak{f}^{-1}(a), \mathfrak{f}^1(a))$ for some $a \in Y_0$. The double regularity of $\mathcal{V}$ implies that p is in any $\mathcal{V}^{itp}$-neighbourhood of a and vice versa $(i = \pm 1)$. Thus the neighbourhood filter pairs of p and a coincide, hence $p = a$, since $\mathcal{V}$ is a reduced extension. Therefore $p \in Y_0$. □

This theorem on unicity is only valid in bitopological context: with $X =]0,1[$ and $Y = [0,1[$, $\mathcal{U} = \mathcal{U}_{\mathrm{so}} \mid X$ is quiet and D-complete, thus $\mathcal{U} = {}^{\mathrm{C}}\mathcal{U}$; but $\mathcal{V} = \mathcal{U}_{\mathrm{so}} \mid Y$ would be another quiet D-complete T_0 extension if $\mathcal{V}^{tp}$-density were only required in the definition of an extension.

3. Complete extensions of totally bounded quasi-uniformities

3.1 Assume that $\mathcal{U}$ is totally bounded and subquiet. Then ${}^{\mathrm{C}}\mathcal{U}$ is C-complete, and also totally bounded, since its totally bounded reflexion is a coarser extension (cf. [13] 1.29) inducing the same trace filter pairs (cf. [13] 1.30), and ${}^{\mathrm{C}}\mathcal{U}$ was already the coarsest one (cf. the first paragraph of 2.1). A weakly quiet totally bounded quasi-uniformity is symmetric (see the paragraph after Definition 1.3), but there do exist non-symmetric subquiet totally bounded quasi-uniformities (e.g. Example 1.3 c); another example, with better separation properties: on $X = [0,1]$, let $d(x,y) = |y - x|$ if $y \neq 0$).

It will be proved in [8] that for any Cauchy filter pair in a totally bounded quasi-uniform space there is a (not necessarily unique) concentrated Cauchy filter pair coarser than it. Hence subquietness means here that the minimal Cauchy filter pairs are not just concentrated, but uniformly so. Thus all those totally bounded quasi-uniform spaces are subquiet in which there are only a finite number of non-convergent minimal Cauchy filter pairs (since a finite family of concentrated filter pairs is clearly uniformly concentrated, while the convergent ones are the same as the neighbourhood filter pairs, which are uniformly concentrated, too).

A construction different form ${}^{\mathrm{C}}\mathcal{U}$ yields a totally bounded C-complete extension for any totally bounded quasi-uniformity, see in [8]. (This is, however, a very bad construction: there can be new points inducing convergent trace filter pairs.)

3.2 A totally bounded non-symmetric quasi-uniformity with infinitely many non-convergent minimal Cauchy filter pairs may, or may not, be subquiet:

EXAMPLES. a) With e denoting the Euclidean metric on $\mathbf{R}^3$, $O=(0,0,0)$, consider on

$$X=\{(x',x'',x''')\in\mathbf{R}^3\colon 0<|x'|=|x''|\leqq 1,\ (x'''=0 \text{ or } 1/x'''\in\mathbf{N})\}\cup\{v,w\}$$

$(v,w\notin\mathbf{R}^3,\ v\neq w)$ the distance

$$d(x,y)=\begin{cases} e(x,y) & \text{if } x=(x',x'',x'''), y=(y',y'',y''') \\ & \qquad\text{and either } x'<0<y' \\ & \qquad\qquad\text{or } x'y'>0<x''y'', \\ e(x,O) & \text{if } x=(x',x'',x'''),\ x'<0,\ y=v, \\ e(O,y) & \text{if } y=(y',y'',y'''),\ 0<y',\ x=w. \end{cases}$$

For $n\in\mathbf{N}$, minimal Cauchy filter pairs $(\mathfrak{f}_n^{-1},\mathfrak{f}_n^{1})$ are defined by

$$\text{(1)}\qquad \mathfrak{f}_n^{i}=\mathrm{fil}\,\{(]0,i\varepsilon[\times\mathbf{R}\times\{1/n\})\cap X:\varepsilon>0\}.$$

(There are others that need not concern us now.) No $U_{(\delta)}$ is quiet for a given $U_{(\varepsilon)}$ and these filter pairs, since if $n>1/\delta$ then $U_{(\delta)}w\in\mathfrak{f}_n^1$, $U_{(\delta)}^{-1}v\in\mathfrak{f}_n^{-1}$, but $(w,v)\notin U_{(\varepsilon)}$. Thus $\mathcal{U}(d)$ is not subquiet.

b) With $\mathcal{U}(d)$ from a), $\mathcal{U}(d)\mid(X\setminus\{v,w\})$ is subquiet. Indeed, $U_{(\varepsilon/2)}$ is quiet for $U_{(\varepsilon)}$ and the minimal Cauchy filter pairs: it is enough to check this for the filter pairs (1) and for $(\mathfrak{f}_0^{-1},\mathfrak{f}_0^{1})$, where

$$\mathfrak{f}_0^{i}=\mathrm{fil}\,\{(]0,i\varepsilon[\times\mathbf{R}\times[0,\varepsilon[)\cap X:\varepsilon>0\},$$

because the other minimal Cauchy filter pairs are linked (i.e. $S_{-1}\cap S_1\neq\emptyset$ for $S_i\in\mathfrak{f}^i$), and such filter pairs are always uniformly weakly concentrated (evident; or see [5b] 7.9). □

3.3 Let $\mathcal{U}$ be a subquiet totally bounded quasi-uniformity, δ and ${}^{\mathrm{C}}\delta$ the quasi-proximities induced by $\mathcal{U}$ and ${}^{\mathrm{C}}\mathcal{U}$, respectively. We are going to describe ${}^{\mathrm{C}}\delta$ in terms of δ. First of all, a filter pair is $\mathcal{U}$-Cauchy iff it is δ-compressed ([5a] Lemma 5.1), where *δ-compressed* means that $A\in\sec\mathfrak{f}^{-1}$, $B\in\sec\mathfrak{f}^1$ imply $A\,\delta\,B$.

More generally, let $(X,\mathcal{U})$ be an arbitrary totally bounded quasi-uniform space, and consider $(Y,{}^4\mathcal{U})$ with some uniformly concentrated trace filter pairs; let δ and ${}^4\delta$ be defined as δ and ${}^{\mathrm{C}}\delta$ above. Recall that

$$\mathcal{S}=\{U_{A,B}=X\times X\setminus A\times B\colon A\,\overline{\delta}\,B\}$$

is a subbase for $\mathcal{U}$ (e.g. [14] 1.33); if $\mathcal{S}_0\subset\mathcal{S}$ is another subbase for $\mathcal{U}$ then $A\,\overline{\delta}\,B$ iff there are finite collections $\mathcal{A}$ and $\mathcal{B}$ such that $A=\bigcup\mathcal{A}$, $B=\bigcup\mathcal{B}$, and for any $A'\in\mathcal{A}$, $B'\in\mathcal{B}$, there exist $A''\supset A'$, $B''\supset B'$ such that $U_{A'',B''}\in\mathcal{S}_0$

(this follows e.g. from [4] 3.39). Denoting the ${}^4\mathcal{U}$-closures by Cl^i, one can easily check that

$${}^4U_{A,B} = Y \times Y \setminus \mathrm{Cl}^{-1} A \times \mathrm{Cl}^1 B.$$

These entourages form a subbase for ${}^4\mathcal{U}$, thus

(1) $A\, \overline{{}^4\delta}\, B$ iff there are finite collections $\mathcal{A}$ and $\mathcal{B}$ such that $A = \bigcup\mathcal{A}$, $B = \bigcup\mathcal{B}$, and for any $A' \in \mathcal{A}$, $B' \in \mathcal{B}$ there exist $A'', B'' \subset X$ with $A''\, \overline{\delta}\, B''$, $A' \subset \mathrm{Cl}^{-1} A''$, $B' \subset \mathrm{Cl}^1 B''$.

(Or, using the trace filters, $A'' \in \sec \mathfrak{f}^{-1}(a)$ for each $a \in A'$, $B'' \in \sec \mathfrak{f}^1(b)$ for each $b \in B'$.)

In the special case when the trace filter pairs are linked (see in Example 3.2 b)), [5b] Theorem 11.2 states the ${}^4\mathcal{U}$ coincides with another construction, denoted there by ${}^0\mathcal{U}$; the description of ${}^0\mathcal{U}$ given in [5a] 5.3 in terms of the induced quasi-proximity yields that

(2) $A\, \overline{{}^4\delta}\, B$ iff $A \cap B = \emptyset$ and there are $A'', B'' \subset X$ such that $A''\, \overline{\delta}\, B''$, $A'' \in \mathfrak{f}^1(a)$ $(a \in A)$, $B'' \in \mathfrak{f}^{-1}(b)$ $(b \in B)$.

The trace filter pairs being linked, the condition $A \cap B = \emptyset$ is now redundant. From (1) and (2) we obtain

(3) $A\, \overline{{}^4\delta}\, B$ iff there are $A'', B'' \subset X$ such that $A''\, \overline{\delta}\, B''$, $A \subset \mathrm{Cl}^{-1} A''$, $B \subset \mathrm{Cl}^1 B''$,

since if $A\, \overline{{}^4\delta}\, B$ then the condition in (3) holds with A'' and B'' from (2) (again because of the linkedness), and conversely the condition in (3) is clearly stronger than the one in (1). (In particular, if all the non-convergent linked round compressed filter pairs are considered then (3) gives one of the several possible descriptions of the usual supremum-compactification of δ.)

The next example shows that, in general, the construction (1) for ${}^C\delta$ (where δ is subquiet) cannot be reduced to (3).

EXAMPLE. Let $X = [-1, 0[\times \{\pm 1, \pm 4, \pm 5\} \bigcup]0, 1] \times \{\pm 2, \pm 3\}$,

$$d((x', x''), (y', y'')) = |y' - x'| \quad \text{if either } x'' = y'' \text{ or } x' < 0 < y',\ |x'' - y''| \leqq 4.$$

$\mathcal{U}(d)$ is totally bounded and subquiet (as there are only a finite number of non-convergent minimal Cauchy filter pairs). The following filter pairs are minimal Cauchy (there are others, too):

$$(\mathfrak{f}^{-1}(q), \mathfrak{f}^1(q)) = (\mathfrak{g}^{-1}\{\pm 1\}, \mathfrak{g}^1\{\pm 2, \pm 3\}),$$

$$(\mathfrak{f}^{-1}(p_j), \mathfrak{f}^1(p_j)) = (\mathfrak{g}^{-1}\{\pm 1, j4, j5\}, \mathfrak{g}^1\{j2, j3\}) \qquad (j = \pm 1),$$

where $\mathfrak{g}^i H = \mathrm{fil}\,\{]0, i\varepsilon[\times H : 0 < \varepsilon < 1\}$. Check now that $\{p_{-1}, p_1\}\,\overline{{}^{\mathrm{C}}\delta}\,\{q\}$, but (3) fails for these two sets. □

REMARK. It follows from [17] Proposition 20 that if δ is not symmetric then ${}^{\mathrm{C}}\delta$ cannot be T_1.

4. Quiet extensions for prescribed bitopologies

4.1 In the spirit of [5a, 5b], we consider the following problem: Let $(X, \mathcal{U})$ be a quasi-uniform space, $(Y, \mathcal{S}^{-1}, \mathcal{S}^1)$ an extension of the bitopological space $(X, \mathcal{U}^{-tp}, \mathcal{U}^{tp})$; under what conditions does there exist a quiet extension $\mathcal{V}$ of $\mathcal{U}$ compatible with $(\mathcal{S}^{-1}, \mathcal{S}^1)$? (In [5a, 5b] and [7] § 3, analogous problems with other properties instead of quietness were investigated.)

THEOREM. *A quasi-uniformity $\mathcal{U}$ has a quiet extension to a prescribed extension of the induced bitopology iff $\mathcal{U}$ is quiet, the bitopological extension is doubly strict, and the trace filter pairs are minimal Cauchy.*

PROOF. *Sufficiency.* Theorem 2.2 and the first paragraph of 2.1, recalling that in a quiet space the minimal Cauchy filter pairs are uniformly concentrated.

Necessity. Assume that $\mathcal{V}$ is a quiet extension. Then $(\mathcal{S}^{-1}, \mathcal{S}^1)$ is doubly regular, hence doubly strict. Quietness is a hereditary property, so $\mathcal{U}$ is quiet, too. The trace filter pairs $(\mathfrak{f}^{-1}(p), \mathfrak{f}^1(p))$ $(p \in Y \setminus X)$ are Cauchy; we have only to show that they are minimal Cauchy; repeat for this purpose the reasoning from the last paragraph of the proof of Theorem 2.2, with $(\mathfrak{f}^{-1}, \mathfrak{f}^1)$ the neighbourhood filter pair of p, and ${}^4X$ replaced by Y. □

PROBLEMS. a) Investigate the question of quiet extensions in a *topological* space.

b) The same problem in topological or bitopological spaces for subquietness or weak quietness. (It may cause difficulties that these properties are not hereditary.)

ACKNOWLEDGEMENT. The author wishes to thank Prof. Á. Császár for spotting a wrong counterexample in the first version of this paper.

REFERENCES

[1] CARTER, K. S. and HICKS, T. L., Some results on quasi-uniform spaces, *Canad. Math. Bull.* **19** (1976), No. 1, 39–51. *MR* **54** # 11284

[2] CSÁSZÁR, Á., Extensions of quasi-uniformities, *Acta Math. Acad. Sci. Hungar.* **37** (1981), No. 1–3, 121–145. *MR* **82f**: 54039

[3] CSÁSZÁR, Á., *D*-complete extensions of quasi-uniform spaces, *Acta Math. Hungar.* **6** (1994), No. 1, 41–54.

[4] DEÁK, J., Preproximities and internal characterizations of complete regularity, *Studia Sci. Math. Hungar.* **24** (1989), No. 2–3, 147–177.

[5A] DEÁK, J., Quasi-uniform extensions for prescribed bitopologies I, *Studia Sci. Math. Hungar.* **25** (1990), No. 1-2, 45–67.

[5B] DEÁK, J., Quasi-uniform extensions for prescribed bitopologies II, *Studia Sci. Math. Hungar.* **25** (1990), No. 1-2, 69–91.

[6] DEÁK, J., On the coincidence of some notions of quasi-uniform completeness defined by filter pairs, *Studia Sci. Math. Hungar.* **26** (1991), 411–413.

[7] DEÁK, J., A survey of compatible extensions (presenting 77 unsolved problems), *Topology, theory and applications II* (Proc. Sixth Colloq., Pécs, 1989), Colloq. Math. Soc. J. Bolyai **55**, North-Holland, Amsterdam, 1993, 127–175.

[8] DEÁK, J., A bitopological view of quasi-uniform completeness I–III, *Studia Sci. Math. Hungar.* (to appear).

[9] DOITCHINOV, D., On completeness of quasi-uniform spaces, *C. R. Acad. Bulgar. Sci.* **41** (1988), No. 7, 5–8.

[10] DOITCHINOV, D., A concept of completeness of quasi-uniform spaces, *Topology Appl.* **38** (1991), 205–217.

[11] FLETCHER, P. and HUNSAKER, W., Completeness using pairs of filters, *Topology Appl.* **44** (1992), 149–155.

[12] FLETCHER, P. and HUNSAKER, W., Uniformly regular quasi-uniformities, *Topology Appl.* **37** (1990), 285–291.

[13] FLETCHER, P. and LINDGREN, W. F., C-complete quasi-uniform spaces, *Arch. Math.* (Basel) **30** (1978), No. 2, 175–180. *MR* **58** #7562

[14] FLETCHER, P. and LINDGREN, W. F., *Quasi-uniform spaces*, Lecture Notes in Pure Appl. Math. **77**, Marcel Dekker, New York, 1982. *MR* **84h**:54052

[15] KOPPERMAN, R. D., Total boundedness and compactness for filter pairs, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **33** (1990), 25–30.

[16] KÜNZI, H.-P. A., Totally bounded quiet quasi-uniformities, *Topology Proc.* **15** (1990), 113–115.

[17] LINDGREN, W. F. and FLETCHER, P., A construction of the pair completion of a quasi-uniform space, *Canad. Math. Bull.* **21** (1978), No. 1, 53–59. *MR* **58** #7562

(Received April 9, 1991)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 363–367

RIGHT LIE ALGEBRAS AND THEIR MULTIPLICATION ALGEBRAS

S. TUMURBAT

In [3] we have introduced, among others, a Hoehnke radical T which turned out to be useful in the study of right Lie algebras.

It is the purpose of this note to characterize T-semisimple right Lie algebras in terms of Properties of the right multiplication subalgebra and of the Baer radical of its multiplication algebra.

A not necessarily associative algebra L over a commutative ring with 1 is called a *right Lie algebra* if it satisfies the additional identity

$$x(yz) = (xy)z + y(xz) \qquad \forall x, y, z \in L.$$

We have proved in [2] that also the identity

$$(1) \qquad (xy + yx)z = 0 \qquad \forall x, y, z \in L$$

is satisfied by every right Lie algebra L. Let Z_L denote the ideal of L generated by the set

$$\{x^2, yz + zy \mid x, y, z \in L\}.$$

Then by (1) we have $Z_L L = 0$.

It is known (see [1], Theorem 1.3.2) that for any Lie algebra L and any element $a \in L$ the right multiplication

$$R_a \colon L \to La$$

defined by $(x)R_a = xa$, $x \in L$, is a derivation of L and the mapping ad: $L \to$ $\to \operatorname{ad} L$ defined by $\operatorname{ad}(x) = R_x$, $x \in L$, is a homomorphism onto $\operatorname{ad} L$ and obviously $\operatorname{Ker} \operatorname{ad} = \{a \in L \mid La = 0\}$.

As is well known, for any not necessarily associative algebra L one can define right and left multiplications correspondingly, and the subalgebra $M(L)$ of endomorphisms of L generated by all right and left multiplications (in [4]) is an associative algebra called the *multiplication algebra* of L. The subalgebra generated by all right multiplications is the *right multiplication algebra* and will be denoted by $R(L)$.

We shall need some calculation rules concerning right and left multiplications of a right Lie algebra.

1991 *Mathematics Subject Classification.* Primary 17A65; Secondary 17D99.

Key words and phrases. Right Lie algebra, right multiplication, Hoehnke and Baer radical.

LEMMA 1. *For any right Lie algebra L the following identities are satisfied for all $x, y, z \in L$.*

$$R_{xy} = R_x R_y + R_y L_x. \tag{2}$$

$$L_x L_y = L_{xy} + L_y L_x. \tag{3}$$

$$R_{xy} = R_y L_x - L_x R_y. \tag{4}$$

For the proof we refer to [2].

LEMMA 2. *Every right Lie algebra L satisfies the identity*

$$R_x^n = (-1)^{n-1} L_x^{n-1} R_x \quad \text{for all} \quad n = 1, 2, \dots. \tag{5}$$

PROOF. For any $y \in L$ by (1) we get

$$(y)R_x^n = (\dots((yx)x)\dots x) = (-1)((\dots(xy)x)\dots x) = \\ \dots = (-1)^{n-1}(x(\dots(x(xy))\dots))x = (-1)^{n-1} L_x^{n-1} R_x.$$

PROPOSITION 1. *$R(L)$ is an ideal in $M(L)$ for every right Lie algebra L.*

PROOF. As is well-known, $M(L)$ is an associative ring. Clearly it suffices to show that

$$R_a L_b + L_c R_d \in R(L)$$

for any $a, b, c, d \in L$. By (2) in Lemma 1 $R_a L_b \in R(L)$ is always true. Hence (4) in Lemma 1 implies also $L_c R_a \in R(L)$.

We shall say that a nonzero ideal H of an associative algebra A is perfect in A, if for any ideal I of A such that $0 \neq I \subseteq H$, $IA \neq 0$ follows, that is, the intersection of H and of the left annihilator ideal $\operatorname{Ann} A$ of A is zero. For a right Lie algebra L, let us define the ideal

$$C(L) = \{a \in L \mid ax - xa = 0, \ \forall x \in L\}.$$

Following [3] we define $T_0(L) = 0$,

$$T_{\alpha+1}(L)/T_\alpha(L) = C(L/T_\alpha(L)),$$

$$T_\gamma(L) = \bigcup_{\beta<\gamma} T_\beta(L)$$

for limit ordinals γ and

$$T(L) = \bigcup_{\alpha \geqq 0} T_\alpha(L).$$

It has been shown in [3] Corollary 1 that the assignment $T\colon L \to T(L)$ is an idempotent Hoehnke radical in the variety U of all right Lie algebras over F. Further, let us consider the class

$$\mathcal{A} = \{L \in U \mid C(L) \neq 0 \text{ and } C(L)L = 0\}$$

and the complement

$$\overline{\mathcal{A}} = U \setminus \mathcal{A}.$$

THEOREM 1. *A right Lie algebra $L \in U$ over a field F of characteristic $\neq 2$ is T-semisimple (that is, $T(L) = 0$) if and only if*

(a) $L \in \overline{\mathcal{A}}$;

(b) *the right multiplication algebra $R(L)$ is a perfect ideal in $M(L)$.*

PROOF. Suppose that $T(L) = 0$. Then, by definition, $L \in \overline{\mathcal{A}}$. Let us assume that $R(L)$ is not perfect in $M(L)$. Now $I = R(L) \cap \operatorname{Ann} M(L) \neq 0$. Hence $IL_z = IR_z = 0$, for all $z \in L$. Any nonzero element $\overline{a} \in I$ has the form

$$\overline{a} = \sum_i \alpha_i \prod_j R_{a_{ij}}$$

and so for every element $y \in L$ we have

$$y\overline{a} = \sum_i \alpha_i (ya_{i1}) \prod_{j\neq 1} a_{ij}.$$

Since $\overline{a} \neq 0$, there exists an element $y_0 \in L$ such that $0 \neq y_0\overline{a} \in L$. Since $\overline{a} \in I$ and $IL_z = IR_z = 0$, we conclude

$$z(y_0\overline{a}) = y_0\overline{a}L_z = 0$$

and

$$(y_0\overline{a})z = y_0\overline{a}R_z = 0$$

for all $z \in L$. Hence it follows $0 \neq y_0\overline{a} \in C(L)$, contradicting $C(L) \subseteqq T(L) = 0$. Thus $R(L)$ is perfect in $M(L)$.

Assume that $T(L) \neq 0$ though (a) and (b) are satisfied. Since $L \in \overline{\mathcal{A}}$, there exists an element $c \in C(L)$ such that $cL \neq 0$ and $cy = yc$ for all $y \in L$. Hence $R_c = L_c \neq 0$. Let H denote the ideal in $M(L)$ generated by L_c. Now we have $H \triangleleft R(L)$, $(yc)x = (cy)x$ and by (1) also $(yc + cy)x = 0$. Since char $F \neq 2$ we conclude $(yc)x = 0$ for all $x, y \in L$, that is,

$$L_cR_x = R_cR_x = 0 \qquad \forall x \in L.$$

By $c \in C(L) \triangleleft L$ we have $yc \in C(L)$ for all $y \in L$, and therefore also yc commutes with all $x \in L$. This implies $L_cL_x = 0$ for all $x \in L$. Thus we have got

$$0 \neq L_c \in H \cap \operatorname{Ann} M(L) \subseteqq R(L) \cap \operatorname{Ann} M(L)$$

contradicting (b).

In the sequel β will stand for the Baer (that is, prime) radical of an associative algebra.

THEOREM 2. *Let L be a right Lie algebra over a field F of characteristic $\neq 2$. If $L \in \overline{\mathcal{A}}$, then the following conditions are equivalent:*

(1) $T(L) = 0$;

(2) *$R(L)$ is a perfect ideal in $M(L)$;*

(3) $\beta(M(L))$ *is a perfect ideal in* $M(L)$.

PROOF. The implication $(2) \Rightarrow (1)$ follows from Theorem 2.

$(1) \Rightarrow (3)$. Assume that $\beta(M(L))$ is not perfect in $M(L)$. Then we have

$$I = \beta(M(L)) \cap \operatorname{Ann} M(L) \neq 0$$

and

$$IM(L) = 0.$$

Hence there exists a nonzero element $a \in I$ and to a an element $y \in L$ such that $(y)a \neq 0$. Since $IM(L) = 0$ it follows

$$((y)a)R_z = ((y)a)L_z = 0$$

and

$$((y)a)z = z((y)a) = 0 \qquad \forall z \in L.$$

This implies $(y)a \in C(L)$. By (1) we have also

$$0 \neq (y)a \in C(L) \subseteqq T(L) = 0,$$

a contradiction.

$(3) \Rightarrow (2)$. Suppose that $R(L)$ is not perfect in $M(L)$. Then

$$I = R(L) \cap \operatorname{Ann} M(L) \neq 0,$$

and also

$$I^2 \subseteqq IM(L) = 0.$$

The latter implies $I \subseteqq \beta(M(L))$. Since by (3) $\beta(M(L))$ is perfect in $M(L)$, we get

$$I \subseteqq \beta(M(L)) \cap \operatorname{Ann} M(L) = 0,$$

a contradiction.

If in Theorem 2 we drop the condition imposed on the characteristic of F, then we can state the following

PROPOSITION 2. *Let L be a right Lie algebra over a commutative ring with* 1, *which is not Lie algebra. If $T(L) = 0$, then $R(L) \cap (M(L)) \neq 0$.*

PROOF. Let $R(Z_L)$ denote the subalgebra of $R(L)$ generated by the set $\{R_a \mid a \in Z_L\}$. Since L is not a Lie algebra, $Z_L \neq 0$, though $Z_L L = 0$. Hence for every nonzero element $a \in Z_L$ we have $La \neq 0$, and consequently $R(Z_L)$ is a nonzero subalgebra of $M(L)$. Let I be the ideal of $M(L)$ generated by $R(Z_L)$. Since $Z_L L = 0$, one can verify that $I^2 = 0$. Therefore $I \subseteqq \beta(M(L))$ and by Proposition 1, we get

$$0 \neq I \subseteqq R(L) \cap \beta(M(L))$$

proving the assertion.

PROPOSITION 3. *Let L be a right Lie algebra over a commutative ring with* 1. *If $R(L)\cap\beta(M(L))=0$, then L satisfies the following three identities*

$$xz^2=0, \qquad x(yz+zy)=0,$$
$$x(y(zt)+z(ty)+t(yz))=0 \quad \textit{for all} \quad x,y,z,t\in L.$$

PROOF. If the identity $xz^2=0$ holds then by linearization we get $x(yz+ + zy)=0$, as well as $x(y(zt)+z(ty)+t(yz))=0$. If $xz^2\neq 0$ for some $x,z\in L$, then obviously also $R(Z_L)\neq 0$. As earlier, the square of the ideal of $M(L)$ generated by $R(Z_L)$ is zero, yielding

$$0\neq R(Z_L)\subseteq R(L)\cap\beta(M(L)),$$

a contradiction.

COROLLARY 1. *Let L be a right Lie algebra over a commutative ring with* 1. *If $\beta(M(L))=0=T(L)$, then L is a Lie algebra such that* $\operatorname{ad} L\cong L$.

PROOF. In virtue of Proposition 2 we have $xz^2=0$. Hence by $T(L)=0$ it follows $z^2=0$, and consequently $\operatorname{ad} L\cong L$.

Finally we give a nilpotency result concerning the multiplication algebra $M(L)$ of a right Lie algebra L.

PROPOSITION 4. *Let L be a right Lie algebra over a commutative ring with* 1 *and let $K(L)$ denote the ideal of $M(L)$ generated by all left multiplications of L. Then*

$$\overline{M}(L)=M(L)/K(L)$$

is a zero algebra (that is, $\overline{M}^2(L)=0$).

PROOF. Obvious.

ACKNOWLEDGEMENT. The author wishes to express his indebtedness and gratitude to Prof. R. Wiegandt for this invaluable comments and fruitful suggestions and discussions and technical assistance.

REFERENCES

[1] БАХТУРИН, Ю. А., Тождества в алгебрах Ли [Identities in Lie algebras], Наука, Москва, 1985 (in Russian). *MR* **86k**:17015

[2] ТУМУРБАТ, С., Об одной алгебре, *Мон. Г. У. Эрдэм шинжилгээний бичиг*, **1990**, No. 102, 164–171.

[3] TUMURBAT, S., Hoehnke radicals for right Lie algebras, *Beiträge Alg. Geom.* **33** (1992), 85–96.

[4] ZHEVLAKOV, K. A., SLIN'KO, A. M., SHESTAKOV, I. P. and SHIRSHOV, A. I., *Rings that are nearly associative*, Pure and Applied Mathematics, 104, Academic Press, New York–London, 1982. *MR* **83i**:17001. For the Russian edition see *MR* **80h**:17002

(Received April 25, 1991)

DEPARTMENT OF ALGEBRA
UNIVERSITY OF MONGOLIA
P.O.BOX 75
ULAN BATOR 20
MONGOLIA

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 369–377

ON THE DENSITY OF RATIONAL COMBINATIONS OF $\{x^{\lambda_n}\}$ FOR SOME COMPLEX SEQUENCE $\{\lambda_n\}$

S. P. ZHOU

Abstract

The present paper investigates the approximation by rational combinations of $\{x^{\lambda_n}\}$ when the complex sequence $\{\lambda_n\}$ is not too close to the imaginary axis.

1. Introduction

Let $C_{[a,b]}$ be the class of all complex continuous functions f on $[a,b]$. For $f \in C_{[0,1]}$,

$$\omega(f,t) = \max\{|f(x+h) - f(x)| : 0 < h \leqq t,\ x \in [0, 1-h]\},$$

for $f \in C_{[0,b]}$,

$$\|f\|_{[0,b]} = \max_{x \in [0,b]} |f(x)|.$$

Given a subspace S of $C_{[0,1]}$, let

$$R(S) = \left\{ \frac{P(x)}{Q(x)} : P(x) \in S, Q(x) \in S, Q(x) > 0, x \in (0,1] \right\},$$

where we assume that $P(0)/Q(0) = \lim_{x \to 0+} P(x)/Q(x)$ is finite in the case $Q(0) = 0$. For a sequence of numbers $\Lambda = \{\lambda_n\}_{n=0}^{\infty}$, write

$$R(\Lambda) = R(\mathrm{span}\{x^{\lambda_n}\}).$$

From Müntz theorem, it is well-known that the linear combinations of $\{x^{\lambda_n}\}$ for

(1) $$0 = \lambda_0 < \lambda_1 < \lambda_2 \ldots$$

1991 *Mathematics Subject Classification.* Primary 41A20; Secondary 41A30.
Key words and phrases. Rational combinations, Müntz approximation, density, complex sequence.

[1]We assume that $\lim_{x \to +\infty} f(x)$ exists if $b = +\infty$.

0081-6906/94/$ 4.00 ©1994 *Akadémiai Kiadó, Budapest*

are dense in $C_{[0,1]}$ if and only if

$$\sum_{n=1}^{\infty} \frac{1}{\lambda_n} = \infty.$$

In 1916, Szász [6] investigated the case when λ_n are distinct complex numbers. He proved that under conditions $\lambda_0 = 0$ and $\operatorname{Re}\lambda_n > 0$, the linear combinations of $\{x^{\lambda_n}\}$ are dense in $C_{[0,1]}$ if

$$\sum_{n=1}^{\infty} \frac{\operatorname{Re}\lambda_n}{1+|\lambda_n|^2} = \infty,$$

while the linear combinations of $\{x^{\lambda_n}\}$ are not dense if

$$\sum_{n=1}^{\infty} \frac{\operatorname{Re}\lambda_n + 1}{1+|\lambda_n|^2} < \infty.$$

At the same time, some works studied the case when λ_n are not too close to the imaginary axis. For example, in 1971 Luxemburg and Korevaar [3] obtained the following result:

Let $|\operatorname{Re}\lambda_n| \geqq \delta|\lambda_n|$ for some $\delta > 0$, then the linear combinations of $\{x^{\lambda_n}\}$ are dense in $C_{[a,b]}$ for $a > 0$ if and only if

$$\sum_{n=1}^{\infty} \frac{1}{|\lambda_n|} = \infty.$$

Concerning the rational case, in 1976, Somorjai [4] showed a beautiful result that under (1), $R(\Lambda)$ is always dense in $C_{[0,1]}$. In 1978, Bak and Newman [2] proved that if $\{\lambda_n\}$ is a sequence of distinct positive numbers, then $R(\Lambda)$ is dense in $C_{[0,1]}$, too. Our recent work [6] established that $R(\Lambda)$ is always dense for any sequence of real numbers $\{\lambda_n\}$ with infinitely many distinct elements.

The present paper will investigate the approximation by rational combinations of $\{x^{\lambda_n}\}$ when the complex sequence $\{\lambda_n\}$ is not too close to the imaginary axis.

2. Result and proof

Write

$$\Lambda_n = \{\lambda_1, \lambda_2, \dots, \lambda_n\},$$

$$R(\Lambda_n) = R(\operatorname{span}\{x^{\lambda_k}\} : \lambda_k \in \Lambda_n),$$

$$R_n(f,\Lambda)_{[0,b]} = \min_{r\in R(\Lambda_n)} \|f-r\|_{[0,b]}.$$

Given a sequence $\{a_n\}$, let

$$\Delta a_1 = a_1,$$
$$\Delta a_k = a_k - a_{k-1} \qquad \text{for} \quad k=2,3,\dots .$$

For $1 \leqq j \leqq 2n-1$,

$$x_j^n := x_j = \begin{cases} \frac{j}{n}, & 1\leqq j \leqq n, \\ \frac{n}{2n-j}, & n+1 \leqq j \leqq 2n-1, \end{cases}$$

$$P_j^n(x) := P_j(x) = x^{\lambda_j} \prod_{l=1}^{j} x_l^{-\Delta\lambda_l}.$$

LEMMA 1. *Let* $\alpha > 0$,

$$\operatorname{Re}\Delta\lambda_k \geqq \alpha k$$

for $k=2,3,\dots$, *and denote*

$$\Delta^* x_k = \begin{cases} 2x_2, & k=2, \\ \Delta x_k, & 3\leqq k \leqq 2n-1, \\ \infty, & k=2n, \end{cases}$$

$$k^* := k^*(x) = \begin{cases} 1, & 0 \leqq x \leqq x_2, \\ k-1, & x_k - \Delta^* x_k/2 \leqq x \leqq x_k,\ 3\leqq k \leqq 2n-1, \\ k, & x_k < x < x_k + \Delta^* x_{k+1}/2,\ 2 \leqq k \leqq 2n-1. \end{cases}$$

Then for $x_k - \Delta^* x_k/2 \leqq x < x_k + \Delta^* x_{k+1}/2$, $k=2,3,\dots,2n-1$, *and* $j \in \in \{1,2,\dots,2n-1\}\setminus\{k-1,k\}$,

$$|P_{k^*}^{-1}(x)P_j(x)| \leqq e^{-\alpha|k-j|/8}.$$

PROOF. We only prove the case when $2 \leqq k \leqq n$. The remaining case can be treated similarly. Let $x \in [x_k - \Delta^* x_k/2, x_k] = [x_k - (2n)^{-1}, x_k]$ for $k \geqq 3$. Direct calculation yields that for $1 \leqq j \leqq k-2$,

$$|P_{k-1}^{-1}(x)P_j(x)| = \Big|\prod_{l=j+1}^{k-1}\Big(\frac{l}{nx}\Big)^{\Delta\lambda_l}\Big| \leqq \prod_{l=j+1}^{k-1}\Big(\frac{2l}{2k-1}\Big)^{\alpha l},$$

by the estimate

$$(2) \qquad 1 - x \leqq e^{-x} \quad \text{for} \quad x \geqq 0,$$

it follows that for $1 \leqq l \leqq k-1$,

$$\Big(\frac{2l}{2k-1}\Big)^{l} \leqq \exp\Big(-\frac{(2k-2l-1)l}{2k-1}\Big) \leqq e^{-1/4},$$

that is,

$$(3)\qquad |P_{k-1}^{-1}(x)P_j(x)| \leqq e^{-\alpha(k-j-1)/4} \leqq e^{-\alpha(k-j)/8}.$$

Meanwhile,

$$|P_{k-1}^{-1}(x)P_j(x)| = \begin{cases} \left|\prod_{l=k}^{j}(\frac{nx}{l})^{\Delta\lambda_l}\right|, & k+1 \leqq j \leqq n, \\ \left|\prod_{l=k}^{l=n}(\frac{nx}{l})^{\Delta\lambda_l}\prod_{l=n+1}^{j}(\frac{(2n-l)x}{n})^{\Delta\lambda_l}\right|, & n+1 \leqq j \leqq 2n-1, \end{cases}$$

using (2) again, for $k+1 \leqq j \leqq n$ we get

$$|P_{k-1}^{-1}(x)P_j(x)| \leqq e^{-\alpha(1+2+\ldots+(j-k))} \leqq e^{-\alpha(j-k)^2/2},$$

for $n+1 \leqq j \leqq 2n-1$,

$$|P_{k-1}^{-1}(x)P_j(x)| \leqq e^{-\alpha(n-k)^2/2}\prod_{l=n+1}^{j}\left(\frac{2n-l}{n}\right)^{\alpha l} \leqq \exp\left(-\frac{\alpha}{2}((n-k)^2+(j-n)^2)\right),$$

altogether for $k+1 \leqq j \leqq 2n-1$,

$$(4)\qquad |P_{k-1}^{-1}(x)P_j(x)| \leqq e^{-\alpha(j-k)^2/4}.$$

Similarly, let $x \in [x_2 - \Delta^* x_2/2, x_2] = [0, x_2]$, we can prove that for $3 \leqq j \leqq$ $\leqq 2n-1$,

$$(4')\qquad |P_1^{-1}(x)P_j(x)| \leqq e^{-\alpha(j-1)^2/4}.$$

When $x_k < x < x_k + \Delta^* x_{k+1}/2$, $k = 2, 3, \ldots, 2n-1$, similar discussion will lead to that for $1 \leqq j \leqq k-2$,

$$(5)\qquad |P_k^{-1}(x)P_j(x)| \leqq e^{-\alpha(k-j-1)/2} \leqq e^{-\alpha(k-j)/4},$$

and for $k+1 \leqq j \leqq 2n-1$,

$$(6)\qquad |P_k^{-1}(x)P_j(x)| \leqq e^{-\alpha(j-k)^2/8}.$$

From the estimates (3), (4), (4′), (5) and (6), the required result of Lemma 1 follows. □

LEMMA 2. *Let* $\operatorname{Re}\Delta\lambda_n > 0$ *and*

$$(7)\qquad |\Delta\lambda_n| \leqq M \operatorname{Re}\Delta\lambda_n$$

for $n = 1, 2, \ldots$. *Then for* $x \in [0, x_k]$,

$$1 - e^{-\frac{\pi}{2M}} \leqq |1 + P_{k-1}^{-1}(x) P_k(x)| \leqq 2,$$

and for $x \in [x_k, +\infty)$,

$$1 - e^{-\frac{\pi}{2M}} \leqq |1 + P_k^{-1}(x) P_{k-1}(x)| \leqq 2.$$

PROOF. Evidently, when $x \in [0, x_k]$,

$$|P_{k-1}^{-1}(x) P_k(x)| = |x_k^{-\Delta\lambda_k} x^{\Delta\lambda_k}| \leqq 1,$$

and

$$x_k^{-i\,\mathrm{Im}\Delta\lambda_k} x^{i\,\mathrm{Im}\Delta\lambda_k} = \exp\left(i\,\mathrm{Im}\,\Delta\lambda_k \log\frac{x}{x_k}\right).$$

For $x_k e^{-\frac{\pi}{2|\Delta\lambda_k|}} \leqq x \leqq x_k$,

$$\left|\mathrm{Im}\,\Delta\lambda_k \log\frac{x}{x_k}\right| \leqq \frac{\pi}{2},$$

that is, for $x \in [x_k e^{-\frac{\pi}{2|\Delta\lambda_k|}}, x_k]$,

$$|1 + P_{k-1}^{-1}(x) P_k(x)| \geqq 1. \tag{8}$$

On the other hand, when $0 \leqq x \leqq x_k e^{-\frac{\pi}{2|\Delta\lambda_k|}}$,

$$x_k^{-\mathrm{Re}\,\Delta\lambda_k} x^{\mathrm{Re}\,\Delta\lambda_k} \leqq e^{-\frac{\pi}{2|\Delta\lambda_k|}\mathrm{Re}\,\Delta\lambda_k},$$

by (7),

$$x_k^{-\mathrm{Re}\,\Delta\lambda_k} x^{\mathrm{Re}\,\Delta\lambda_k} \leqq e^{-\frac{\pi}{2M}},$$

therefore, together with (8), for every $x \in [0, x_k]$,

$$|1 + P_{k-1}^{-1}(x) P_k(x)| \geqq 1 - e^{-\frac{\pi}{2M}}.$$

The argument for the second inequality is similar. □

We now establish our main result

THEOREM. *Let* $M > 0$ *be a given number, and* $\alpha > 0$ *be any number satisfying*

$$\sum_{j=1}^{\infty} e^{-\alpha j/8} < \frac{1}{2}(1 - e^{-\frac{\pi}{2M}}).$$

Let $\{\lambda_n\}$ *be a sequence of complex numbers which satisfies*

$$0 \leqq \mathrm{Re}\,\lambda_1 < \mathrm{Re}\,\lambda_2 < \mathrm{Re}\,\lambda_3 < \ldots,$$

for $k=1,2,\dots,$

$$\operatorname{Re}\Delta\lambda_k \geqq \alpha k,$$

and

$$|\Delta\lambda_k| \leqq M \operatorname{Re}\Delta\lambda_k.$$

Then for any $f \in C_{[0,+\infty]}$,

$$R_n(f,\Lambda)_{[0,+\infty)} \leqq C_\alpha(\omega(f,n^{-1})+\omega(g,n^{-1})),$$

where C_α *is a constant only depending upon* α, *and*

$$g(x)=f\left(\frac{1}{x}\right), \qquad x \geqq 1.$$

PROOF. Define

$$r_n(f,x)=\frac{\sum_{j=1}^{2n-1} f(x_j)P_j(x)}{\sum_{j=1}^{2n-1} P_j(x)},$$

clearly, $r_n(f,x) \in R_{2n-1}(\Lambda)$. First assume $x \in [x_k - \Delta^* x_k/2, x_k]$ for $2 \leqq k \leqq 2n-1$. Applying Lemmas 1 and 2, together with the conditions of the theorem, we have

(9)

$$\left|\sum_{j=1}^{2n-1} P_j(x)\right| = |P_{k-1}(x)|\left|1+P_{k-1}^{-1}(x)P_k(x)+\sum_{j=1,j\neq k-1,k}^{2n-1} P_{k-1}^{-1}(x)P_j(x)\right| \geqq$$

$$\geqq |P_{k-1}(x)|\left(1-e^{-\frac{\pi}{2M}}-2\sum_{j=1}^{\infty}e^{-\alpha j/8}\right) \geqq C_\alpha|P_{k-1}(x)|.$$

Write

$$f(x)-r_n(f,x)=\frac{(f(x)-f(x_{k-1}))P_{k-1}(x)+(f(x)-f(x_k))P_k(x)}{\sum_{j=1}^{2n-1} P_j(x)}+$$

$$+\frac{\sum_{j=1,j\neq k-1,k}^{2n-1}(f(x)-f(x_j))P_j(x)}{\sum_{j=1}^{2n-1} P_j(x)} := \Sigma_1+\Sigma_2.$$

Then by (9) and Lemma 2,

$$|\Sigma_1| = \frac{|(f(x_k)-f(x_{k-1}))P_{k-1}(x)+(f(x)-f(x_k))(P_{k-1}(x)+P_k(x))|}{\left|\sum_{j=1}^{2n-1} P_j(x)\right|} \leqq$$

$$\leqq \frac{3\omega(f,\Delta x_k)|P_{k-1}(x)|}{C_\alpha |P_{k-1}(x)|} \leqq 3C_\alpha^{-1}(\omega(f,n^{-1})+\omega(g,n^{-1})).$$

At the same time, from (9) and Lemma 1,

$$|\Sigma_2| \leqq \frac{|P_{k-1}(x)|\Big(\sum_{j=1}^{k-2}\omega(f,x_k-x_j)e^{-\alpha(k-j)/8}+\sum_{j=k+1}^{n}\omega(f,x_j-x_{k-1})e^{-\alpha(j-k)/8}\Big)}{C_\alpha|P_{k-1}(x)|}+$$

$$+\frac{|P_{k-1}(x)|\sum_{j=n+1}^{2n-1}\omega(g,\frac{j-k+1}{n})e^{-\alpha(j-k)/8}}{C_\alpha|P_{k-1}(x)|} \leqq$$

$$\leqq C_\alpha^{-1}(\omega(f,n^{-1})+\omega(g,n^{-1}))\sum_{j=1}^{\infty} je^{-\alpha j/8},$$

that is, there is a constant A such that

$$|f(x)-r_n(f,x)| \leqq AC_\alpha^{-1}(\omega(f,(2n)^{-1})+\omega(g,(2n)^{-1})).$$

As to the case when $x \in [x_k, x_k+\Delta^* x_{k+1}/2]$ for $2 \leqq k \leqq 2n-1$, Theorem follows in a similar way by using $P_k(x)$ instead of $P_{k-1}(x)$. The Theorem is proved. □

3. Corollaries and remarks

COROLLARY 1. *Let $M>0$ be a given number, and $\alpha>0$ be any number satisfying*

$$\sum_{j=1}^{\infty} e^{-\alpha j/8} < \frac{1}{2}(1-e^{-\frac{\pi}{2M}}).$$

Let $\{\lambda_n\}$ be a sequence of complex numbers which satisfies

$$0 \leqq \operatorname{Re}\lambda_1 < \operatorname{Re}\lambda_2 < \operatorname{Re}\lambda_3 < \dots,$$

for $k=1,2,\dots$,

$$\operatorname{Re}\Delta\lambda_k \geqq \alpha k,$$

and

$$|\Delta\lambda_k| \leqq M \operatorname{Re}\Delta\lambda_k.$$

Then for any $f \in C_{[0,1]}$,

$$R_n(f,\Lambda)_{[0,1]} \leqq C_\alpha \omega(f,n^{-1}).$$

Corollary 2. *Let $M>0$ be a given number, and $\alpha>0$ be any number satisfying*

$$\sum_{j=1}^{\infty} e^{-\alpha j/8} < \frac{1}{2}(1-e^{-\frac{\pi}{2M}}).$$

Let $\{\lambda_n\}$ be a sequence of complex numbers which satisfies

$$0 \geqq \operatorname{Re}\lambda_1 > \operatorname{Re}\lambda_2 > \operatorname{Re}\lambda_3 > \dots,$$

for $k=1,2,\dots,$

$$\operatorname{Re}\Delta\lambda_k \leqq -\alpha k,$$

and

$$|\Delta\lambda_k| \leqq M|\operatorname{Re}\Delta\lambda_k|.$$

Then for any $f\in C_{[0,1]}$,

$$R_n(f,\Lambda)_{[0,1]} \leqq C_\alpha \omega(f,n^{-1}).$$

Corollary 3. *Let $M>0$ be a given number, and $\{\lambda_n\}$ be a sequence of complex numbers which satisfies*

$$0 \leqq \operatorname{Re}\lambda_1 < \operatorname{Re}\lambda_2 > \operatorname{Re}\lambda_3 > \dots < \operatorname{Re}\lambda_n \to +\infty, \qquad n\to\infty,$$

and for $k=1,2,\dots,$

$$|\lambda_k| \leqq M\operatorname{Re}\lambda_k.$$

Then $R(\Lambda)$ is dense in $C_{[0,1]}$.

Corollary 4. *Let $M>0$ be a given number, and $\{\lambda_n\}$ be a sequence of complex numbers which satisfies*

$$0 \geqq \operatorname{Re}\lambda_1 > \operatorname{Re}\lambda_2 > \operatorname{Re}\lambda_3 > \dots > \operatorname{Re}\lambda_n \to -\infty, \qquad n\to\infty,$$

and for $k=1,2,\dots,$

$$|\lambda_k| \leqq M|\operatorname{Re}\lambda_k|.$$

Then $R(\Lambda)$ is dense in $C_{[0,1]}$.

Remark 1. If $\{\lambda_n\}$ is an increasing nonnegative sequence, we can easily see that every $P_j(x)>0$ for $x\in(0,1]$. Therefore the condition

$$\operatorname{Re}\Delta\lambda_k \geqq \alpha k$$

can be reduced to the condition

$$\Delta\lambda_k \geqq k$$

in this case. Corollary 1 then becomes the result of Bak [1].

REMARK 2. Different from the result of the polynomial case (Luxemburg and Korevaar's result), which is valid in $C_{[a,b]}$ for $a > 0$, our theorem is established for continuous function space on an interval containing the origin.

REMARK 3. We do not know whether or not the rational combinations of $\{x^{\lambda_n}\}$ are dense in $C_{[0,1]}$ for a complex sequence $\{\lambda_n\}$ whose real parts are bounded away from the infinity and which is not too close to the imaginary axis.

REFERENCES

[1] BAK, J., On the efficiency of general rational approximation, *J. Approximation Theory* **20** (1977), 46–50. *MR* **56** #6206

[2] BAK, J. and NEWMAN, D. J., Rational combinations of x^{λ_k}, $\lambda_k \geqq 0$ are always dense in $C_{[0,1]}$, *J. Approximation Theory* **23** (1978), 155–157. *MR* **58** #6840

[3] LUXEMBURG, W. A. J. and KOREVAAR, J., Entire functions and Müntz–Szász type approximation, *Trans. Amer. Math. Soc.* **157** (1971), 23–37. *MR* **43** #7643

[4] SOMORJAI, G., A Müntz-type problem for rational approximation, *Acta Math. Acad. Sci. Hungar.* **27** (1976), 197–199. *MR* **55** #3622

[5] SZÁSZ, O., Über die Approximation stetiger Funktionen durch lineare Aggregate von Potenzen, *Math. Ann.* **77** (1916), 482–496. *Jahrbuch Fortschritte Math.* **46**, 419.

[6] ZHOU, S. P., On Müntz rational approximation, *Constr. Approx.* **9** (1993), 435–444.

(Received April 26, 1991)

DEPARTMENT OF MATHEMATICS, STATISTICS
AND COMPUTING SCIENCE
DALHOUSIE UNIVERSITY
HALIFAX, NOVA SCOTIA
B3H 3J5
CANADA

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 379–386

DREHFLÄCHEN ZWEITER ORDNUNG DURCH EINEN KEGELSCHNITT

O. RÖSCHEL

In [6] und [7] hat H. Schaal unlängst alle Drehzylinder untersucht, die vier allgemeine Punkte des dreidimensionalen euklidischen Raumes E_3 enthalten. Dabei liegt sein Hauptaugenmerk auf der von den Drehachsen gebildeten Regelfläche. Dadurch angeregt hat U. Strobl in [8] die Fläche der Spitzen aller Drehkegel durch vier gegebene Punkte betrachtet. Es sind dies Untersuchungen, die im ,,ebenen Fall" — die gegebenen vier Punkte gehören einer festen Ebene an — eng mit den sogenannten *Fokalkegelschnitten* [4, S. 104] verbunden sind. Es erscheint daher lohnenswert, die einen festen Kegelschnitt k enthaltenden Drehflächen zweiter Ordnung zu studieren. Als Spezialfall stellen sich dabei Drehkegel ein, deren Scheitel dem Fokalkegelschnitt von k angehören.[1] Es gelingt zu zeigen, daß die Mittelpunkte der k enthaltenden Drehflächen zweiter Ordnung mit festem Kehl- bzw. Äquatorkreisradius auf zwei Kegelschnitten s_κ und t_λ liegen, die zu den Fokalkegelschnitten s_0 sowie t_0 von k homothetisch sind. Wenn k eine Parabel ist, tritt als entsprechender Mittelpunktsort eine Parabel s_κ auf, die zur Fokalparabel von k schiebungsgleich ist. Die Achsen der betreffenden Drehquadriken sind dann Tangenten von s_κ bzw. t_λ.

1. Im projektiv abgeschlossenen, komplex erweiterten, reellen dreidimensionalen euklidischen Raum E_3 beschreiben wir Punkte in Bezug auf ein kartesisches Normalkoordinatensystem (x, y, z). In der Normalform wird eine *Ellipse* k dann etwa durch

$$(1) \qquad k\colon\ \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \qquad z = 0$$

beschrieben, wobei $a^2 - b^2 > 0$ gewählt werden können.[2] Mit der üblichen Abkürzung $e^2 := a^2 - b^2$ erhalten die beiden reellen *Brennpunkte* $F_{1,2}$ von k die Koordinaten $(\pm e, 0, 0)$.

1991 *Mathematics Subject Classification.* Primary 51N20.
Key words and phrases. Quadrics of revolution.

[1]Die Achsen der Drehflächen zweiter Ordnung, die fünf allgemeine Punkte des Raumes enthalten, wurden bereits von E. Laguerre [3] betrachtet, wobei jedoch der hier studierte ebene Fall — die fünf Punkte allgemeiner Lage definieren den Kegelschnitt k — nicht untersucht wurde. Dies deshalb, weil dann die Drehachsen der betreffenden Drehquadriken in den Symmetrieebenen von k liegen.

[2]Wir werden im folgenden Kreise von unseren Betrachtungen ausschließen.

Die Frage nach dem *Ort der Scheitel aller k enthaltenden Drehkegel* ist eng mit dem Begriff der *Dandelinschen Kugeln* (vgl. [1, S. 171 f.]) verbunden und kann elementar beantwortet werden (vgl. [4, S. 104]). Die entsprechenden Scheitel sind an die sogenannte *Fokalhyperbel* s_0 mit der Gleichung

$$s_0: \quad \frac{x^2}{e^2} - \frac{z^2}{b^2} = 1, \qquad y = 0 \tag{2}$$

bzw. den *nullteiligen Fokalkegelschnitt* t_0 mit der Gleichung

$$t_0: \quad x = 0, \qquad \frac{y^2}{e^2} + \frac{z^2}{a^2} + 1 = 0 \tag{3}$$

gebunden. Die Punkte von t_0 führen allerdings auf komplexe Drehkegel durch k. s_0 besitzt die Hauptscheitel von k als Brennpunkte, die Brennpunkte von k als Scheitel.

Vollkommen gleichartige Verhältnisse herrschen, wenn k eine *Hyperbel* ist. In unserer Darstellung (1) haben wir bloß b^2 durch $-b^2$ zu ersetzen und erhalten entsprechend die *Fokalellipse* (2) der Hyperbel k.

2. Wir fragen nun allgemeiner nach allen *Drehflächen zweiter Ordnung, die den Kegelschnitt k (1) enthalten.* Ihre Drehachsen d gehören aus Symmetriegründen entweder der $[x, z]$-Ebene $y = 0$ oder der $[y, z]$-Ebene $x = 0$ an. Wir studieren vorerst den Fall, daß die Ebene $y = 0$ auch Symmetrieebene der k enthaltenden Drehflächen zweiter Ordnung ist: Da der Schnitt mit der Ebene $z = 0$ unser Ausgangskegelschnitt k sein soll, dürfen wir diese Quadriken in der Gestalt

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 + z[2a_{03} + 2a_{13}x + a_{33}z] \tag{4}$$

mit noch zu bestimmenden reellen Konstanten a_{03}, a_{13}, a_{33} ansetzen. Eine Drehfläche zweiter Ordnung liegt in (4) genau dann vor, wenn die beiden durch

$$\begin{aligned} \frac{x^2}{a^2} + \frac{y^2}{b^2} &= z[2a_{13}x + a_{33}z] \\ x^2 + y^2 + z^2 &= 0 \end{aligned} \tag{5}$$

erfaßten *reellen Kreisschnittstellungen* von (4) zusammenfallen. Nach Elimination von y aus (5) entsteht eine homogene quadratische Gleichung in $x : z$, die genau für

$$a_{33} = \frac{a^2 b^4 a_{13}^2 - e^2}{e^2 b^2} \tag{6}$$

eine Doppellösung besitzt. Die zugehörigen Kreisschnittebenen sind parallel zur Ebene

$$e^2 x + a^2 b^2 a_{13} z = 0. \tag{7}$$

Die *Drehachse* der dann durch die Gleichung

$$\text{(8)} \qquad \Phi(a_{03}, a_{13}): \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 + z\left[2a_{03} + 2a_{13}x + \frac{a^2b^4a_{13}^2 - e^2}{e^2b^2}z\right]$$

beschriebenen *Drehflächen zweiter Ordnung durch* k besitzen die *Richtungsvektoren*

$$\text{(9)} \qquad \mathbf{d} = (e^2, 0, a^2b^2a_{13})^t.$$

Die *Mittelpunkte dieser Drehquadriken* besitzen die Koordinaten

$$\text{(10)} \qquad M(a_{03}, a_{13}): \quad \left(-\frac{a_{03}a_{13}a^2b^2e^2}{a^4b^2a_{13}^2 - e^2}, 0, -\frac{a_{03}b^2e^2}{a^4b^2a_{13}^2 - e^2}\right) := (m_x, 0, m_z).$$

Sie sind genau für die k enthaltenden *Drehparaboloide* nicht definiert, die sich für

$$\text{(11)} \qquad a^4b^2a_{13}^2 - e^2 = 0$$

einstellen. Unsere Drehquadriken (8) sind genau dann *Drehkegel*, wenn $b^2e^2a_{03}^2 = a^4b^2a_{13}^2 - e^2$ gilt. Die Mittelpunkte (10) werden zu den Scheiteln, die dem Fokalkegelschnitt s_0 (2) von k angehören. Die Drehachsen dieser Drehkegel sind in Übereinstimmung mit [4, S. 104] Tangenten von s_0.

Bei festem a_{13}, aber variablem $a_{03} \in \Re$ besitzen die Drehquadriken (8) feste Achsenrichtung; ihre Mitten liegen auf einer Geraden durch U: $(0,0,0)$. Da sie überdies wegen der festen Achsenrichtung dieselbe Fernkurve besitzen, liegt ein *Büchsel von Drehflächen zweiter Ordnung vor.*

Wenn die Drehquadriken durch k die Symmetrieebene $x = 0$ besitzen, führt eine analoge analytische Betrachtung auf die durch

$$\text{(12)} \qquad \Phi(a_{03}, a_{23}): \quad \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 + z\left[2a_{03} + 2a_{23}y - \frac{a^4b^2a_{23}^2 + e^2}{e^2a^2}z\right]$$

beschriebene zweiparametrige Quadrikenschar. Die zugehörigen Drehachsen besitzen Richtungsvektoren

$$\text{(13)} \qquad \mathbf{d} = (0, -e^2, a^2b^2a_{23})^t,$$

während die Mittelpunkte durch

$$\text{(14)} \qquad M(a_{03}, a_{23}): \quad \left(0, \frac{a_{03}a_{23}a^2b^2e^2}{a^2b^4a_{23}^2 + e^2}, \frac{a_{03}a^2e^2}{a^2b^4a_{23}^2 + e^2}\right)$$

erfaßt werden.

Unter Beachtung, daß die vorliegenden Ergebnisse nach Ersetzen von b^2 durch $-b^2$ auch für eine Hyperbel als Ausgangskegelschnitt k gelten, haben wir den

Satz 1. *Alle Drehflächen zweiter Ordnung des E_3, die den in der Normalform* (1) *vorliegenden Mittelpunktskegelschnitt k enthalten, gehören zwei zweiparametrigen Scharen an und werden durch die Darstellung* (8) *bzw.* (12) *erfaßt. Die zugehörigen Mittelpunkte besitzen Koordinaten* (10) *bzw.* (14), *während die Richtungsvektoren der Drehachsen durch* (9) *bzw.* (13) *beschrieben werden.*

3. Die Mittelpunkte $M(a_{03}, a_{13})$ unserer Drehquadriken $\Phi(a_{03}, a_{13})$ (8) und die zugehörigen Drehachsen (9) definieren in der Ebene $y=0$ ein *Richtungsfeld R* mit der charakteristischen Differentialgleichung $b^2 x dx = e^2 z dz$. Aus den Beziehungen (9) und (10) lesen wir ab, daß R bei *zentrischen Ähnlichkeiten mit dem Mittelpunkt U*: $(0,0,0)$ *von k als Fixpunkt in sich übergeht.* Die Integralkurven dieses Richtungsfeldes sind daher entweder Geraden durch U oder werden von diesen zentrischen Ähnlichkeiten bloß vertauscht. Die Integralkurven besitzen die Gleichung

$$s_\kappa: \quad \frac{x^2}{e^2} - \frac{z^2}{b^2} = 1 - \kappa, \quad y = 0, \tag{15}$$

wobei κ eine reelle Integrationskonstante bezeichnet. Für $\kappa = 0$ stellt sich der *reelle Fokalkegelschnitt* s_0 (2) ein ([4, S. 104]). Der zu s_0 *konjugierte Kegelschnitt* ist $s_{\kappa=2}$. Analoges gilt für die k enthaltenden Drehquadriken $\Phi(a_{03}, a_{23})$ (12) mit der Symmetrieebene $x=0$. Hier erhalten wir ein Richtungsfeld S, dessen Integralkurven durch die Gleichung

$$t_\lambda: \quad x = 0, \quad \frac{y^2}{e^2} + \frac{z^2}{a^2} + 1 - \lambda = 0 \tag{16}$$

mit der Integrationskonstante $\lambda \in \Re$ erfaßt werden. Für $\lambda = 0$ und $\lambda = 2$ stellen sich der nullteilige Fokalkegelschnitt t_0 und der dazu konjugierte (einteilige) t_2 ein.

Überraschend läßt sich nun der folgende Satz beweisen:

Satz 2. *Jene einen Mittelpunktskegelschnitt k enthaltenden Drehflächen zweiter Ordnung, deren Mittelpunkte einer Integralkurve s_κ (t_λ) des Richtungsfeldes R (S) angehören, besitzen festen — nur von κ (λ) abhängigen — Kehl- bzw. Äquatorkreisradius ϱ.*

Beweis. Der *Kehl- bzw. Äquatorkreis* der allgemeinen Drehflächen $\Phi(a_{03}, a_{13})$ liegt in der *Polarebene δ des Fernpunktes der zugehörigen Drehachsen.* Mit (9) erhalten wir für δ die Gleichung

$$\delta: \quad \frac{x}{a^2}[a^4 b^2 a_{13}^2 - e^2] + \frac{b^2 a_{13} z}{e^2}[a^4 b^2 a_{13}^2 - e^2] + a^2 b^2 a_{03} a_{13} = 0. \tag{17}$$

Die zusätzlich in der Ebene $z=0$ gelegenen Punkte des Kehl- bzw. Äquatorkreises besitzen daher die x-Koordinate

$$k_x = -\frac{a^4 b^2 a_{03} a_{13}}{a^4 b^2 a_{13}^2 - e^2}, \tag{18}$$

womit die y-Koordinate k_y aus $k_y^2 = \frac{b^2}{a^2}(1 - k_x^2)$ berechnet werden kann. Der Kehl- bzw. Äquatorkreisradius $\varrho(a_{03}, a_{13})$ berechnet sich daher über $\varrho^2(a_{03}, a_{13}) = (k_x - m_x)^2 + k_y^2 + m_z^2$ unter Beachtung von (10) zu

$$\varrho(a_{03}, a_{13}) = b^2 - \frac{b^4 e^2 a_{03}^2}{a^4 b^2 a_{13}^2 - e^2}. \tag{19}$$

Der Punkt $M(a_{03}, a_{13})$ (10) liegt genau dann auf s_κ (15), wenn

$$a_{03}^2 = \frac{1-\kappa}{b^2 e^2}[a^4 b^2 a_{13}^2 - e^2] \tag{20}$$

gilt. Dadurch ist eine einparametrige Schar von Drehquadriken $\Phi(a_{13})$ in (8) ausgezeichnet. Genau für diese ist nach Vergleich von (19) und (20) der Kehl- bzw. Äquatorkreisradius $\varrho(a_{03}, a_{13})$ eine feste Konstante mit $\varrho^2 = \kappa b^2$.

Für den Fall, daß die k enthaltenden Drehflächen zweiter Ordnung durch $\Phi(a_{02}, a_{23})$ (12) erfaßt werden, berechnet sich der Kehl- bzw. Äquatorkreisradius $\varrho(a_{03}, a_{23})$ aus der Gleichung $\varrho^2 = \lambda a^2$, wenn der Mittelpunkt der Quadrik auf der Integralkurve s_λ gewählt war, q.e.d.

Zu festem Kehl- bzw. Äquatorkreisradius ϱ_0 gehören demnach die Kegelschnitte s_κ und t_λ mit $\kappa = \frac{\varrho_0^2}{a^2}$. Überraschend schneiden sich diese beiden Kegelschnitte i.a. auf der Schnittgeraden der Symmetrieebenen (z-Achse) nicht.[3]

4. Analoge Überlegungen lassen sich auch dann durchführen, wenn der Ausgangskegelschnitt k eine *Parabel* ist. In der Normalform setzen wir

$$k: \quad y^2 = 2px, \quad z = 0, \quad p = \text{konst.} \in \Re \tag{21}$$

und treffen wegen der Symmetrie zur Ebene $y = 0$ für die k enthaltenden Drehflächen zweiter Ordnung den Ansatz

$$y^2 + z[2a_{03} + 2a_{13}x + a_{33}z] = 2px \tag{22}$$

mit noch zu bestimmenden reellen Konstanten a_{03}, a_{13} und a_{33}. Analog zur Gleichung (5) und (6) liegen in (22) genau dann *Drehflächen* vor, wenn

$$a_{33} = 1 - a_{13}^2 \tag{23}$$

[3] Dies ist für eine Ausgangsellipse k auch synthetisch leicht einzusehen: Wir wählen einen Punkt M: $(0, 0, m_z)$ der z-Achse als Mittelpunkt entsprechender Drehquadriken Φ_1 und Φ_2. Die zugehörigen Drehachsen sind entweder parallel zur Haupt- bzw. Nebenachse von k. Im ersten Fall erhalten wir als Äquatorkreisradius den Wert $\varrho_1 = \sqrt{m_z^2 + b^2}$, im zweiten jedoch $\varrho_2 = \sqrt{m_z^2 + a^2}$. Die beiden stimmen für $a^2 - b^2 > 0$ nicht überein.

gilt. Die wieder zweiparametrige Schar von Drehflächen zweiter Ordnung durch die Parabel k wird daher durch die Gleichung

$$\Psi(a_{03}, a_{13}): \quad y^2 + z[2a_{03} + 2a_{13}x + (1 - a_{13}^2)z] = 2px \tag{24}$$

erfaßt. Die Richtungsvektoren der Drehachsen berechnen sich zu $\mathbf{d} = (1, 0, -a_{13})^t$, die Koordinaten der Mittelpunkte $M(a_{03}, a_{13})$ zu

$$(m_x, 0, m_y) := \left(\frac{p(a_{13}^2 - 1) - a_{03}a_{13}}{a_{13}^2}, 0, \frac{p}{a_{13}}\right). \tag{25}$$

Dabei ist darauf zu achten, daß $a_{13} \neq 0$ ist. Bei $a_{13} = 0$ beschreibt (24) die k *enthaltenden Drehparaboloide.*

Wieder bestimmt k zusammen mit den Drehachsen ein Richtungsfeld R, das diesmal die Schiebungen parallel zur x-Achse gestattet. Eine der Integralkurven ist wieder bekannt (vgl. [4, S. 105]): Es ist dies die *Fokalparabel* s_0 von k, die als Ort der Scheitel der k enthaltenden Drehkegel auftritt. Sie besitzt die Gleichung

$$s_0: \quad y = 0, \quad z^2 = p^2 - 2px. \tag{26}$$

Die übrigen Integralkurven s_κ von R gehen aus s_0 durch x-parallele Schiebung hervor und lassen sich daher durch

$$s_\kappa: \quad y = 0, \quad z^2 = p^2 - 2p(x - \kappa) \tag{27}$$

beschreiben. Die in (24) erfaßten Drehquadriken durch k, deren Mittelpunkte auf s_κ (27) mit festgehaltenem $\kappa \in \Re$ liegen, sind durch

$$2a_{03}a_{13} = p(a_{13}^2 - 1) - 2\kappa a_{13}^2 \tag{28}$$

gekennzeichnet.

Wieder interessieren wir uns für den *Kehlkreisradius* ϱ der so ausgezeichneten Drehflächen zweiter Ordnung. Nach kurzer — Abschnitt 3 analoger — Rechnung erhalten wir die elegante Beziehung

$$\varrho^2 = 2\kappa p = \text{konstant}. \tag{29}$$

Damit gilt der

SATZ 3. *Die zweiparametrige Schar von Drehflächen zweiter Ordnung* $\Psi(a_{03}, a_{13})$, *die die gegebene Parabel* k (21) *enthalten, wird durch* (24) *erfaßt. Jene mit festem Kehlkreisradius* ϱ *besitzen Mittelpunkte auf einer Parabel* s_κ (27), *die zur Fokalparabel* s_0 *von* k *schiebungsgleich ist. Die Drehachsen der zugehörigen Drehflächen zweiter Ordnung berühren* s_κ.

Abbildung 1 zeigt die Situation im eben beschriebenen Fall in einem Kreuzriß. Für zwei Mitten M_1, M_2 auf s_κ sind die Meridianhyperbeln m_1, m_2 der k enthaltenden Drehhyperboloide Ψ_1, Ψ_2 angedeutet. Die Richtdrehkegel Δ_1^*, Δ_2^* dieser Drehhyperboloide sind parallel zu entsprechenden k enthaltenden Drehkegeln Δ_1, Δ_2.

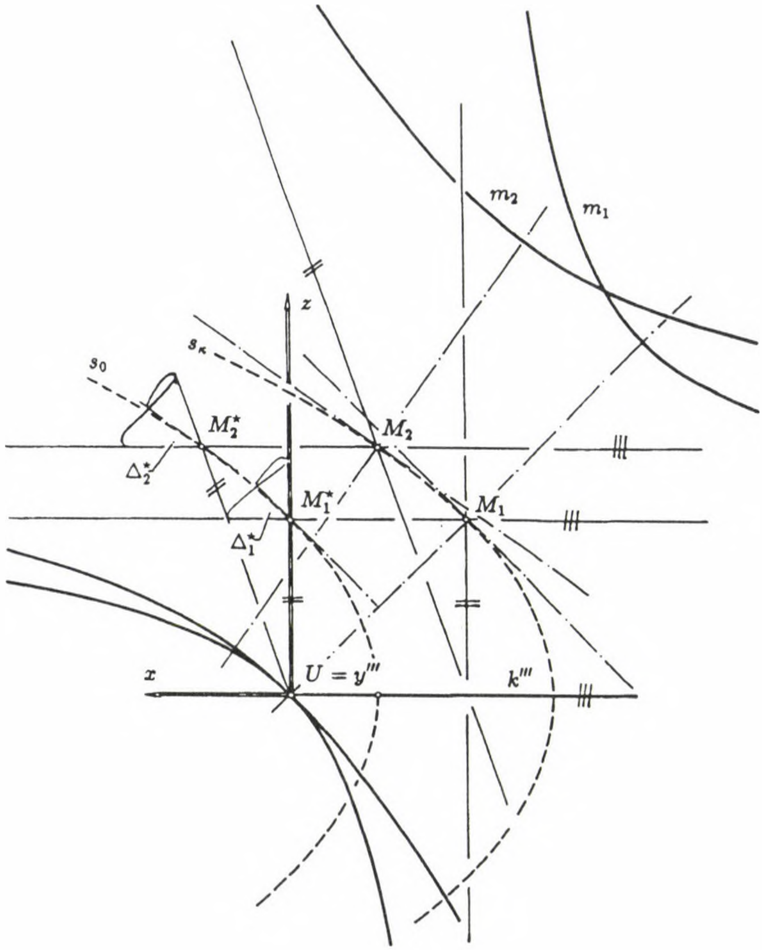


Abb. 1

LITERATURVERZEICHNIS

[1] Brauner, H., *Lehrbuch der konstruktiven Geometrie*, Springer-Verlag, Wien–New York, 1986. *MR* **87h**:51001

[2] Hoffman, L., Über ein bei den Cliffordschen Flächen bestehendes Analogon des Satzes von Dandelin, *Monatsh. Math.* **62** (1958), 1–15. *MR* **20** #4810

[3] Laguerre, E., Sur la courbe enveloppée par les axes des coniques qui passent par quatre points donnés et sur les axes des surfaces de révolution du second ordre qui passent par cinq points donnés. Sur les lignes spiriques, *Nouv. Ann. Math.* (2) **18** (1879), 206–218. *Jb. Fortschritte Math.* **11**, 403

[4] Müller, E. und Kruppa, E., *Lehrbuch der darstellenden Geometrie*, Springer, Wien, 1961.

[5] Pottmann, H., Über Scheitel von Normalrissen einer Raumkurve, *Österreich. Akad. Wiss. Math.-Natur. Kl. Sitzungsber. II* **196** (1987), 39–48. *MR* **89c**:53007

[6] SCHAAL, H., Ein geometrisches Problem der metrischen Getriebesynthese, *Österreich. Akad. Wiss. Math.-Natur. Kl. Sitzungsber. II* **194** (1985), 39–53. *MR* **87j**: 53015

[7] SCHAAL, H., Konstruktion der Drehzylinder durch vier Punkte einer Ebene, *Österreich. Akad. Wiss. Math.-Natur. Kl. Sitzungsber. II* **195** (1986), 405–418. *MR* **88g**:53020

[8] STROBL, U., Die Drehkegel durch vier Punkte, Dissertation, Stuttgart, 1990.

Weitere Literaturhinweise finden sich in [6] und [7].

(Eingegangen am 16. Dezember, 1990)

INSTITUT FÜR GEOMETRIE
TU GRAZ
KOPERNIKUSGASSE 24
A-8010 GRAZ
AUSTRIA

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 387–395

BOUNDEDNESS AND CONTINUITY FOR BILINEAR OPERATORS

P. ANTOSIK and C. SWARTZ

1. Introduction

In [2] and [5] we showed that the Antosik–Mikusiński Diagonal Theorem can be used to treat boundedness and equicontinuity properties of bilinear mappings between topological vector spaces. In [4] the Antosik–Mikusiński Diagonal Theorem was extended to topological groups and employed to extend the Uniform Boundedness Principle (UBP) to a class of topological vector spaces, called $\mathcal{A}$-spaces, which seem to be a very natural class of spaces for which UBP holds. In this paper we show that the diagonal theorem and the class of $\mathcal{A}$-spaces also have applications to bilinear operators between topological vector spaces. Some of our results extend and clarify results on bilinear operators which are contained in § 6 of the monograph [2].

In Section 2 we consider the various types of boundedness properties which are satisfied by families of separately continuous bilinear operators between topological vector spaces; we show that the general UBP established in [2] can be used to establish various boundedness relationships for such families of bilinear operators. In Section 3 we consider the various types of equicontinuity conditions that can be satisfied by families of bilinear operators. The diagonal theorem, being a result about infinite matrices, can be used to treat various types of sequential equicontinuity for bilinear operators; these types of sequential equicontinuity are treated by employing the diagonal theorem in Section 4. Finally, in Section 5 we compare the various types of boundedness and equicontinuity properties.

Our notation and terminology are standard and can be found, for example, in [3].

Throughout we let X, Y, Z be topological vector spaces and $\mathcal{F}$ a family of separately continuous bilinear operators from $X \times Y$ into Z. If $b : X \times Y \to Z$ is a bilinear map, we write $b(x,\cdot)(b(\cdot,y))$ for the linear operator from Y into Z (X into Z) defined by

$$b(x,\cdot)(y) = b(x,y)(b(\cdot,y)(x) = b(x,y))$$

for each $x \in X$ $(y \in Y)$.

1991 *Mathematics Subject Classification.* Primary 47A05.

Key words and phrases. Bilinear operators, bounded, continuous, equicontinuous.

2. Boundedness

In this section we consider the various properties of boundedness which can be satisfied by the family $\mathcal{F}$.

Let $\mathcal{M}(\mathcal{N})$ be a family of bounded subsets of $X(Y)$. We have the following notions of boundedness for the family $\mathcal{F}$ of bilinear operators.

(B1) $\mathcal{F}$ is pointwise bounded, i.e., $\{B(x,y) : B \in \mathcal{F}\}$ is bounded for each $x \in X$, $y \in Y$.

(B2) $\mathcal{F}$ is $\mathcal{M}$-uniformly bounded, i.e., $\{B(x,y) : B \in \mathcal{F}, x \in M\}$ is bounded for each $M \in \mathcal{M}, y \in Y$. [$\mathcal{N}$-uniformly bounded is defined similarly.]

(B3) $\mathcal{F}$ is $(\mathcal{M},\mathcal{N})$-uniformly bounded if $\mathcal{F}$ is both $\mathcal{M}$- and $\mathcal{N}$-uniformly bounded.

(B4) $\mathcal{F}$ is $\mathcal{M} \times \mathcal{N}$-uniformly bounded if $\{B(x,y) : B \in \mathcal{F}, x \in M, y \in N\}$ is bounded for each $M \in \mathcal{M}$, $N \in \mathcal{N}$.

The most important case is when $\mathcal{M}(\mathcal{N})$ is the family of all bounded subsets of $X(Y)$. We denote the family of all bounded subsets of X by $\mathcal{B}(X)$, and we consider the relationships between the types of boundedness above for this case.

It is clear that if $\cup\mathcal{M} = X$ and $\cup\mathcal{N} = Y$ (B4) implies (B3) implies (B2) implies (B1). We give examples showing the reverse implications are in general false when $\mathcal{M}$ and $\mathcal{N}$ are the families of bounded subsets of X and Y.

Example 1. Let c_{00} be the space of all real-valued sequences which are zero eventually and equip c_{00} with the sup-norm. Further, let e_k be the sequence which has 1 in the k^{th} coordinate and 0 elsewhere. Define $b_n : l^\infty \times c_{00} \to \mathbf{R}$ by $b_n(x,y) = \sum_{k=1}^{n} x_k y_k$ where $x = \{x_i\}$, $y = \{y_i\}$. If $m > n$, $b_n\left(e, \sum_{k=1}^{m} e_k\right) = n$, where $e = (1,1,\ldots)$, so $\{b_n\}$ satisfies (B1) but not (B2) for $\mathcal{N} = \mathcal{B}(c_{00})$. Also note that $\{b_n\}$ is $\mathcal{B}(l^\infty)$-uniformly bounded but not $\mathcal{B}(c_{00})$-uniformly bounded so $\{b_n\}$ satisfies (B2) but not (B3).

Example 2. Define $b_n : c_{00} \times c_{00} \to \mathbf{R}$ by $b_n(x,y) = \sum_{k=1}^{n} x_k y_k$. Then $\{b_n\}$ is $\mathcal{B}(c_{00})$-uniformly bounded (i.e., satisfies (B3)) but does not satisfy (B4) since $b_n\left(\sum_{k=1}^{n} e_k, \sum_{k=1}^{n} e_k\right) = n$ for each n.

The examples above show that without some additional assumptions the implications (B1) implies (B2) implies (B3) implies (B4) do not hold for the class of all bounded subsets. As is the case for the UBP, we show that the classes of $\mathcal{K}$-convergent sequences and $\mathcal{K}$-bounded sets can be used to obtain general results on boundedness of bilinear operators which hold without any completeness or barrelledness assumptions.

Recall that a sequence $\{x_k\}$ in a topological vector space X is $\mathcal{K}$-*convergent* if every subsequence of $\{x_k\}$ has a further subsequence $\{x_{n_k}\}$ such that the subseries $\sum x_{n_k}$ is convergent to an element of X ([2]). A $\mathcal{K}$-convergent sequence obviously converges to 0, but the converse is false in general although it does hold in complete metric linear spaces (see [2] § 3). A subset $A \subseteq X$ is said to be $\mathcal{K}$-*bounded* if whenever $\{x_k\} \subseteq A$ and $\{t_k\}$ is a scalar sequence converging to 0, the sequence $\{t_k x_k\}$ is $\mathcal{K}$-convergent ([1], [2]). A $\mathcal{K}$-bounded set is bounded but in general the converse does not hold ([2] § 3).

Let $\mathcal{KB}(X)(\mathcal{KS}(X))$ denote the $\mathcal{K}$-bounded subsets of X ($\mathcal{K}$-convergent sequences in X).

THEOREM 3. *Assume $\mathcal{F}$ is pointwise bounded. Then*
(i) *$\mathcal{F}$ is $\mathcal{KS}(X) \times \mathcal{KS}(Y)$-uniformly bounded;*
(ii) *$\mathcal{F}$ is $\mathcal{KB}(X) \times \mathcal{KS}(Y)$-uniformly bounded;*
(iii) *$\mathcal{F}$ is $\mathcal{KB}(X) \times \mathcal{KB}(Y)$-uniformly bounded.*

PROOF. (i) Let $\{x_k\}, \{y_k\}$ be $\mathcal{K}$-convergent sequences in X and Y, respectively. Fix $y \in Y$. Then $\mathcal{F}_y = \{b(\cdot, y) : b \in \mathcal{F}\}$ is pointwise bounded in $L(X, Z)$ so $\{b(x_k, y) : b \in \mathcal{F}, k \in \mathbb{N}\}$ is bounded by the General UBP, Theorem 2(1) of [4]. That is, $\{b(x_k, \cdot) : b \in \mathcal{F}, k \in \mathbb{N}\}$ is pointwise bounded in $L(Y, Z)$. Again, by Theorem 2(1) of [4], $\{b(x_k, y_j) : b \in \mathcal{F}, k, j \in \mathbb{N}\}$ is bounded and (i) holds.

(ii) and (iii) follow in a similar fashion using Theorem 2(2) of [4].

The class of $\mathcal{A}$-spaces was introduced in [3] and shown to be a natural class of spaces for which the UBP holds. (Recall X is an $\mathcal{A}$-space if every bounded subset of X is $\mathcal{K}$-bounded; examples of $\mathcal{A}$-spaces are given in [4].) For $\mathcal{A}$-spaces, we obtain from Theorem 3 the important

COROLLARY 4. *If X is an $\mathcal{A}$-space and $\mathcal{F}$ is pointwise bounded, then $\mathcal{F}$ is $\mathcal{B}(X)$-uniformly bounded.* [That is (B1) implies (B2).] *If X and Y are $\mathcal{A}$-spaces, then $\mathcal{F}$ is $\mathcal{B}(X) \times \mathcal{B}(Y)$-uniformly bounded.* [That is, (B1) implies (B4) if $\mathcal{M} = \mathcal{B}(X)$ and $\mathcal{N} = \mathcal{B}(Y)$.]

From the proof of Theorem 3 and the General UBP of [4], we also have

THEOREM 5. *If $\mathcal{F}$ is $\mathcal{B}(X)$-uniformly bounded, then $\mathcal{F}$ is $\mathcal{B}(X) \times \mathcal{KB}(Y)$ and $\mathcal{B}(X) \times \mathcal{KS}(Y)$-uniformly bounded.*

Again, for $\mathcal{A}$-spaces, we obtain

COROLLARY 6. *Let Y be an $\mathcal{A}$-space. If $\mathcal{F}$ is $\mathcal{B}(X)$-uniformly bounded, then $\mathcal{F}$ is $\mathcal{B}(X) \times \mathcal{B}(Y)$-uniformly bounded.* [That is, (B2) implies (B4) if $\mathcal{M} = \mathcal{B}(X)$.]

3. Equicontinuity

In this section we consider the various equicontinuity properties that $\mathcal{F}$ can satisfy and the relationship which hold between these properties. The definitions and examples are used mostly to motivate and contrast the situations with sequential equicontinuity which are considered in Section 4.

We consider the following types of continuity:

(E1) $\mathcal{F}$ is left equicontinuous, i.e., $\{b(\cdot,y): b\in\mathcal{F}\}\subseteqq L(X,Z)$ is equicontinuous $\forall y\in Y$. [Right equicontinuity is defined similarly.]

(E2) $\mathcal{F}$ is $\mathcal{N}$-equihypocontinuous if for each $N\in\mathcal{N}$ the family $\{b(\cdot,y): b\in \in\mathcal{F}, y\in N\}$ is equicontinuous in $L(X,Z)$.

(E3) $\mathcal{F}$ is $(\mathcal{M},\mathcal{N})$-equihypocontinuous if $\mathcal{F}$ is both $\mathcal{M}$- and $\mathcal{N}$-equihypocontinuous.

(E4) $\mathcal{F}$ is equicontinuous.

It is clear that if $\mathcal{M}=\mathcal{B}(X)$ and $\mathcal{N}=\mathcal{B}(Y)$, then (E4) implies (E3) implies (E2) implies (E1). We give examples to show that the reverse implications do not hold in general when $\mathcal{M}$ and $\mathcal{N}$ are the families of all bounded subsets of X and Y, respectively.

EXAMPLE 7. That (E1) does not imply (E2) follows from Example 1 since if $y=(y_1,\dots,y_m,0,\dots)\in c_{00}$, then $b_n(\cdot,y)=(y_1,\dots,y_m,0,\dots)\in(l^\infty)'$ for $n\geqq m$. Thus, $\|b_n(\cdot,y)\|=\sum_{k=1}^{m}|y_k|$ and $\{b_n(\cdot,y)\}$ is equicontinuous, i.e., $\{b_n\}$ is left equicontinuous. If $y_n=\sum_{k=1}^{n}e_k$, then $b_n(\cdot,y_n)=y_n$ so $\|b_n(\cdot,y)\|_1= =n$ and $\{b_n(\cdot,y_n):n\}$ is not equicontinuous. Hence, $\{b_n\}$ is not $\mathcal{B}(c_{00})$-equihypocontinuous. Note also that $\{b_n\}$ is not right equicontinuous since $\|b_n(e,\cdot)\|_1=n$ for each n.

EXAMPLE 8. To show that (E2) does not imply (E3) we give an example of a single separately continuous bilinear functional which is $\mathcal{B}(Y)$-hypocontinuous but is not $\mathcal{B}(X)$-hypocontinuous.

Let E be a Hausdorff locally convex space and consider the bilinear pairing $\langle\ ,\ \rangle$ between E and E'_b where E'_b is E' with the strong topology. This bilinear functional is always $\mathcal{B}(E'_b)$-hypocontinuous; for if $A\subseteqq E$ is bounded, A^0, the polar of A, is a strong neighbourhood of 0. Therefore if $\varepsilon>0$, $|\langle\,\varepsilon A^0,A\rangle|\leqq\varepsilon$.

However, we claim that $\langle\ ,\ \rangle$ is $\mathcal{B}(E)$-hypocontinuous if and only if E is quasi-barrelled. First, assume that E is quasi-barrelled. Let $B\subseteqq E'_b$ be bounded. Then B is equicontinuous since E is quasi-barrelled. Therefore, B^0 is a neighbourhood of 0 in E and if $\varepsilon>0$, $|\langle B,\varepsilon B^0\rangle|\leqq\varepsilon$ and $\langle\ ,\ \rangle$ is $\mathcal{B}(E)$-hypocontinuous.

Conversely, assume that $\langle\ ,\ \rangle$ is $\mathcal{B}(E)$-hypocontinuous. Let $B\subseteqq E'_b$ be bounded. Since $\langle\ ,\ \rangle$ is $\mathcal{B}(E)$-hypocontinuous, there is a neighbourhood of $0, U$, in E such that $|\langle B,U\rangle|\leqq 1$. Then $B\subseteqq U^0$ so B is equicontinuous, and E is quasi-barrelled.

Thus, for any locally convex space E which is not quasi-barrelled, $\langle\ ,\ \rangle$ furnishes an example of a bilinear map satisfying (E2) but not (E3).

EXAMPLE 9. To show that (E3) does not imply (E4), we give an example of a single separately continuous bilinear functional which is both $\mathcal{B}(X)$ and $\mathcal{B}(Y)$-hypocontinuous but not continuous.

We use the bilinear functional, $\langle\ ,\ \rangle$, of Example 8. We claim that $\langle\ ,\ \rangle$ is continuous if and only if E is normable. If E is normable, then $\langle\ ,\ \rangle$ is clearly continuous. Conversely, assume that $\langle\ ,\ \rangle$ is continuous. Then there exist a convex neighbourhood U of 0 in E and a closed, bounded, absolutely convex set B in E such that $|\langle B^0, U\rangle| \leqq 1$. Thus, $U \subseteqq (B^0)^0 = B$ by the Bipolar Theorem. Hence, U is a bounded convex neighbourhood of 0, and E is normable by Kolmogorov's Theorem.

Thus, if E is a non-normable, quasi-barrelled space, $\langle\ ,\ \rangle$ is hypocontinuous but not continuous.

4. Sequential equicontinuity

In this section we consider the various forms of sequential equicontinuity that can be satisfied by the family $\mathcal{F}$. These definitions are modelled on the corresponding equicontinuity conditions given in Section 3.

(S1) $\mathcal{F}$ is sequentially left equicontinuous, i.e., if $x_i \to 0$ in X, then $\lim b(x_i, y) = 0$ uniformly for $b \in \mathcal{F}$ for each $y \in Y$.

(S2) $\mathcal{F}$ is sequentially $\mathcal{N}$-equihypocontinuous, i.e., if for each $N \in \mathcal{N}$ when $x_i \to 0$ in X, then $\lim b(x_i, y) = 0$ uniformly for $b \in \mathcal{F}, y \in N$.

(S3) $\mathcal{F}$ is sequentially $(\mathcal{M}, \mathcal{N})$-equihypocontinuous if $\mathcal{F}$ is both $\mathcal{M}$- and $\mathcal{N}$-equihypocontinuous.

(S4) $\mathcal{F}$ is sequentially equicontinuous, i.e., $x_i \to 0$ in X $y_i \to 0$ in Y implies $\lim b(x_i, y_i) = 0$ uniformly for $b \in \mathcal{F}$.

Of course, the condition (Ei) implies the corresponding sequential condition (Si).

The situaton with sequential equicontinuity is much different than with (topological) equicontinuity. We have the following obvious implications which should be contrasted with those in Section 3.

THEOREM 10. (S3) *implies* (S2) *provided* $X = \cup\mathcal{M}$; *if* $\cup\mathcal{N} = Y$, *then* (S2) *implies* (S1); *if each convergent sequence in* Y *is contained in an element of* $\mathcal{N}$, *then* (S2) *implies* (S4) *(this holds, for example, if* $\mathcal{N} = \mathcal{B}(Y)$*)*.

As in Section 3, (S1) does not imply (S2) when $\mathcal{N} = \mathcal{B}(Y)$.

EXAMPLE 11. Let $b_n : c_{00} \times c_{00} \to \mathbf{R}$ be given by $b_n(x, y) = \sum_{k=1}^{n} x_k y_k$. If $y = (y_1, \ldots, y_N, 0, \ldots)$, then $|b_n(x, y)| \leqq \|x\|_\infty \sum_{i=1}^{N} |y_i|$ so $\{b_i\}$ is left equicontinuous. However, $\{b_i\}$ is not $\mathcal{B}(c_{00})$-equihypocontinuous since if $x_n = \sum_{k=1}^{n} e_k$, then $(1/\sqrt{n})x_n \to 0$, $\{x_n\}$ is bounded, but $b_n((1/n)x_n, x_n) = \sqrt{n}$.

In marked contrast to the situation in Section 3, we have

EXAMPLE 12. To show that (S4) does not imply (S2) when $\mathcal{N}$ is the family of bounded sets, let $b : l^\infty \times l^1 \to \mathbf{R}$ be defined by $b(x,y) = \sum_{i=1}^{\infty} x_i y_i$ if $x = (x_i)$ and $y = (y_i)$. Equip l^1 with the weak topology, $\sigma(l^1, l^\infty)$, and l^∞ with the weak* topology, $\sigma(l^\infty, l^1)$. If $y^i \to 0$ weakly in l^1, then $y^i \to 0$ in $\| \|_1$ by Schur's Lemma ([2] 8.2), and if $x^i \to 0$ in l^∞, then $\{x^i\}$ is $\| \|_\infty$ bounded so $b(x^i, y^i) \to 0$. However, $e_i \to 0$ weak* in l^∞ and $\{e_i\}$ is bounded in l^1 but $b(e_i, e_i) = 1$ so (S2) fails.

Finally, we have

EXAMPLE 13. To show that (S1) does not imply (S4), consider any separately continuous bilinear map on a metric linear space which is not continuous. For example, $b : c_{00} \times c_{00} \to \mathbf{R}$ defined by $b(x,y) = \sum_{i=1}^{\infty} x_i y_i$ where c_{00} has the sup-norm.

As was the case in boundedness, the families of $\mathcal{K}$-convergent sequences and $\mathcal{K}$-bounded sets can be used to obtain general results which hold without any completeness or barrelledness assumptions. From the proof of Theorem 6.17 of [2] and the general form of the Antosik–Mikusiński Diagonal Theorem from [4], we have the following general result.

THEOREM 14. *If $\mathcal{F}$ is sequentially left equicontinuous, then $\mathcal{F}$ is sequentially $\mathcal{KS}(Y)$ and $\mathcal{KB}(Y)$-equihypocontinuous.*

A sequence $\{x_k\}$ in X which converges to 0 is said to be *Mackey-convergent* or *M-convergent* if there is a sequence of scalars $\{t_i\}$ such that $t_i \to \infty$ and $t_i x_i \to 0$. The space X is called an *M-space* if every sequence which converges to 0 is M-convergent. Any metric linear space is an M-space ([3] 28.1(5)). For M-spaces, we have, from the proof of Theorem 6.17 of [2],

THEOREM 15. *Let X be an M-space. If $\mathcal{F}$ is sequentially left equicontinuous, then $\mathcal{F}$ is sequentially $\mathcal{KS}(Y)$- and $\mathcal{KB}(Y)$-equihypocontinuous.*

Of course, if X is metrizable, then sequential continuity can be replaced by continuity in both Theorems 14 and 15. Further, if Y is an $\mathcal{A}$-space, then $\mathcal{KB}(Y)$-equihypocontinuity can be replaced by $\mathcal{B}(Y)$-equihypocontinuity in Theorem 15.

5. Boundedness and equicontinuity

In this section we compare the notions of boundedness and equicontinuity which were introduced earlier. First we consider left equicontinuity.

THEOREM 16. *If $\mathcal{F}$ is sequentially left equicontinuous, then $\mathcal{F}$ is $\mathcal{B}(X)$-uniformly bounded.* [That is, (S1) implies (B2).]

PROOF. Let $y \in Y$ and $A \subseteqq X$ be bounded. Let $\{x_k\} \subseteqq A$, $\{b_k\} \subseteqq \mathcal{F}$ and $t_k \to 0$. Since $t_k x_k \to 0$, we have $b_k(t_k x_k, y) = t_k b_k(x_k, y) \to 0$ so $\{b(x,y) : x \in \in A, b \in \mathcal{F}\}$ is bounded.

REMARK 17. This result gives an improvement of 6.14 of [2]. The sequence in Example 11 also shows that Theorem 16 cannot be improved to $\mathcal{F}$ is $\mathcal{B}(X) \times \mathcal{B}(Y)$ uniformly bounded.

The following example shows that the converse of Theorem 16 does not hold.

EXAMPLE 18. Define $b_n : l^1 \times l^\infty \to \mathbf{R}$ by $b_n(x,y) = \sum_{k=1}^{n} x_k y_k$. Equip l^1 with the weak* topology, $\sigma(l^1, c_0)$, and l^∞ with the sup norm topology. Then each b_n is continuous and $\{b_n\}$ is uniformly bounded on the products of bounded sets so, in particular, $\{b_n\}$ is $\mathcal{B}(l^\infty)$-uniformly bounded. However, $\{b_n\}$ is not sequentially left equicontinuous since $e_k \to 0$ weak* in l^1 while $b_n(e_n, e) = 1$ for each n.

We do have a partial converse for Theorem 16.

THEOREM 19. *Let X be an $\mathcal{M}$-space. If $\mathcal{F}$ is $\mathcal{B}(X)$-uniformly bounded, then $\mathcal{F}$ is sequentially left equicontinuous.*

PROOF. Let $y \in Y$ and $x_k \to 0$ in X. Pick $t_k \to \infty$ such that $t_k x_k \to 0$. If $\{b_k\} \subseteqq \mathcal{F}$, then $\{b_k(t_k x_k, y)\}$ is bounded since $\{t_k x_k\}$ is bounded so $(1/t_k) b_k(t_k x_k, y) = b_k(x_k, y) \to 0$, and $\mathcal{F}$ is sequentially left equicontinuous.

Similarly,

THEOREM 20. *Let X be quasibarrelled. If $\mathcal{F}$ is $\mathcal{B}(X)$-uniformly bounded, then $\mathcal{F}$ is left equicontinuous.*

PROOF. For $y \in Y$, $\mathcal{F}_y = \{b(\cdot, y) : b \in \mathcal{F}\}$ is uniformly bounded on bounded subsets of X so $\mathcal{F}_y$ is equicontinuous by the quasibarrelled assumption (Proposition 11 of [4]).

Corollary 4 gives conditions when a pointwise bounded family is $\mathcal{B}(X)$-uniformly bounded so Theorems 19 and 20 are applicable.

THEOREM 21. *If $\mathcal{F}$ is sequentially $\mathcal{B}(Y)$-equihypocontinuous, then $\mathcal{F}$ is $\mathcal{B}(X) \times \mathcal{B}(Y)$ uniformly bounded.*

PROOF. Let $A \subseteqq X$ and $B \subseteqq Y$ be bounded. Let $\{x_k\} \subseteqq A$, $\{y_k\} \subseteqq \subseteqq B$, $\{b_k\} \subseteqq \mathcal{F}$ and $t_k \to 0$. Since $t_k x_k \to 0$, $b_k(t_k x_k, y_k) = t_k b_k(x_k, y_k) \to 0$ so $\{b(x,y) : b \in \mathcal{F}, x \in A, y \in B\}$ is bounded.

The converse is false.

EXAMPLE 22. We give an example of a single separately continuous bilinear map which is uniformly bounded on products of bounded sets but not sequentially $\mathcal{B}(Y)$-hypocontinuous. Define $b : l^\infty \times l^1 \to \mathbf{R}$ by $b(x,y) = \sum_{k=1}^{\infty} x_k y_k$. Equip l^∞ with the weak* topology, $\sigma(l^\infty, l^1)$, and l^1 with $\| \|_1$. Then b is uniformly bounded on products of bounded sets but $e_k \to 0$ weak* in l^∞ and $\{e_k\}$ is bounded in l^1 with $b(e_k, e_k) = 1$ so b is not $\mathcal{B}(Y)$-hypocontinuous.

As a partial converse, we have

THEOREM 23. *Let X be an M-space. If $\mathcal{F}$ is $\mathcal{B}(X) \times \mathcal{B}(Y)$ uniformly bounded, then $\mathcal{F}$ is sequentially $\mathcal{B}(Y)$-equihypocontinuous.*

PROOF. Let $x_k \to 0, \{b_k\} \subseteqq \mathcal{F}$ and $\{y_k\}$ be bounded in Y. Pick $t_k \to \infty$ such that $t_k x_k \to 0$. Then $\{t_k x_k\}$ is bounded so $\{b_k(t_k x_k, y_k)\}$ is bounded and $(1/t_k) b_k(t_k x_k, y_k) = b_k(x_k, y_k) \to 0$ as desired.

Similarly,

THEOREM 24. *Let X be quasibarrelled. If $\mathcal{F}$ is $\mathcal{B}(X) \times \mathcal{B}(Y)$ uniformly bounded, then $\mathcal{F}$ is $\mathcal{B}(Y)$-equihypocontinuous.*

PROOF. Let $B \subseteqq Y$ be bounded. Then $\{b(\cdot, y) : b \in \mathcal{F}, y \in B\}$ is uniformly bounded on bounded subsets of X and is, therefore, equicontinuous by Proposition 11 of [4].

Corollaries 4 and 6 give conditions under which a family is $\mathcal{B}(X) \times \mathcal{B}(Y)$-uniformly bounded and Theorems 23 and 24 are applicable.

Finally we can combine some of the results above to obtain a version of the Banach–Steinhaus Theorem for sequences of bilinear operators (see 6.12 of [2]).

THEOREM 25. *Let $b_i : X \times Y \to Z$ be bilinear and separately continuous and suppose that $\lim_i b_i(x,y) = b(x,y)$ exists for each $x \in X, y \in Y$. If X is an M-space and both X and Y are $\mathcal{A}$-spaces, then $\{b_i\}$ is sequentially equicontinuous and b is sequentially continuous.*

PROOF. Let $x_j \to 0$ in X and $y_j \to 0$ in Y. Since $\{b_i\}$ is pointwise bounded by Corollaries 4 and 6, $\{b_i\}$ is $\mathcal{B}(X) \times \mathcal{B}(Y)$-bounded and by Theorem 23, $\{b_i\}$ is sequentially equicontinuous. Hence, $\lim_j b(x_j, y_j) = \lim_j \lim_i b_i(x_j, y_j) = \lim_i \lim_j b_i(x_j, y_j) = 0$, and the result follows.

REMARK 26. This generalizes 6.12 of [2].

REFERENCES

[1] ANTOSIK, P., On uniform boundedness of families of mappings, *Proceedings of the Conference on Convergence Structures* (Szczyrk, 1979), Polsk. Akad. Nauk, Oddzial Katowicah, Katowice, 1980, 1–16. *MR* **83i**:46006a

[2] ANTOSIK, P. and SWARTZ, C., *Matrix methods in analysis*, Lecture Notes in Mathematics, 1113, Springer-Verlag, Berlin–New York, 1985. *MR* **87b**:46079

[3] KÖTHE, G., *Topological vector spaces.* I, Die Grundlehren der mathematischen Wissenschaften, Band 159, Springer-Verlag, N. Y., 1969. *MR* **40** #1750

[4] LI, R. and SWARTZ, C., Spaces for which the uniform boundedness principle holds, *Studia Sci. Math. Hungar.* **27** (1992), 379–384.

[5] SWARTZ, C., Continuity and hypocontinuity for bilinear maps, *Math. Z.* **186** (1984), 321–329. *MR* **85h**:46006

(Received May 28, 1991)

DEPARTMENT OF MATHEMATICAL SCIENCES
BOX 30001, DEPT. MB
LAS CRUCES, NM 88003–0001
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 397–413

ON RANDOM WALKS WITH BARRIERS AND THEIR APPLICATION TO QUEUES

W. BÖHM and S. G. MOHANTY

Abstract

The n-step transition probabilities of a random walk with two barriers, each being either reflecting or absorbing are considered on the basis of a simple renewal argument. The relation of these walks to queuing problems is pointed out and the distributions of the queue length in the finite capacity case, the same during a busy period and of the maximum queue length are derived for discrete time models. By taking the limit the solutions of continuous time models are derived, verifying some known results.

1. Introduction

In this paper we will be concerned with the one-dimensional random walk $Z(n)$ having state space the set of all integers and one-step transition probabilities

$$\begin{aligned} P(Z(n)=k+1|Z(n-1)=k) &= \alpha \\ P(Z(n)=k-1|Z(n-1)=k) &= \gamma \\ P(Z(n)=k|Z(n-1)=k) &= \beta, \end{aligned}$$

where $\beta = 1-\alpha-\gamma$. It is well known that there is an intimate connection between the random walk $Z(n)$ and the discrete time queuing process $Q(n)$, in particular the following relation holds:

$$Q(n) = Z(n) - \min_{0\leqq i\leqq n} Z(i) \wedge 0,$$

which states that $Q(n)$ is derived from $Z(n)$ by introducing a reflecting barrier at zero.

Suppose now that $Z(n)$ is restricted by two barriers, one at zero and the other located at $h>0$, where each one may be either reflecting or absorbing. Here four possible cases arise and each has its own interpretation in a queuing context. In the simplest case there are two absorbing barriers, one at zero and the other one located at h. The transitions of $Z(n)$ between these barriers

1991 *Mathematics Subject Classification*. Primary 60J15, 60K25; Secondary 05A15.

Key words and phrases. Transient solution, maximum queue length, busy period, discrete time queues, finite capacity queues, random walk, absorbing and reflecting barriers, generating function.

coincide with the transitions of the process $Q(n)$ during a period where the service facility is continuously busy and the maximum queue length is less than h. As the second case we will consider a reflecting barrier at zero and an absorbing barrier at h. This time $Z(n)$ is equivalent to a queuing process with maximum queue length less than h and such that the server need not be continuously busy. In this context we will also discuss the distribution of the stopping time τ_{mh}, the time the process $Q(n)$ requires to reach a maximum value of h for the first time, provided that there are m customers initially waiting. In the third case there will be an absorbing barrier at zero and a reflecting barrier at h. Now $Z(n)$ corresponds to a queuing process $Q(n)$, which is continuously busy up to time n and has a waiting room with finite capacity h. In the remaining fourth case there will be two reflecting barriers, located at zero and at h. This arrangement gives rise to the distribution of the queue length of a queuing process with finite waiting room capacity h, the server need not be continuously busy.

At a first glance it seems that the introduction of various barriers is a purely formal procedure, changing only the boundary conditions of the system of partial difference equations, which has to be solved in order to find the distribution of $Q(n)$. However, the presence of barriers may be given an interesting probabilistic interpretation, which provides us with additional insight into the structure of the process $Q(n)$ and with a simple method of solving all four possible cases mentioned above.

To give this interpretation and to outline the main argument which we will use in subsequent sections, the case where $Z(n)$ is restricted by one reflecting barrier at zero, thus giving rise to the simple queuing process $Q(n)$, will be discussed shortly.

Suppose that $Z(0) = m > 0$, then it is apparent that the transitions of $Z(n)$ and $Q(n)$ coincide as long as $Z(n)$ remains positive. At the time when $Z(n)$ touches the barrier for the first time, a busy period of $Q(n)$ terminates and therefore the stopping time

$$T_{m0} = \inf\{n : Z(n) = 0 \mid Z(0) = m\}$$

equals the duration of a busy period initiated by m customers. The end of a busy period marks the beginning of an idle period which will terminate with the arrival of the next customer. Let A_1 denote the length of this idle period. It equals the number of trials up to the first success in a sequence of Bernoulli experiments having success probability α and hence has a geometric distribution. With the arrival of the next customer the random walk $Z(n)$ starts anew and continues up to the time when the server becomes free again. Thus the queuing process $Q(n)$ exhibits a typical repetitive pattern. It is regenerative with renewal points those time instants, where the queue becomes empty. Any two consecutive renewal points include exactly one idle period.

Let us now assume that there are $i \geq 0$ completed idle periods and consider the event $\{Q(n)=k \mid Q(0)=m\}$. This event has probability

$$P(Z(n)=k,\, T_{m0}>n \mid Z(0)=m) \quad \text{if } \; i=0,$$

and

$$P(T_{m0}=n) * [P(A_1=n) * P(T_{10}=n)]^{(i-1)*} *$$
$$* P(A_1=n) * P(Z(n)=k, T_{10}>n \mid Z(0)=1) \quad \text{if } \; i>0,$$

where "$*$" denotes convolution and $(\;)^{i*}$ means i-fold convolution. Noting that by the Markov property of $Z(n)$ we have

$$P(T_{a0}=n) * P(T_{b0}=n) = P(T_{a+b,0}=n)$$

and

$$P(A_1=n)^{i*} = P(A_i=n),$$

where A_i is the number of trials up to the i-th success in a sequence of Bernoulli experiments, we obtain for $k>0$ after summing on i:

$$P(Q(n)=k \mid Q(0)=m) = P(Z(n)=k, T_{m0}>n \mid Z(0)=m) +$$

$$+ P(Z(n)=k, T_{10}>n \mid Z(0)=1) * \sum_{i\geq 0} P(A_{i+1}=n) * P(T_{m+i,0}=n).$$

Since A_i has a negative binomial distribution and the distribution of T_{m0} is given by

$$P(T_{m0}=n) = \gamma P(Z(n-1)=1, T_{m0}>n-1 \mid Z(0)=m),$$

the transient distribution of $Q(n)$ is primarily determined by the zero-avoiding transition probabilities

$$P(Z(n)=k, T_{m0}>n \mid Z(0)=m). \tag{1}$$

The key point in the above analysis is that it takes into account the structure of the process by splitting the sample paths into repetitive patterns at renewal points and possibly identifying a distribution on which the determination of the transient distribution depends.

In this paper we will follow the same approach for two-barrier cases. Clearly the zero-avoiding transition probabilities will be replaced by $\{0,h\}$-avoiding transition probabilities, which are needed in each of the four situations.

In Karlin and McGregor [3] this type of random walks has been treated through the method of spectral decomposition. This method may be used to determine the transient solution of general birth-death processes. It exhibits its full power if the transition probability matrices are of infinite dimension which gives rise to continuous spectra. In the finite dimensional case it reduces to the determination of eigenvalues and eigenvectors of a tridiagonal matrix, which is not always the simplest way to find a solution. This method works from the top to the bottom in the sense that first the one-step transition matrix is set up, then its eigenvalues, eigenvectors and their norms are determined leading to the final evaluation of transition probabilities. The results so obtained for one type of barriers cannot be utilized for another type for which the whole procedure has to be repeated right from the beginning. In this sense it is unable to exploit the structural properties of the processes.

In contrast our method proceeds from the bottom to the top, as explained hereafter. First we look for appropriate renewal points, and then we decompose the sample paths according to those points into subsegments whose generating functions are known from the two-absorbing-barrier case. Our method does not only give the solutions to the various cases, but also utilizes the structure common to two-barrier situations. Furthermore it is simple and unifying in the sense that only one generating function is needed, from which all results are derived in a straightforward manner. Its use is not restricted to the discussion of simple random walks, but also applies to more general Markov processes. It may further be noted that our technique resembles a powerful principle from enumerative combinatorics: in order to count a set of configurations one usually looks for a decomposition into simpler subconfigurations. This gives rise to a factorization of the generating function carrying the enumerative information into simpler functions, which may be known (Goulden and Jackson [2], pp. 34).

In Mohanty and Panny [7] an elegant geometric-combinatoric method is suggested for the random walk with one reflecting barrier at zero. It is unfortunate that the same method cannot be adapted for two-barrier cases, if at least one is reflecting. In another paper (Mohanty and Panny [8]) for the same one-barrier problem, they started with the generating function of the random walk having a reflecting barrier at zero and an absorbing barrier at h. However, they neither used the structure of the process nor gave the solution to the two-barrier problem.

The paper is organized as follows: in the next section the basic generating function is given and used to derive the transient solutions of the discrete time queuing processes described above. In the last section we will show how these results translate to continuous time random walks.

Usually solutions for continuous time models are considered and are used as approximations for discrete time models. What we have achieved in this paper is to provide exact transient solutions for discrete time models, and by taking the limit, the solutions of continuous time models are derived, verifying some known results.

2. The results in discrete time

2.1. The basic generating function. Let $p_n(h,m,k) = P(Z(n) = k, 0 < Z(i) < h, i = 0, \dots, n \mid Z(0) = m)$ and define the pgf.

$$G_{mk}(s) = \sum_{n \geqq 0} s^n p_n(h,m,k).$$

Then it can be shown that (Panny [11])

$$(2)\qquad \begin{aligned} G_{mk}(s) &= \frac{(\alpha v^2+\beta v+\gamma)(\varrho v)^{k-m}(1-(\varrho v^2)^{h-k})(1-(\varrho v^2)^m)}{\gamma(1-\varrho v^2)(1-(\varrho v^2)^h)} && (k \geqq m) \\ &= \frac{(\alpha v^2+\beta v+\gamma)v^{m-k}(1-(\varrho v^2)^{h-m})(1-(\varrho v^2)^k)}{\gamma(1-\varrho v^2)(1-(\varrho v^2)^h)} && (k \leqq m), \end{aligned}$$

where the substitution $s = (\alpha v + \beta + \gamma/v)^{-1}$ has been used and $\varrho = \alpha/\gamma$. Two useful pgf's may be derived readily from (2): define the stopping times

$$\begin{aligned} T_{m0} &= \inf\{n : Z(n) = 0 \mid Z(0) = m, Z(i) < h, i = 0, \dots, n\} \\ T_{mh} &= \inf\{n : Z(n) = h \mid Z(0) = m, Z(i) > 0, i = 0, \dots, n\}, \end{aligned}$$

with $H_{m0}(s) = \sum_{n \geqq 0} s^n P(T_{m0} = n)$ and $H_{mh}(s) = \sum_{n \geqq 0} s^n P(T_{mh} = n)$. Then we have

$$(3)\qquad H_{m0}(s) = s\gamma G_{m1}(s) = \frac{v^m(1-(\varrho v^2)^{h-m})}{1-(\varrho v^2)^h},$$

and

$$(4)\qquad H_{mh}(s) = s\gamma G_{m,h-1}(s) = \frac{(\varrho v)^{h-m}(1-(\varrho v^2)^m)}{1-(\varrho v^2)^h}.$$

For brevity we omit the arguments in the pgf's, for example we write G_{mk} for $G_{mk}(s)$.

2.2. Absorbing barriers at 0 and h. This case has been dealt with several times in the literature, e.g. by Kemperman [6] for random walks with only two types of steps. We note that the following identity holds:

$$\begin{aligned} &\{Z(n) = k, 0 < Z(i) < h, 0 \leqq i \leqq n \mid Z(0) = m\} \equiv \\ &\equiv \{Q(n) = k, 0 < Q(i) < h, 0 \leqq i \leqq n \mid Q(0) = m\} \equiv \\ &\equiv \{Q(n) = k, \max_{0 \leqq i \leqq n} Q(i) < h, Q(i) > 0, 0 \leqq i \leqq n \mid Q(0) = m\}. \end{aligned}$$

Thus

$$P(Q(n) = k, \max_{0 \leqq i \leqq n} Q(i) < h, Q(i) > 0, 0 \leqq i \leqq n \mid Q(0) = m) = p_n(h,m,k).$$

Expanding (2) into partial fractions and using Theorem 1 (pp. 9–10) in Panny [11], an explicit expression for $p_n(h,m,k)$ may be given as follows:
(5)

$$p_n(h,m,k)=\frac{2}{h}\varrho^{\frac{k-m}{2}}\sum_{\nu=1}^{h-1}\sin\frac{(h-m)\nu\pi}{h}\sin\frac{(h-k)\nu\pi}{h}\Big(\beta+2\sqrt{\alpha\gamma}\cos\frac{\nu\pi}{h}\Big)^n.$$

It will be both convenient and instructive to have an alternative representation of $p_n(h,m,k)$ in terms of generalized trinomial coefficients, which are defined by

$$\binom{n;\alpha,\beta,\gamma}{k}=[v^k](\alpha v^2+\beta v+\gamma)^n \tag{6}$$

and are, as we shall see later, the direct discrete time analogues of the modified Bessel functions. So in terms of generalized trinomial coefficients we have
(7)

$$p_n(h,m,k)=\sum_{\nu=-\infty}^{\infty}\varrho^{h\nu}\left[\binom{n;\alpha,\beta,\gamma}{n+k-m-2\nu h}-\varrho^{h-m}\binom{n;\alpha,\beta,\gamma}{n+k+m-2h(\nu+1)}\right],$$

which may be derived from (2) by applying Cauchy's integral theorem. This formula has a striking combinatorial flavour, since it can be derived by repeatedly applying the reflection principle to the sample paths of $Z(n)$. However, expression (7) cannot be derived directly by the method of spectral decomposition.

2.3. Absorbing barrier at h and reflecting barrier at 0. As we remarked in the introduction, in this case the n-step transition probabilities of the random walk $Z(n)$ equal

$$P(Q(n)=k,\max_{0\leqq i\leqq n}Q(i)<h\mid Q(0)=m). \tag{8}$$

Let us exploit the renewal properties of the process $Q(n)$. We note that a convenient set of renewal points are those time instants where the queue becomes empty. From the discussion in the introduction it is clear that the lengths of idle periods have a geometric distribution with success probability α. Let $W(s)$ be the corresponding pgf. Then upon using the substitution $s=(\alpha v+\beta+\gamma/v)^{-1}$ we have

$$W(s)=\frac{\alpha s}{1-(1-\alpha)s}=\frac{\varrho v}{\varrho v^2-v+1}.$$

Suppose now m and k are positive and let $S(h,m,k)$ be the pgf. of (8). Then

$$S(h,m,k)=\sum_{i\geqq 0}S_i(h,m,k),$$

where $S_i(h,m,k)$ is the pgf. of (8), given that there are i completed idle periods. By using the renewal properties, we get

(9) $$S_i(h,m,k) = H_{m0}W^i H_{10}^{i-1} G_{1k} \qquad (i \geqq 1)$$

and

(10) $$S_0(h,m,k) = G_{mk}.$$

Hence we find

(11) $$S(h,m,k) = G_{mk} + H_{m0}G_{1k}W \sum_{i\geqq 0} W^i H_{10}^i = G_{mk} + \frac{H_{m0}G_{1k}W}{1-H_{10}W}.$$

Similar reasoning shows that

(12) $$\begin{aligned} S(h,m,0) &= \frac{H_{m0}W}{\alpha s(1-H_{10}W)} \\ S(h,0,k) &= \frac{G_{1k}W}{1-H_{10}W} \\ S(h,0,0) &= \frac{W}{\alpha s(1-H_{10}W)}. \end{aligned}$$

The generating functions occurring in the formulas above are all known (see (2), (3)). Inserting their expressions we obtain for $m \leqq k$:

(13)
$$S(h,m,k) = \frac{(\varrho v)^{k-m}(\alpha v^2+\beta v+\gamma)[1-(\varrho v^2)^{h-k}][1-v+v(1-\varrho v)(\varrho v^2)^m]}{\gamma(1-\varrho v^2)[1-v+v(1-\varrho v)(\varrho v^2)^h]},$$

and for $m \geqq k$:

(14) $$S(h,m,k) = \frac{v^{m-k}(\alpha v^2+\beta v+\gamma)[1-(\varrho v^2)^{h-m}][1-v+v(1-\varrho v)(\varrho v^2)^k]}{\gamma(1-\varrho v^2)[1-v+v(1-\varrho v)(\varrho v^2)^h]}.$$

It is easily checked that (13) and (14) cover also the cases where m and (or) k are equal to zero. It is possible to expand (13) and (14) into partial fractions, however, this requires the knowledge of the roots of the polynomial equation

$$1 - v + \varrho^h v^{2h+1} - \varrho^{h+1} v^{2h+2} = 0.$$

Unfortunately the roots (and hence the eigenvalues of the transition matrix), except for the trivial case $\varrho = 1$, cannot be given in closed form. However, for completeness, we will sketch how a partial fraction expansion may be obtained. Using the substitution $v = \varrho^{-1/2}e^{\theta i}$, we find for $m \geqq k$:

(15)
$$S(h,m,k) = \frac{\varrho^{\frac{m-k-1}{2}}(\beta + 2\sqrt{\alpha\gamma}\cos\theta)\sin(h-k)\theta[\sin(m+1)\theta - \varrho^{-1/2}\sin m\theta]}{\gamma\sin\theta[\sin(h+1)\theta - \varrho^{-1/2}\sin h\theta]}.$$

The roots θ_ν of the denominator have to be determined numerically (one of the angles θ_ν will be complex if $\varrho^{1/2} < h/(h+1)$) and after expansion into partial fractions:

$$P(Q(n)=k, \max_{0\leqq i\leqq n} Q(i) < h \mid Q(0)=m) =$$

(16)

$$= -2\varrho^{\frac{k-m}{2}} \sum_{\nu=1}^{h} \frac{(\beta+2\sqrt{\alpha\gamma}\cos\theta_\nu)^n \sin(h-k)\theta_\nu[\sqrt{\alpha}\sin(m+1)\theta_\nu - \sqrt{\gamma}\sin m\theta_\nu]}{\sqrt{\alpha}(h+1)\cos(h+1)\theta_\nu - \sqrt{\gamma}h\cos h\theta_\nu},$$

which holds also for $m \leqq k$.

Root finding may be avoided if we expand the denominator of (13) directly. For this purpose we note that the function $(1-v+v(1-\varrho v)(\varrho v^2)^h)^{-1}$ is analytic in a disc around the origin. Therefore it has an expansion of the form

$$(1-v+v(1-\varrho v)(\varrho v^2)^h)^{-1} = \sum_{a\geqq 0} d_{a,h} v^a \qquad (|v| \leqq 1),$$

where an application of the binomial theorem yields

$$d_{a,h} = (-\varrho)^a \sum_{i\leqq a} \sum_{j\geqq 0} \binom{i}{j} \binom{j}{a-i-2hj} (-1)^{i+j} \varrho^{-i-hj}. \tag{17}$$

Let us now expand (14) by means of Cauchy's integral theorem. We obtain:

$$P(Q(n)=k, \max_{0\leqq i\leqq n} Q(i) < h \mid Q(0)=m) = \tag{18}$$

$$= \frac{1}{2\pi i} \oint \frac{v^{m-k}}{v^{n+1}} \frac{(\alpha v^2+\beta v+\gamma)^n [1-(\varrho v^2)^{h-m}][1-v+v(1-\varrho v)(\varrho v^2)^k]}{1-v+v(1-\varrho v)(\varrho v^2)^h} dv =$$

$$= \sum_{a\geqq 0} d_{a,h} \left[\binom{n;\alpha,\beta,\gamma}{n-m+k-a} - \binom{n;\alpha,\beta,\gamma}{n-m+k-a-1} \right] -$$

$$-\varrho^{h-m} \sum_{a\geqq 0} d_{a,h} \left[\binom{n;\alpha,\beta,\gamma}{n+m+k-2h-a} - \binom{n;\alpha,\beta,\gamma}{n+m+k-2h-a-1} \right] +$$

$$+\varrho^k \sum_{a\geqq 0} d_{a,h} \left[\binom{n;\alpha,\beta,\gamma}{n-m-k-a-1} - \varrho \binom{n;\alpha,\beta,\gamma}{n-m-k-a-2} \right] -$$

$$-\varrho^{h-m+k} \sum_{a\geqq 0} d_{a,h} \left[\binom{n;\alpha,\beta,\gamma}{n+m-k-2h-a-1} - \varrho \binom{n;\alpha,\beta,\gamma}{n+m-k-2h-a-2} \right].$$

It can be shown that (18) covers also the case $m \leqq k$.

At this point it is worthwhile to remark that in the spectral analysis method the derivation of the pgf. is not explicit and therefore the method may not lead to (18) directly. For the same reason the method cannot utilize the information obtained on various segments in the pgf. and thus is forced to start all over again when a new situation arises, as illustrated in Section 2.5. However, our approach, in which the explicit derivation of the pgf. is important, uses the basic pgf. (2) and its special cases (3) and (4) in deriving (13), which in turn will appear in Section 2.5.

From the pgf. $S(h,m,k)$ additional interesting information may be obtained. Let $\tau_{mh} = \inf\{n : Q(n) = h \mid Q(0) = m\}$, the time until a maximum queue length of h is reached for the first time. We note that the pgf. of τ_{mh} equals $s\alpha S(h,m,h-1)$ and therefore $P(\tau_{mh} < \infty) = 1$, since $S(h,m,h-1)|_{s=1} = 1$. The probability function of τ_{mh} is clearly

$$P(\tau_{mh} = n) = \alpha P(Q(n-1) = h-1, \max_{0 \leqq i \leqq n-1} Q(i) < h \mid Q(0) = m),$$

and therefore an explicit expression may be obtained from (18). For the expected value of τ_{mh} we find

$$E(\tau_{mh}) = \frac{d}{ds} s\alpha S(h,m,h-1)|_{s=1} =$$
$$= \frac{d}{ds} \frac{(\varrho v)^{h-m}(1-v+v(1-\varrho v)(\varrho v^2)^m)}{1-v+v(1-\varrho v)(\varrho v^2)^h}\Bigg|_{s=1}.$$

Now

$$\frac{dv}{ds} = \frac{(\alpha v^2 + \beta v + \gamma)^2}{\gamma(1-\varrho v^2)},$$

and therefore

$$E(\tau_{mh}) = \frac{\varrho^m - \varrho^h - \varrho^{h+m}(h-m)(1-p)}{\gamma(1-\varrho)^2 \varrho^{h+m}}, \qquad (\varrho \neq 1). \tag{19}$$

The results of this section have, as far as we know, not been reported in the literature until now.

2.4. Absorbing barrier at 0 and reflecting barrier at h. In the case where the random walk $Z(n)$ is restricted by a reflecting barrier at h and an absorbing barrier at 0, it behaves like a queuing process with finite capacity h during a period where the server is continuously busy. To fix notation we write $Q_h(n)$ for the queuing process with finite capacity h at time n. Then

$$P(Q_h(n) = k, Q(i) > 0, 0 \leqq i \leqq n \mid Q_h(0) = m)$$

can be simply derived from the results of the previous section by observing the following duality relation: consider the reversed sample paths of the queuing process $Q(n)$ restricted by an absorbing barrier at h and a reflecting barrier at 0. Then it is immediately seen that

$$(20)\qquad \begin{aligned} &P(Q(n)=k, \max_{0\leqq i\leqq n} Q(i)\mid Q(0)=m)=\\ &=P(Q_h^*(n)=h-m, Q_h^*(i)>0, 0\leqq i\leqq n\mid Q_h^*(0)=h-k), \end{aligned}$$

where the process $Q_h^*(n)$ is obtained from $Q_h(n)$ by interchanging the arrival and departure probabilities α and γ.

Let us denote the length of a busy period under finite capacity h by $\tau_{m0}^{(h)}$. Its distribution is found by arguments similar to those we used to derive the distribution of τ_{mh} in the previous section and by exploiting the duality relation (20). In particular we have

$$\{\tau_{m0}^{(h)}=n\}\equiv\{\tau_{h-m,h}^{*}=n\},$$

where the star indicates that the probabilities α and γ have to be interchanged. Therefore

$$(21)\qquad E(\tau_{m0}^{(h)})=E(\tau_{h-m,h}^{*})=\frac{\varrho^{h+m+1}-\varrho^{h+1}+m\varrho^m(1-\varrho)}{\gamma(1-\varrho)^2\varrho^m},\qquad(\varrho\neq 1).$$

2.5. Reflecting barriers at 0 and h. In this case the renewal argument may be formulated in terms of the random walk $Z(n)$ restricted by a reflecting barrier at zero, as it was discussed in Section 2.3. As renewal points we choose those time instants where $Z(n)$ touches the barrier at h. At the moment when $Z(n)$ touches this barrier, the waiting room is full and the system closes in the sense that customers arriving now are not allowed to enter the system and disappear. The system remains closed until the next departure which frees space in the waiting room. It is clear that the time the system remains closed has a geometric distribution with pgf.

$$(22)\qquad V(s)=\frac{\gamma s}{1-(1-\gamma)s}=\frac{v}{\varrho v^2-\varrho v+1},$$

where again $s=(\alpha v+\beta+\gamma/v)^{-1}$. From Section 2.3 it is known that the pgf. of τ_{mh}, the time required to reach a maximum value of h for the first time, is $\alpha s S(h,m,h-1)$. Suppose now that the system has been closed $i\geqq 0$ times and the sample path of the queuing process $Q_h(n)$ leads from m to k, where $m,k<h$. Then the corresponding pgf. is given by

$$(23)\qquad \begin{aligned} &T_i(h,m,k)=S(h,m,h-1)(\alpha sV)^i S^{i-1}(h,h-1,h-1)S(h,h-1,k)\quad(i\geqq 1)\\ &T_0(h,m,k)=S(h,m,k). \end{aligned}$$

Let $T(h,m,k)=\sum_{i\geqq 0}T_i(h,m,k)$. It follows that $T(h,m,k)$ is the pgf. of $P(Q_h(n)=k\mid Q_h(0)=m)$, in particular

$$(24)\qquad T(h,m,k)=S(h,m,k)+\frac{s\alpha S(h,m,h-1)VS(h,h-1,k)}{1-s\alpha VS(h,h-1,h-1)}.$$

This expression may be simplified further by observing that

$$(\varrho v^2)^a R(h-a) = R(h) - (1-v)(1-(\varrho v^2)^a)$$

and

$$(\varrho v^2 - \varrho v + 1)R(a) = (1-v)(1-\varrho v)(1-(\varrho v^2)^{a+1}),$$

where

$$R(a) = 1 - v + v(1-\varrho v)(\varrho v^2)^a.$$

Thus we find for $m \leqq k$:

(25)
$$T(h,m,k) =$$
$$= \frac{(\alpha v^2+\beta v+\gamma)(\varrho v)^{k-m}[1-v+v(1-\varrho v)(\varrho v^2)^m][1-\varrho v+\varrho v(1-v)(\varrho v^2)^{h-k}]}{\gamma(1-\varrho v^2)(1-v)(1-\varrho v)(1-(\varrho v^2)^{h+1})},$$

and similarly we have for $m \geqq k$:

(26)
$$T(h,m,k) =$$
$$= \frac{(\alpha v^2+\beta v+\gamma)v^{m-k}[1-v+v(1-\varrho v)(\varrho v^2)^k][1-\varrho v+\varrho v(1-v)(\varrho v^2)^{h-m}]}{\gamma(1-\varrho v^2)(1-v)(1-\varrho v)(1-(\varrho v^2)^{h+1})}.$$

In (25) and (26) the denominator unlike (13) exhibits a very simple structure, its roots being 1, $\pm\varrho^{-1/2}$, ϱ^{-1} and $\varrho^{-1/2}e^{\theta_\nu i}, \nu = 0, 1, \ldots, h$, where $\theta_\nu = \frac{\nu\pi}{h+1}$. Using the substitution $v = \varrho^{-1/2}e^{\theta i}$ and noting that there is again one complex angle, viz. $\theta = \frac{1}{2i}\log\varrho$, which is due to the factor $(1-v)$, one obtains after routine calculations:

$$P(Q_h(n) = k \mid Q_h(0) = m) =$$

(27)
$$= \frac{\varrho - 1}{\varrho^{h+1}-1}\varrho^k + \frac{2\alpha}{h+1}\sum_{\nu=1}^{h}\frac{(\beta+2\sqrt{\alpha\gamma}\cos\theta_\nu)^n}{1-\beta-2\sqrt{\alpha\gamma}\cos\theta_\nu}C_{mk}(\nu),$$

where

$$C_{mk}(\nu) = \left[\varrho^{-\frac{m+1}{2}}\sin(m+1)\theta_\nu - \varrho^{-\frac{m}{2}-1}\sin m\theta_\nu\right]\left[\varrho^{\frac{k+1}{2}}\sin(k+1)\theta_\nu - \varrho^{\frac{k}{2}}\sin k\theta_\nu\right],$$

which is valid even if k and (or) m are equal to h.

From (27) it may be immediately deduced that the steady state distribution is given by

(28)
$$\lim_{n\to\infty} P(Q_h(n) = k \mid Q_h(0) = m) = \frac{\varrho-1}{\varrho^{h+1}-1}\varrho^k.$$

On the other hand, if we expand the terms of the summation in (27) as geometric series, then representation (5) may be applied to give an expansion in terms of generalized trinomial coefficients, which turns out to be an interesting relation linking the transition probabilities of a random walk restricted by two absorbing barriers and the steady-state distribution (28):

$$
\begin{aligned}
P(Q_h(n)=k \mid Q_h(0)=m) &= \frac{\varrho-1}{\varrho^{h+1}-1}\varrho^k + \\
(29) \qquad +\alpha \sum_{m \geqq n} &[p_m(h+1,h-k,h-m) - p_m(h+1,h-k+1,h-m)] - \\
-\gamma \sum_{m \geqq n} &[p_m(h+1,h-k,h-m+1) - p_m(h+1,h-k+1,h-m+1)].
\end{aligned}
$$

At this point it may be instructive to see how the spectral decomposition method works. Let V_n denote the n-th component of an eigenvector associated with an eigenvalue σ of the transition probability matrix. Then it is verified that the three term recurrence relation induced by the eigenvector equation has the general solution (see Karlin and Taylor [4], pp. 10–18)

$$V_n(\theta) = Ae^{n\theta i} + Be^{-n\theta i},$$

and the eigenvalues will be of the following form:

$$\sigma = \beta + 2\sqrt{\alpha\gamma}\cos\theta.$$

The unknowns θ, A and B have to be determined according to the boundary conditions, which lead to the following homogeneous system:

$$A(\sqrt{\alpha\gamma}e^{-\theta i} - \gamma) + B(\sqrt{\alpha\gamma}e^{\theta i} - \gamma) = 0$$

$$A(\sqrt{\alpha\gamma}e^{(h+1)\theta i} - \alpha e^{h\theta i}) + B(\sqrt{\alpha\gamma}e^{-(h+1)\theta i} - \alpha e^{-h\theta i}) = 0.$$

After elimination of A and B we obtain the following equation, which determines the angles θ:

$$\varrho^{1/2}\sin(h+2)\theta - (1+\varrho)\sin(h+1)\theta + \varrho^{1/2}\sin h\theta = 0.$$

Since the sine function is odd, we immediately find that the roots are $\theta_\nu = \frac{\nu\pi}{h+1}$, $\nu = 1,2,\dots,h$. The root $\theta = 0$ has to be excluded, since in this case the corresponding eigenvector would be identically zero. However, it is not immediately apparent, that there is also one complex angle, viz. $\theta_0 = \frac{1}{2i}\log\varrho$, which we detected by simple inspection of (25). These angles determine the eigenvalues σ:

$$\sigma_\nu = \beta + 2\sqrt{\alpha\gamma}\cos\theta_\nu \qquad (\nu = 0,1,\dots,h).$$

Note that $\sigma_0 = 1$, which gives rise to the first term in (27) and hence to the steady state distribution. The eigenvectors corresponding to the eigenvalues

σ_ν are found by solving for the indeterminates A and B, where one of them is arbitrary. In particular, we find after some algebraic simplification that

$$V_n(\theta_0) = \varrho^{n/2}$$

$$V_n(\theta_\nu) = \sin n\theta_\nu - \varrho^{1/2}\sin(n+1)\theta_\nu \qquad (\nu > 0).$$

These eigenvectors form an orthogonal bases. If we divide them by their euclidean norm and denote their (normalized) components by $V_n^*(\theta_\nu)$, the transient solution is found to be

$$P(Q_h(n) = k \mid Q_h(0) = m) = \varrho^{\frac{k-m}{2}} \sum_{\nu=0}^{h} \sigma_\nu V_m^*(\theta_\nu) V_k^*(\theta_\nu),$$

which after some manipulations yields (27). It may be realized that the whole procedure has to be carried out right from the beginning for Sections 2.2 and 2.3. Moreover it is evident that our approach is reasonably elementary in contrast to the spectral analysis method.

3. Continuous time results

In the previous sections we have derived various results for the discrete time random walk $Z(n)$ and its associated queuing process $Q(n)$. As remarked in the introduction, it is possible to pass from discrete time to continuous time by a Poisson type limiting procedure. For this purpose we consider the time interval $(0,t)$ and split it into n subintervals of equal length $\Delta = t/n$. Set $\alpha = \lambda t/n$ and $\gamma = \mu t/n$. Now let $n \to \infty$ or, equivalently $\Delta \to 0$, while keeping λ, μ and t fixed. We expect that the finite dimensional distributions of the processes $Z(n)$ and $Q(n)$ converge to the distributions of the continuous time Markov processes $Z(t)$ and $Q(t)$, which have jump intensities λ and μ. Actually a much stronger result can be proved: it can be shown that even the infinite dimensional distributions converge in the weak sense, which entails the convergence of functionals of the discrete time processes such as stopping times. For details the reader is referred to Ethier and Kurtz [1], Chapter 2, Theorem 2.6, pp. 168–169. To derive the limiting forms of the finite dimensional distributions we will use the following results, which may be found in Mohanty and Panny [7], [8]:

As $n \to \infty$:

$$\binom{n;\alpha,\beta,\gamma}{n+a} = e^{-(\lambda+\mu)t}\varrho^{a/2} I_a(2t\sqrt{\lambda\mu})\Big(1+O\Big(\frac{1}{n^\varepsilon}\Big)\Big) \qquad (\varepsilon > 0),$$

(30)

$$(\beta + 2\sqrt{\alpha\gamma}\cos\theta)^n = e^{-(\lambda+\mu)t} e^{2t\sqrt{\lambda\mu}\cos\theta}\Big(1+O\Big(\frac{1}{n^\varepsilon}\Big)\Big) \qquad (\varepsilon > 0),$$

where $I_a(2t\sqrt{\lambda\mu})$ denotes the modified Bessel function of order a. Additionally we will require the following result:

$$\lim_{n\to\infty} \alpha \sum_{\nu\geqq n} \binom{\nu;\alpha,\beta,\gamma}{\nu+a} = \lambda\varrho^{a/2} \int_t^\infty e^{-(\lambda+\mu)s} I_a(2s\sqrt{\lambda\mu})ds. \tag{31}$$

Equation (31) follows directly from (30). In particular we have, observing that $n=t/\Delta$:

$$\lim_{n\to\infty} \alpha \sum_{\nu\geqq n} \binom{\nu;\alpha,\beta,\gamma}{\nu+a} = \lim_{n\to\infty} \lambda\Delta \sum_{i\geqq 0} e^{-(\lambda+\mu)(t+i\Delta)} \varrho^{a/2} I_a(2(t+i\Delta)\sqrt{\lambda\mu}) +$$
$$+ \lim_{n\to\infty} \lambda\Delta \sum_{i\geq 0} e^{-(\lambda+\mu)(t+i\Delta)} \varrho^{a/2} I_a(2(t+i\Delta)\sqrt{\lambda\mu}) O\Big(\frac{\Delta^\varepsilon}{(t+i\Delta^2)^\varepsilon}\Big).$$

The first sum is the Riemann sum of the integral in (31) and the second sum tends to zero as $n\to\infty$, thus (31) follows.

Using (30), we find in the case of two absorbing barriers through equation (5):

$$p_t(h,m,k) = \tag{32}$$
$$= \frac{2}{h}\varrho^{\frac{k-m}{2}} e^{-(\lambda+\mu)t} \sum_{\nu=1}^{h-1} \sin\frac{(h-k)\nu\pi}{h} \sin\frac{(h-m)\nu\pi}{h} \exp\Big[2t\sqrt{\lambda\mu}\cos\frac{\nu\pi}{h}\Big],$$

a result, which may be found in Neuts [10]. Alternatively we obtain from (7):

$$p_t(h,m,k) = \varrho^{\frac{k-m}{2}} e^{-(\lambda+\mu)t} \sum_{\nu=-\infty}^{\infty} [I_{k-m-2\nu h} - I_{k+m-2\nu(h+1)}], \tag{33}$$

where I_a is an abbreviation of $I_a(2t\sqrt{\lambda\nu})$. To deal with the case of a reflecting barrier at zero and an absorbing barrier at h, we first consider formula (16). Observe that the roots θ_ν do not depend on n. Hence using (30) we find

$$P(Q(t)=k, \max_{0\leqq s\leqq t} Q(s)<h \mid Q(0)=m) = \tag{34}$$
$$= -2\varrho^{\frac{k-m}{2}} e^{-(\lambda+\mu)t} \sum_{\nu=1}^{h} e^{2t\sqrt{\lambda\mu}\cos\theta_\nu} D_{mk}(\nu),$$

where

$$D_{mk}(\nu) = \frac{\sin(h-k)\theta_\nu[\sqrt{\lambda}\sin(m+1)\theta_\nu - \sqrt{\mu}\sin m\theta_\nu]}{\sqrt{\lambda}(h+1)\cos(h+1)\theta_\nu - \sqrt{\mu}h\cos h\theta_\nu}.$$

To derive the limiting form of (18) we note that the coefficients $d_{a,h}$ are independent of n, hence they enter into the limiting expression without change. In particular we find:

$$P(Q(t)=k, \max_{0\leqq s\leqq t} Q(s)<h \mid Q(0)=m)=$$

$$= e^{-(\lambda+\mu)t}\varrho^{\frac{k-m}{2}}\sum_{a\geqq 0} d_{a,h}\varrho^{-a/2}\times \tag{35}$$

$$\times\Big[I_{k-m-a}-\varrho^{-1/2}I_{k-m-a-1}-I_{k+m-2h-a}+\varrho^{-1/2}I_{k+m-2h-a-1}-$$

$$-I_{k+m+a-2}+\varrho^{-1/2}I_{k+m+a+1}+I_{k-m+2h+a+2}-\varrho^{-1/2}I_{k-m+2h+a+1}\Big].$$

It remains to show the absolute convergence of the series above. For this purpose we recall that

$$d_{a,h}=[v^a](1-v+(1-\varrho v)(\varrho v^2)^h)^{-1}.$$

The function $(1-v+(1-\varrho v)(\varrho v^2)^h)^{-1}$ is analytic in a disc around the origin and has a positive radius of convergence $R\leqq 1$. Hence there is an integer A, such that

$$d_{a,h}<\left(\frac{1}{R}+\varepsilon\right)^a \qquad (\varepsilon>0 \text{ and } a>A).$$

Thus setting $z=\varepsilon+1/R$ we estimate

$$\sum_{a\geqq A} d_{a,h}\varrho^{-a/2}I_{k+a}(2t\sqrt{\lambda\mu})\leqq\sum_{a\geqq A} z^a\varrho^{-a/2}(t\sqrt{\lambda\mu})^{k+a}\sum_{i\geqq 0}\frac{(t^2\lambda\mu)^i}{i!(i+k+a)!}\leqq$$

$$\leqq e^{t^2\lambda\mu}\sum_{a\geqq A}\frac{z^a\varrho^{-a/2}(t\sqrt{\lambda\mu})^{k+a}}{a!}\quad\to 0 \quad\text{as}\quad A\to\infty.$$

In the case of two reflecting barriers at zero and at h we get by means of (30) the limit of (27):

$$P(Q_h(t)=k\mid Q_h(0)=m)=$$

$$=\frac{\varrho-1}{\varrho^{h+1}-1}\varrho^k+2\lambda\frac{e^{-(\lambda+\mu)t}}{h+1}\sum_{\nu=1}^{h}\frac{e^{2t\sqrt{\lambda\mu}\cos\theta_\nu}}{\lambda-2\sqrt{\lambda\mu}\cos\theta_\nu+\mu}C_{mk}(\nu), \tag{36}$$

where again

$$C_{mk}(\nu)=\left[\varrho^{-\frac{m+1}{2}}\sin(m+1)\theta_\nu-\varrho^{-\frac{m}{2}-1}\sin m\theta_\nu\right]\left[\varrho^{\frac{k+1}{2}}\sin(k+1)\theta_\nu-\varrho^{\frac{k}{2}}\sin k\theta_\nu\right].$$

Formula (36) is well known, it may be found in Morse [9] or Takács ([12], p. 13). Finally the limit of (29) is found using (31) and expansion (7):

$$P(Q_h(t)=k \mid Q_h(0)=m)=\frac{\varrho-1}{\varrho^{h+1}-1}\varrho^k+\varrho^{\frac{k-m}{2}}\sum_{\nu=-\infty}^{\infty}\int_t^{\infty}e^{-(\lambda+\mu)s}\times$$

(37)
$$\times\Big[I_{k-m-2\nu(h+1)}+2\sqrt{\lambda\mu}I_{k+m+2\nu(h+1)}-\sqrt{\lambda\mu}I_{k-m-1-2\nu(h+1)}-$$
$$-\sqrt{\lambda\mu}I_{k-m+1-2\nu(h+1)}-\lambda I_{k+m+2+2\nu(h+1)}-\mu I_{k+m+2\nu(h+1)}\Big]ds.$$

A formula similar to (37) is given by Kashyap [5].

Acknowledgements. This work was done when the first author visited McMaster University during summer 1990 and was supported by the grant from the NSERC, Canada to the second author.

REFERENCES

[1] Ethier, S. N. and Kurtz, T. G., *Markov processes. Characterization and convergence*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley and Sons, Inc., New York, 1986. *MR* **88a**:60130

[2] Goulden, I. P. and Jackson, D. M., *Combinatorial enumeration*, Wiley-Interscience Series in Discrete Mathematics, John Wiley and Sons, New York, 1983. *MR* **84m**:05002

[3] Karlin, S. and McGregor, J. L., Random walks, *Illinois J. Math.* **3** (1959), 66–81. *MR* **20** #7352

[4] Karlin, S. and Taylor, H. K., *A second course in stochastic processes*, Academic Press, New York–London, 1981. *MR* **82j**:60003

[5] Kashyap, B. R. K., The random walk with partially reflecting barriers with application to queuing theory, *Proc. Nat. Inst. Sci. India Part A* **31** (1965), 527–535. *MR* **35** #6225

[6] Kemperman, J. H. B., The passage problem for a stationary Markov chain, Statistical Research Monographs, Vol. I, The University of Chicago Press, Chicago, Ill., 1961. *MR* **22** #9992

[7] Mohanty, S. G. and Panny, W., A discrete time analogue of the $M/M/1$ queue and the transient solution: a geometric approach, *Sankhyā Ser. A* (1990), 364–370.

[8] Mohanty, S. G. and Panny, W., A discrete time analogue of the $M/M/1$ queue and the transient solution: an analytic approach, *Proc. 3rd Hungarian Colloquium on Limit Theorems in Probability and Statistics*, P. Révész ed., North-Holland, Amsterdam, 1990.

[9] Morse, P. M., *Queues, inventories and maintenance. The analysis of operational systems with variable demand and supply*, Publications in Operations Research, Operations Research Society of America, No. 1, John Wiley and Sons, Inc., New York; Chapman and Hall, Ltd., London, 1958. *MR* **19**–930

[10] Neuts, M. F., The distribution of the maximum length of a Poisson queue during a busy period, *Operations Res.* **12** (1964), 281–285. *MR* **28** #5491

[11] Panny, W., *The maximal deviation of lattice paths*, Mathematical Systems in Economics, 91, Verlagsgruppe Athenäum/Hain/Hanstein, Königstein Ts., 1984. *MR* **86e**:60012

[12] Takács, L., *Introduction to the theory of queues*, University Texts in the Mathematical Sciences, Oxford University Press, New York, 1962. *MR* **24** #A3704

(Received July 9, 1991)

MATHEMATISCHES INSTITUT DER
WIRTSCHAFTSUNIVERSITÄT WIEN
AUGASSE 2–6
A–1090 WIEN
AUSTRIA

DEPARTMENT OF MATHEMATICS
McMASTER UNIVERSITY
HAMILTON, ONTARIO
L8S 4K1
CANADA

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 415–432

SOME 2-PERIODIC TRIGONOMETRIC INTERPOLATION PROBLEMS ON EQUIDISTANT NODES II: CONVERGENCE

A. SHARMA, J. SZABADOS[1] and R. S. VARGA

1. Introduction

Let $n, p, q \geqq 1$ be integers with

$$N := n(p+q), \quad M = \left[\frac{n}{2}\right]$$

and let

$$(1) \qquad x_n = x_k(n) := \frac{k\pi}{n} \quad (k = 0, 1, \ldots, 2n-1).$$

Let $\mathbf{m}_1 = (m_1, \ldots, m_p)$, $\mathbf{m}_2 = (m_{p+1}, \ldots, m_{p+q})$ be two sequences of integers such that

$$(2) \qquad 0 = m_1 < m_2 < \cdots < m_p, \quad 0 \leqq m_{p+1} < \cdots < m_{p+q}.$$

The problem of 2-periodic trigonometric interpolation is to reconstruct the unique trigonometric polynomial

$$(3) \quad t_M(x) = a_0 + \sum_{k=1}^{M} (a_k \cos kx + b_k \sin kx) \quad (a_M b_M = 0 \quad \text{if} \quad N \quad \text{is even}),$$

from the data

$$t_M^{(m_\mu)}(x_{2k}), \quad t_M^{(m_\nu)}(x_{2k+1}) \quad (\mu = 1, \ldots, p;\ \nu = p+1, \ldots, p+q),$$

for given sequences of integers $\mathbf{m}_1, \mathbf{m}_2$ satisfying (2).

1980 *Mathematics Subject Classification* (1985 Revision). Primary 42A15; Secondary 41A10.

Key words and phrases. 2-periodic interpolation, fundamental polynomials, convergence.

[1]Research was supported by Hungarian National Research Fund Grant No. 1801, and by NSERC Grant No. A3094.

Recently in [2], we gave necessary and (separately) sufficient conditions for the regularity (or unique solvability) of the above problem. For studying the convergence problem, we shall suppose that the sufficient conditions in [2] are satisfied. We shall state these conditions for the sake of completeness, but in a different form.

We shall say that a finite sequence of non-negative integers $\mathbf{m} = (m_1, m_2, \dots, m_p)$ such that

$$(3) \qquad 0 \leqq m_1 < m_2 < \cdots < m_p \quad \text{and} \quad m_i + m_{i+1} \quad \text{is odd} \quad (i = 1, \dots, p-1)$$

is EE, EO, OE or OO according as m_1, m_p are both even; m_1 is even, m_p is odd; m_1 is odd, m_p is even or m_1, m_p are both odd, respectively. A sequence with only one element will be either EE or OO depending on the parity of the element. It is clear from this definition that EE and OO sequences have odd cardinality while EO and OE sequences have even cardinality. With this definition we can now state

THEOREM A [2]. *The 2-periodic trigonometric interpolation problem on the nodes* (1) *corresponding to the sequences* $\mathbf{m}_1, \mathbf{m}_2$ *is regular in the following cases:*

		$\mathbf{m}_1$	$\mathbf{m}_2$	n	*Type of* T_M
I	(a)	EO	EE	*odd*	—
II	(a)	EE	OE	*odd*	—
III	(b)	EO	EO	*arb.*	$a_M = 0$
IV	(b)	EE	OO	*odd*	$a_M = 0$
V	(b)	EE	EE	*arb.*	$b_M = 0$
VI	(b)	EO	OE	*arb.*	$a_M = 0$
VII	(c)	EE	EO	*even*	$b_M = 0$

REMARK. Since by supposition $\mathbf{m}_1$ begins with an even number, there are only 8 possible combinations of $\mathbf{m}_1, \mathbf{m}_2$ in the table. The pair $\mathbf{m}_1 = EO$, and $\mathbf{m}_2 = OO$ is not in the above table, because in this case even the necessary conditions of regularity of Theorems 1 and 2 in [2] are not satisfied. The conditions which are necessary for regularity as given in [2] imply that

$$(4) \qquad e - o = 0, 1 \quad \text{or} \quad 2$$

where e and o denote the cardinality of even and odd numbers in the set $\mathbf{m}_1 \cup \mathbf{m}_2$, respectively. But if $\mathbf{m}_1 = EO$, $\mathbf{m}_2 = OO$, then $o - e = 1$, which contradicts (4).

These conditions are sufficient, but not necessary as can be seen by the examples in [2] and also by the results in [4]. In [3] it was shown that if $\mathbf{m}_1 := (0, m_1, \dots, m_p)$ and if $\mathbf{m}_2 = (m_1, \dots, m_p)$, then the problem is regular if and only if p is even and then $b_M = 0$. In this case no condition like $m_i + m_{i+1}$ odd is needed.

The object of this note is twofold: First we want to construct the fundamental polynomials of interpolation in all the above cases. Secondly, we want to examine the convergence of the interpolant.

In Section 2, we find the fundamental polynomials when n is odd and $p+q=2s+1$ (cases covered in (a)). In Section 3, we find the fundamental polynomials when $p+q$ is even and n is arbitrary. This covers the four situations listed in (b) in the Table. Lastly, Section 4 is devoted to the last case (c) in the Table when n is even and $p+q$ is odd. Section 5 deals with the problem of convergence.

2. Fundamental polynomials (n odd, $p+q$ odd)

Here $n = 2r+1$, $p+q = 2s+1$ and $M = ns + r$. This case covers the first two cases in the table. We shall denote the fundamental polynomials by $\varrho_\nu(x)$ which are going to be determined by the followig conditions for $\nu = 1, \dots, p$:

$$(2.1)\qquad \begin{cases} \varrho_\nu^{(m_j)}(x_{2k}) = \delta_{k0}\delta_{\nu j}, & j = 1, \dots, p; \quad k = 0, 1, \dots, n-1 \\ \varrho_\nu^{(m_j)}(x_{2k+1}) = 0, & j = p+1, \dots, p+q; \quad k = 0, 1, \dots, n-1. \end{cases}$$

For $\nu = p+1, \dots, p+q$, conditions (2.1) will be replaced by

$$(2.2)\qquad \begin{cases} \varrho_\nu^{(m_j)}(x_{2k}) = 0, & j = 1, \dots, p; \ k = 0, 1, \dots, n-1 \\ \varrho_\nu^{(m_j)}(x_{2k+1}) = \delta_{k0}\delta_{\nu j}, & j = p+1, \dots, p+q; \ k = 0, 1, \dots, n-1. \end{cases}$$

Putting $z = e^{ix}$, we may set

$$(2.3)\qquad \varrho_\nu(x) = z^{-M} \sum_{\lambda=0}^{2s} z^{\lambda n} Q_{\nu\lambda}(Z), \quad Q_{\nu\lambda}(z) = \sum_{j=0}^{n-1} a_{\lambda j}(\nu) z^j.$$

Conditions (2.1) (equivalently (2.2)) give the following system of $2s+1$ differential equations to determine $Q_{\nu\lambda}(z)$:

(2.4)

$$\begin{cases} \sum_{\lambda=0}^{2s} (\Theta + \lambda n - M)^{m_\mu} Q_{\nu\lambda}(z) = \dfrac{i^{-m_\nu}}{n} \dfrac{z^n - 1}{z - 1} \delta_{\mu\nu}, & \mu = 1, \dots, p \\ \sum_{\lambda=0}^{2s} (-1)^\lambda (\Theta + \lambda n - M)^{m_\mu} Q_{\nu\lambda}(z) = -\dfrac{i^{-m_\nu}}{n} \dfrac{z^n + 1}{z + 1} \delta_{\mu\nu}, & \mu = p+1, \dots, p+q \end{cases}$$

where $\Theta = z\frac{d}{dz}$. If we denote the determinant of this system by $\Delta(\Theta)$ and the co-factor of the $(\nu, \lambda+1)$ term by $\Delta_{\nu,\lambda+1}(\Theta)$, then we have

$$Q_{\nu\lambda}(z) = \begin{cases} \dfrac{i^{-m_\nu}}{n}\dfrac{\Delta_{\nu,\lambda+1}(\Theta)}{\Delta(\Theta)}\dfrac{z^n-1}{z-1}, & \nu = 1,\dots,p \\ -\dfrac{i^{-m_\nu}}{n}\dfrac{\Delta_{\nu,\lambda+1}(\Theta)}{\Delta(\Theta)}\dfrac{z^n+1}{z+1}, & \nu = p+1,\dots,p+q. \end{cases}$$

Since $\Theta z^j = jz^j$, we see easily that

$$\frac{\Delta_{\nu,\lambda+1}(\Theta)}{\Delta(\Theta)}z^j = \frac{1}{n^{m_\nu}}\frac{D_{\nu,\lambda+1}(\alpha_j)}{D(\alpha_j)}z^j, \quad \alpha_j = \frac{j-r}{n}$$

where

$$D(\alpha) = \begin{vmatrix} 1 & 1 & \dots & 1 \\ (\alpha-s)^{m_2} & (\alpha-s+1)^{m_2} & \dots & (\alpha+s)^{m_2} \\ \dots & \dots & \dots & \dots \\ (\alpha-s)^{m_p} & (\alpha-s+1)^{m_p} & \dots & (\alpha+s)^{m_p} \\ (\alpha-s)^{m_{p+1}} & -(\alpha-s+1)^{m_{p+1}} & \dots & (\alpha+s)^{m_{p+1}} \\ \dots & \dots & \dots & \dots \\ (\alpha-s)^{m_{p+q}} & -(\alpha-s+1)^{m_{p+q}} & \dots & (\alpha+s)^{m_{p+q}} \end{vmatrix} \tag{2.5}$$

and $D_{\nu,\lambda+1}(\alpha)$ is the cofactor of $(\nu, \lambda+1)$ terms in $D(\alpha)$. In the last q rows in $D(\alpha)$, the columns have alternately positive, negative signs.

Setting

$$\alpha_{\lambda j}(\nu) := \frac{D_{\nu,\lambda+1}(\alpha_j)}{D(\alpha_j)}, \qquad \begin{array}{l} \nu = 1,\dots,p+q; \quad \lambda = 0,1,\dots,2s \\ j = 0,1,\dots,n-1 \end{array} \tag{2.6}$$

we see that

$$Q_{\nu\lambda}(z) = \begin{cases} \dfrac{i^{-m_\nu}}{n^{1+m_\nu}}\displaystyle\sum_{j=0}^{n-1} a_{\lambda j}(\nu)z^j, & \nu = 1,\dots,p \\ -\dfrac{i^{-m_\nu}}{n^{1+m_\nu}}\displaystyle\sum_{j=0}^{n-1}(-1)^j a_{\lambda j}(\nu)z^j, & \nu = p+1,\dots,p+q \end{cases} \tag{2.7}$$

since n is odd.

It has been shown in [2] that under the hypothesis of Theorem 1 the determinant $D(\alpha) \neq 0$ for $|\alpha| \leq 1/2$. Also, if e and o denote the number of even and odd integers in the sequence $m_1,\dots,m_{p+q}$, then

$$e - 1 = o$$

is necessary for regularity. Multiplying the rows in $D(-\alpha)$ corresponding to odd m_ν's by (-1) and then doing s elementary column operators, we see that

$$(-1)^o D(-\alpha) = (-1)^s D(\alpha)$$

which proves that

$$D(-\alpha) = D(\alpha). \tag{2.8}$$

Similarly, we can see that

$$D_{\nu,\lambda+1}(-\alpha) = (-1)^{m_\nu} D_{\nu,2s+1-\lambda}(\alpha), \quad \nu = 1, \dots, p+q. \tag{2.9}$$

From (2.6), (2.8) and (2.9), we have

$$a_{\lambda j}(\nu) = (-1)^{m_\nu} a_{2s-\lambda,2r-j}(\nu), \quad \nu = 1, \dots, p+q. \tag{2.10}$$

Combining (2.7) and (2.10), we obtain

$$\begin{aligned} \varrho_\nu(x) = \frac{(-1)^{m_\nu/2}}{n^{1+m_\nu}} \Big[a_{sr}(\nu) + 2\sum_{j=1}^{r} a_{s,r-j}(\nu) \cos jx + \\ + 2\sum_{\lambda=1}^{s} \sum_{j=0}^{n-1} a_{s-\lambda,j}(\nu) \cos(\lambda n + r - j)x \Big] \end{aligned} \tag{2.11}$$

when m_ν is even and $1 \leqq \nu \leqq p$. When m_ν is odd, because of (2.10) we see that $\varrho_\nu(x)$ is a sine series. More precisely, we have

$$\begin{aligned} \varrho_\nu(x) = \frac{2(-1)^{\frac{m_\nu-1}{2}}}{n^{1+m_\nu}} \Big[\sum_{j=1}^{r} a_{s,r-j} \sin jx + \\ + \sum_{\lambda=1}^{s} \sum_{j=0}^{n-1} a_{s-\lambda,j}(\nu) \sin(\lambda n + r - j)x \Big] \end{aligned} \tag{2.12}$$

when m_ν is odd and $1 \leqq \nu \leqq p$.

For $\nu = p+1, \dots, p+q$, $\varrho_\nu(x)$ is obtained from (2.11) or (2.12) according as m_ν is even or odd, respectively, by replacing x by $x - \frac{\pi}{n}$.

3. Fundamental polynomials ($p+q$ even)

Here $p+q = 2s+2$ and $M = ns+n$. So we set

$$\varrho_\nu(x) := z^{-M} \Big[\sum_{\lambda=0}^{2s+1} z^{\lambda n} Q_{\nu\lambda}(z) + C_\nu z^{(2s+2)n} \Big], \quad Q_{\nu\lambda}(z) = \sum_{j=0}^{n-1} a_{\lambda j}(\nu) z^j. \tag{3.1}$$

The conditions which determine $\varrho_\nu(x)$ are given by (2.1) or (2.2) and lead to the following systems of equations:

$$(3.2)\quad \begin{cases} \sum_{\lambda=0}^{2s+1}(\Theta+\lambda n-M)^{m_\mu}Q_{\nu\lambda}(z)+C_\nu M^{m_\mu}=\frac{i^{-m_\nu}}{n}\frac{z^n-1}{z-1}\delta_{\mu\nu} \\ \qquad (\mu=1,2,\dots,p), \\ \sum_{\lambda=0}^{2s+1}(-1)^\lambda(\Theta+\lambda n-M)^{m_\mu}Q_{\nu\lambda}(z)+C_\nu M^{m_\mu}=\frac{i^{-m_\nu}}{n}\frac{z^n+1}{z-\omega}\omega^{n-1}\delta_{\mu\nu} \\ \qquad (\mu=p+1,\dots,p+q), \\ \delta_0 Q_{\nu 0}+(-1)^{1+\varepsilon}C_\nu=0, \end{cases}$$

where δ_0 in the last equation is the point evaluation at 0 and $\omega=\exp\frac{i\pi}{n}$. The last condition is a consequence of the fact that the last term in $\varrho_\nu(x)$ is $a_M\cos(Mx+\frac{\varepsilon\pi}{2})$, where $\varepsilon=0$ or 1.

If we denote the determinant of this system by $\Delta^*(\Theta)$ and the cofactors by $\Delta^*_{kl}(\Theta)$, then we have

$$(3.3)\qquad Q_{\nu,\lambda}(z)=\begin{cases} \frac{i^{-m_\nu}}{n}\frac{\Delta^*_{\nu,\lambda+1}(\Theta)}{\Delta^*(\Theta)}\frac{z^n-1}{z-1}, & \nu=1,\dots,p \\ \frac{i^{-m_\nu}}{n}\frac{\Delta^*_{\nu,\lambda+1}(\Theta)}{\Delta^*(\Theta)}\frac{Z^n-1}{Z-1}, & \nu=p+1,\dots,p+q \end{cases}$$

for $\lambda=0,1,\dots,2s+1$, where we have put $Z=ze^{-\frac{i\pi}{n}}$. Also C_ν is given by formulae (3.3) when $\lambda=2s+2$.

Because of the point evaluation operator δ_0 in the determinant $\Delta^*(\Theta)$, it is easy to see that

$$(3.4)\qquad \frac{\Delta^*_{\nu,\lambda+1}(\Theta)}{\Delta^*(\Theta)}z^j=\begin{cases} \frac{1}{n^{m_\nu}}\frac{\bar{D}_{\nu,\lambda+1}(\alpha_j)}{\bar{D}(\alpha_j)}z^j, & \lambda=0,1,\dots,2s+1 \\ 0, & \lambda=2s+2 \end{cases}$$

for $j=1,2,\dots,n-1$ where $\alpha_j=\frac{j}{n}$ $(j=1,\dots,n-1)$ and

$$(3.5)\qquad \bar{D}(\alpha):=\begin{vmatrix} 1 & 1 & \dots & 1 \\ (\alpha-s-1)^{m_1} & (\alpha-s)^{m_1} & \dots & (\alpha+s)^{m_1} \\ \dots & \dots & \dots & \dots \\ (\alpha-s-1)^{m_p} & (\alpha-s)^{m_p} & \dots & (\alpha+s)^{m_p} \\ (\alpha-s-1)^{m_{p+1}} & -(\alpha-s)^{m_{p+1}} & \dots & -(\alpha+s)^{m_{p+1}} \\ \dots & \dots & \dots & \dots \\ (\alpha-s-1)^{m_{p+q}} & -(\alpha-s)^{m_{p+q}} & \dots & -(\alpha+s)^{m_{p+q}} \end{vmatrix}$$

and $\bar{D}_{\nu,\lambda+1(\alpha)}$ is a cofactor of $\bar{D}(\alpha)$. The order of $\bar{D}(\alpha)$ is $2s+2$ and the last q rows have alternating sign in the columns.

When $j=0$, we see that

$$\frac{\Delta^*_{\nu,\lambda+1}(\Theta)}{\Delta^*(\Theta)}z^0 = \frac{1}{n^{m_\nu}}\frac{\bar{D}^*_{\nu,\lambda+1}}{\bar{D}^*}, \quad \lambda=0,1,\dots,2s+2 \tag{3.6}$$

where

$$\bar{D}^* := \begin{vmatrix} 1 & 1 & \dots & 1 & 1 \\ (-s-1)^{m_1} & (-s)^{m_1} & \dots & s^{m_1} & (s+1)^{m_1} \\ \dots & \dots & \dots & \dots & \dots \\ (-s-1)^{m_p} & (-s)^{m_p} & \dots & s^{m_p} & (s+1)^{m_p} \\ (-s-1)^{m_{p+1}} & -(-s)^{m_{p+1}} & \dots & -s^{m_{p+1}} & (s+1)^{m_{p+1}} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 0 & \dots & 0 & (-1)^{1+\varepsilon} \end{vmatrix} \tag{3.7}$$

and $\bar{D}^*_{\nu,\lambda+1}$ is the cofactor of $\bar{D}^*$. It follows from (3.3) and (3.6) that C_ν is a constant given by

$$C_\nu = \frac{i^{-m_\nu}}{n}\frac{D^*_{\nu,2s+3}}{D^*}, \quad \nu=1,\dots,p+q. \tag{3.8}$$

If we now set

$$a_{\lambda j}(\nu) := \begin{cases} \dfrac{D_{\nu,\lambda+1}(\alpha_j)}{D(\alpha_j)}, & 0<\alpha_j\leqq 1 \\ \dfrac{D^*_{\nu,\lambda+1}}{D^*}, & \alpha_j=0 \end{cases} \qquad (\lambda=0,1,\dots,2s+2), \tag{3.9}$$

then we have

$$Q_{\nu\lambda}(z) = \frac{i^{-m_\nu}}{n^{1+m_\nu}}\sum_{j=0}^{n=1}\alpha\lambda_j(\nu)z^j \qquad (\lambda=0,1,\dots,2s+1),$$

$$C_\nu := \frac{i^{-m_\nu}}{n}\frac{D^*_{\nu,2s+3}}{D^*}$$

so that

$$\varrho_\nu(x) = \frac{i^{-m_\nu}}{n}z^{-sn-n}\left[\sum_{\lambda=0}^{2s+1}z^{\lambda n}\sum_{j=0}^{n-1}a_{\lambda j}(\nu)z^j + a_{2s+2,n}(\nu)z^{(2s+2)n}\right] \tag{3.10}$$

From (3.5), we see easily that

$$\begin{cases} D(\alpha)=D(1-\alpha) \\ D_{\nu,\lambda+1}(\alpha) = (-1)^{m_\nu}D_{\nu,2s+2-\lambda}(1-\alpha) \quad (\lambda=0,1,\dots,2s+1). \end{cases} \tag{3.11}$$

Similarly,

$$D^*_{\nu,\lambda+1} = (-1)^{m_\nu} D^*_{\nu,2s+3-\lambda},$$

so that from (3.9) and (3.11) we have

$$a_{\lambda j}(\nu) = (-1)^{m_\nu} a_{2s+2-\lambda,n-j}(\nu), \quad j = 1,\dots,n-1$$
$$a_{\lambda 0}(\nu) = (-1)^{m_\nu} a_{2s+s-\lambda,n}(\nu).$$

Thus we have from (3.10)

(3.12) $$\varrho_\nu(x) =$$

$$= \begin{cases} \dfrac{2(-1)^{\frac{m_\nu-1}{2}}}{n^{1+m_\nu}} \Big[\displaystyle\sum_{\lambda=0}^{s} \sum_{j=0}^{n=1} a_{\lambda j}(\nu) \sin((s+1-\lambda)n-j)x \Big], & m_\nu \text{ odd} \\ \dfrac{(-1)^{\frac{m_\nu}{2}}}{n^{1+m_\nu}} \Big[a_{s+1,0}(\nu) + 2 \displaystyle\sum_{\lambda=0}^{s} \sum_{j=0}^{n-1} a_{\lambda j}(\nu) \cos((s+1-\lambda)n-j)x \Big], & m_\nu \text{ even.} \end{cases}$$

4. Fundamental polynomials (n even, $p+q$ odd)

In this case $n = 2r, p+q = 2s+1$ and $M = ns+r$ so that we may set

(4.1) $$\varrho_\nu(x) = z^{-M} \sum_{\lambda=0}^{2s} z^{\lambda n} Q_\lambda(z) + C_\nu z^M, \quad Q_{\nu\lambda}(z) \in \pi_{n-1}.$$

The system of differential equations as in Section 3 is given by

(4.2) $$\begin{cases} \displaystyle\sum_{\lambda=0}^{2s} (\Theta + \lambda n - M)^{m_\mu} Q_{\nu\lambda}(z) + C_\nu M^{m_\mu} = \frac{i^{-m_\nu}}{n} \frac{z^n - 1}{z-1} \delta_{\mu\nu} \\ \qquad (\mu = 1,\dots,p) \\ \displaystyle\sum_{\lambda=0}^{2s} (-1)^\lambda (\Theta + \lambda n - M)^{m_\mu} Q_{\nu\lambda}(z) + C_\nu M^{m_\mu} = -\frac{i^{-m_\nu}}{n} \frac{z^n + 1}{z+1} \delta_{\mu\nu} \\ \qquad (\mu = p+1,\dots,p+q) \\ \delta_0 Q_{\nu,0}(z) + (-1)^{1+\varepsilon} C_\nu = 0. \end{cases}$$

Then, as in Section 3, we obtain

$$Q_{\nu\lambda}(z) = \begin{cases} \dfrac{i^{-m_\nu}}{n^{1+m_\nu}} \displaystyle\sum_{j=0}^{n-1} a_{\lambda j}(\nu) z^j, & \nu = 1,\dots,p \\ \dfrac{i^{-m_\nu}}{n^{1+m_\nu}} \displaystyle\sum_{j=0}^{n-1} (-1)^j a_{\lambda j}(\nu) z^j, & \nu = p+1,\dots,p+q \end{cases}$$

where

$$a_{\lambda j}(\nu) = \frac{D_{\nu,\lambda+1}(\alpha_j)}{D(\alpha_j)}, \quad \alpha_j = \frac{j-r}{n} \quad (\lambda = 0,1,\ldots,2s)$$
$$(\nu = 1,\ldots,p+q; \quad j = 1,\ldots,n-1),$$

the determinant $D(\alpha)$ being the same as in (2.5). However, for $j=0$, we have

$$a_{\lambda 0}(\nu) = \frac{D^*_{\nu,\lambda+1}}{D^*} \quad (\lambda = 0,1,\ldots,2s), \tag{4.4}$$

where

(4.5)

$$D^* = \begin{vmatrix} 1 & 1 & \ldots & 1 & 1 \\ (\alpha_0 - s)^{m_1} & (\alpha_0+1-s)^{m_1} & \ldots & (\alpha_0+s)^{m_1} & (\alpha_0+s+1)^{m_1} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ (\alpha_0 - s)^{m_p} & (\alpha_0+1-s)^{m_p} & \ldots & (\alpha_0+s)^{m_p} & (\alpha_0+s+1)^{m_p} \\ (\alpha_0 - s)^{m_{p+1}} & -(\alpha_0+1-s)^{m_{p+1}} & \ldots & (\alpha_0+s)^{m_{p+1}} & -(\alpha_0+s+1)^{m_{p+1}} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ (\alpha_0 - s)^{m_{p+q}} & -(\alpha_0+1-s)^{m_{p+q}} & \ldots & (\alpha_0+s)^{m_{p+q}} & -(\alpha_0+s+1)^{m_{p+q}} \\ 1 & 0 & \ldots & 0 & (-1)^{1+\epsilon} \end{vmatrix}$$

and $D^*_{\nu,\lambda+1}$ is its cofactor. From (4.2), we see, as in Section 3, that

$$C_\nu = \frac{i^{-m_\nu}}{n} \frac{D^*_{\nu,2s+2}}{D^*} \quad (\nu = 1,\ldots,p+q). \tag{4.6}$$

As in Section 2, (2.10) is valid in this case also for $\nu = 1,\ldots,p+q$. Thus after some simplification, we obtain

$$\varrho_\nu(x) = \begin{cases} \dfrac{(-1)^{\frac{m_\nu}{2}}}{n^{1+m_\nu}} \Big[a_{sr}(\nu) + 2\sum_{j=1}^{r} a_{s,r-j}(\nu)\cos jx + \\ \quad + 2\sum_{\lambda=1}^{s}\sum_{j=0}^{n-1} a_{s-\lambda,j}(\nu)\cos(\lambda n + r - j)x\Big], & m_\nu \text{ even} \\ \dfrac{2(-1)^{\frac{m_\nu-1}{2}}}{n^{1+m_\nu}} \Big[\sum_{j=1}^{r} a_{s,r-j}(\nu)\sin jx + \\ \quad + \sum_{\lambda=1}^{s}\sum_{j=0}^{n=1} a_{s-\lambda,j}(\nu)\sin(\lambda n + r - j)x\Big], & m_\nu \text{ odd.} \end{cases} \tag{4.7}$$

5. Convergence of 2-periodic interpolation

The definition of 2-periodic trigonometric interpolation suggests that in order to prove a general convergence result for continuous 2π-periodic functions, it is enough to consider the linear operator of the form

$$(5.1)\quad L_n(f;x) := \sum_{k=0}^{n-1} [f(x_{2k})\varrho_1(x-x_{2k}) + \delta_{0,m_{p+1}} f(x_{2k+1})\varrho_{m_{p+1}}(x-x_{2k})]$$

which has the following properties:

$$\begin{matrix} L_n^{(m_\nu)}(f;x_{2k}) = \delta_{\nu 1} f(x_{2k}) \quad (\nu=1,\dots,p) \\ L_n^{(m_\nu)}(f;x_{2k+1}) = \delta_{\nu,p+1} f(x_{2k+1})\delta_{0,m_{p+1}} \ (\nu=1,\dots,p+q) \end{matrix} \ (k=0,1,\dots,n-1).$$

The convergence properties of this operator depend on the order of magnitude of the fundamental functions $\varrho_\nu(x)$ determined in the previous sections. We first prove a lemma on the fundamental polynomials.

LEMMA 1. *Under the conditions of Theorem A in case (a) and (c) with the additional assumption that* $m_{p+1} > 0$, *we have*

$$(5.2)\qquad \Big\| \sum_{k=0}^{2n-1} |\varrho_\nu(x-x_k)| \Big\| = O(n^{-m_\nu}\log n), \quad \nu = 1\dots,p+q.$$

(Here and in what follows $\|\cdot\|$ means the sup norm.)

PROOF. Assume that m_ν is even. (The proof for the case m_ν odd is similar.) Then $\varrho_\nu(x)$ is given by the form (2.11) or (4.7). Since the coefficients $a_{\lambda j}(\nu)$ of this polynomial are obtained as the ratio of two determinants where the determinant in the denominator is a non-vanishing function of the variable α in the closed interval $[-\frac{1}{2},\frac{1}{2}]$, we have

$$(5.3)\qquad |\alpha_{\lambda j}(\nu)| = O(1).$$

Hence separating terms corresponding to $j=0$, we can write $\varrho_\nu(x)$ in the form

$$(5.4)\qquad \varrho_\nu(x) = \frac{(-1)^{m_\nu/2}}{n^{1+m_\nu}} \Big[2\sum_{\lambda=0}^{s}{}' \sum_{j=1}^{n-1} a_{s-\lambda,j}(\nu)\cos(\lambda n - j)x + O(1)\Big],$$

where $\sum'$ indicates that when $\lambda=0$, the factor 2 should be dropped. For a fixed λ, all the coefficients $a_{s-\lambda,j}(\nu)$ are determined by the same formula (e.g. (2.6) in Section 2 and Section 4). Therefore

$$\Delta a_{s-\lambda,j}(\nu) := a_{s-\lambda,j}(\nu) - a_{s-\lambda,j+1}(\nu) = O(n^{-1}),$$

$$\text{where} \begin{cases} j=1,\dots,r-1, & \text{if } \lambda = 0 \\ j=1,\dots,n-2, & \text{if } \lambda = 1,\dots,s. \end{cases}$$

Using Abel-transform on (5.4), we obtain

$$\sum_{j=1}^{n-1} a_{s-\lambda,j}(\nu)\cos(\lambda n+r-j)x=$$

$$=\sum_{j=1}^{n-2}\Delta a_{s-\lambda,j}(\nu)\sum_{l=1}^{j}\cos(\lambda n+r-l)x+a_{s-\lambda,n-1}(\nu)\sum_{l=1}^{n-1}\cos(\lambda n+r-l)x=$$

$$=O(n^{-1})\sum_{j=1}^{n-2}\left|\frac{\sin\frac{j}{2}x\sin(\lambda n+r-\frac{j+1}{2})x}{\sin\frac{x}{2}}\right|+O(1)\left|\frac{\sin\frac{n-1}{2}x\sin(\lambda n+r-\frac{n}{2})x}{\sin\frac{x}{2}}\right|,$$

whence again from (5.4), we have

$$\sum_{k=0}^{2n-1}|\varrho_\nu(x-x_k)|=$$

$$=O(n^{-1-m_\nu})\sum_{\lambda=0}^{s}\left\{O(n^{-1})\sum_{j=1}^{n-2}\sum_{k=0}^{2n-1}\left|\frac{\sin\frac{j}{2}(x-x_k)}{\sin\frac{x-x_k}{2}}\right|+\right.$$

$$\left.+O(1)\sum_{k=1}^{2n-1}\left|\frac{\sin\frac{n-1}{2}(x-x_k)}{\sin\frac{x-x_k}{2}}\right|\right\}+O(n^{-m_\nu})=$$

$$=O(n^{-1-m_\nu})\{O(n^{-1})O(n)O(n\log n)+O(1)O(n\log n)\}+O(n^{-m_\nu})=$$

$$=O(n^{-m_\nu}\log n).$$

We can now state our convergence theorem.

THEOREM 1. *Under the conditions of Theorem A in cases (a) and (c), we have*

$$\|f(x)-L_n(f,x)\|=O(E_m(f)\log n)+ +O\Big(\frac{\log n}{n^\mu}\Big)\sum_{k=0}^{n}(k+1)^{\mu-1}E_k(f), \tag{5.5}$$

where

$$\begin{cases}\mu=\begin{cases}\min(m_2,m_{p+2}) & \text{if } m_{p+1}=0,\\ \min(m_2,m_{p+1}) & \text{if } m_{p+1}>0,\end{cases}\\ m=\dfrac{n}{(\log n)1/\mu},\end{cases} \tag{5.6}$$

and $E_k(f)$ is the error of best trigonometric approximation of order k to $f(x)$.

PROOF. Let $p_m(x)$ be the trigonometric polynomial of best approximation of order m to $f(x)$. Since in the cases concerned the problem of 2-periodic interpolation is regular, we evidently have

$$p_m(f,x)=L_n(p_m,x)+\sum_{k=0}^{n-1}\Big\{\sum_{\nu=2}^{p}p_m^{(m_\nu)}(x_{2k})\varrho_\nu(x-x_{2k})+ \tag{5.7}$$
$$+\sum_{\nu=p+1}^{p+q}{}'p_m^{(m_\nu)}(x_{2k+1}\varrho_\nu(x-x_{2k})\Big\}$$

where and in what follows the prime on the summation indicates that the term corresponding to $\nu=p+1$ should be omitted if $m_{p+1}=0$.

According to Lemma 2 [5], we have

$$\|p_m^{(j)}(x)\|=O\Big(\sum_{k=0}^{m}(k+1)^{j-1}E_k(f)\Big).$$

Thus from Lemma 1 and the definitions of m and μ in (5.6) we see from (5.7) that

$$\|p_m-L_n(p_m,x)\|=O\Big(\sum_{\nu=2}^{p+q}{}'\|p_m^{(m_\nu)}\|\|\sum_{k=0}^{2n-1}\varrho_\nu(x-x_k)\|\Big)=$$
$$=O\Big(\sum_{\nu=2}^{p+q}{}'\frac{\log n}{n^{m_\nu}}\sum_{k=0}^{m}(k+1)^{m_\nu-1}E_k(f)\Big)=$$
$$=O\Big(\sum_{\nu=2}^{p+q}{}'\frac{\log n}{n^{m_\nu}}m^{m_\nu-\mu}\sum_{k=0}^{m}(k+1)^{\mu-1}E_k(f)\Big)=$$
$$=O\Big(\sum_{\nu=2}^{p+q}{}'\frac{(\log n)^{2-\frac{m_\nu}{\mu}}}{n^{\mu}}\sum_{k=0}^{m}(k+1)^{\mu-1}E_k(f)\Big)=$$
$$=O\Big(\frac{\log n}{n^{\mu}}\sum_{k=0}^{m}(k+1)^{\mu-1}E_k(f)\Big).$$

Hence it follows that

$$\|f-L_n(f)\|\leqq\|f-p_m\|+\|p_m-L_n(p_m)\|+\|L_n(p_m-f)\|\leqq$$
$$\leqq E_m(f)+O\Big(\frac{\log n}{n^{\mu}}\sum_{k=0}^{m}(k+1)^{\mu+1}E_k(f)\Big)+O(E_m(f)\log n). \qquad \square$$

6. Convergence (continued)

Theorem 1 shows that in order to have convergence in cases (a) and (c) of Theorem A, we have to assume $E_n(f) = o(\frac{1}{\log n})$. However, in some cases, this condition can be dropped. We shall now turn to these cases.

LEMMA 2. *Under the conditions of Theorem A (in cases III, V, and VI), with n even and $m_{p+1} > 0$, we have*

$$(6.1)\qquad \Big\| \sum_{k=0}^{2n-1} |\varrho_1(x - x_k)| \Big\| = O(1), \quad \Big\| \sum_{k=0}^{2n-1} |\varrho_\nu(x - x_k)| \Big\| = O\Big(\frac{\log n}{n^{m_\nu}}\Big)$$

$$(\nu = 2, \dots, n-1).$$

PROOF. Set $\alpha_j = \frac{j}{n}$ $(j = 0, 1, \dots, n-1)$ and

$$(6.2)\qquad a_{\lambda j}(\nu) := \frac{D_{\nu,\lambda+1}(\alpha_j)}{D(\alpha_j)}, \qquad \begin{matrix} 0 \leqq \alpha_j \leqq 1; \ \nu = 1, \dots, p+q \\ \lambda = 0, 1, \dots, 2s+1. \end{matrix}$$

Then from (3.11) we get for $\nu = 1$,

$$(6.3)\quad \varrho_1(x) = \frac{1}{n}\Big[a_{s,n-1}(1) + 2\sum_{\lambda=0}^{s}\sum_{j=0}^{n-1} a_{\lambda j}(1)\cos((s+1-\lambda)n - j)x\Big] + O\Big(\frac{1}{n}\Big).$$

Since $a_{\lambda j}(\nu)$ $(\lambda = 0, 1, \dots, 2s+1)$ is an analytic function of α_j in a domain containing the interval $[0, 1]$ and since $\alpha_{j+1} - \alpha_{j-1} = O(\frac{1}{n})$, it follows easily by the mean-value theorem that

$$(6.4)\qquad \begin{gathered} \Delta^2 a_{\lambda j}(\nu) := a_{\lambda,j-1}(\nu) - 2a_{\lambda,j}(\nu) + a_{\lambda,j+1}(\nu) + O\Big(\frac{1}{n^2}\Big) \\ \nu = 1, \dots, p+q; \quad \lambda = 0, 1, \dots, 2s+1; \\ j = 1, \dots, n-2. \end{gathered}$$

From (3.4) we see that

$$(6.5)\qquad \tilde{D}_{1,\lambda+1}(1) = 0, \quad \lambda \neq s$$

since the cofactor of $(1, \lambda+1)$ term in $\tilde{D}(1)$ will contain a zero column (here we use the fact that $m_{p+1} > 0$). This together with (3.11) yields

$$(6.6)\qquad \tilde{D}_{1,\lambda+2}(0) = \tilde{D}_{1,2s+1-\lambda}(1) = 0, \quad \lambda \neq s.$$

Therefore we have

$$
\begin{aligned}
(6.7)\qquad &|a_{\lambda,n-2}(1)-2a_{\lambda,n-1}(1)+a_{\lambda+1,0}(1)|\leqq\\
&\leqq|a_{\lambda,n-2}(1)-a_{\lambda,n-1}(1)|+|a_{\lambda,n-1}(1)-a_{\lambda+1,0}(1)|=\\
&=O\Big(\frac{1}{n}\Big)+\Big|\frac{\tilde D_{1,\lambda+1}(\alpha_{n-1})}{\tilde D(\alpha_{n-1})}-\frac{\tilde D_{1,\lambda+2}(0)}{\tilde D(0)}\Big|=\\
&=O\Big(\frac{1}{n}\Big)+\Big|\frac{\tilde D_{1,\lambda+1}(\alpha_{n-1})}{\tilde D(\alpha_{n-1})}\Big|\quad\text{on using (6.6)}\\
&=O\Big(\frac{1}{n}\Big),\quad \lambda\neq s.
\end{aligned}
$$

Similarly, we have

$$
(6.8)\qquad |a_{\lambda,n-1}(1)-2a_{\lambda+1,0}(1)+a_{\lambda+1,n}(1)|=O\Big(\frac{1}{n}\Big),\quad \lambda\neq s.
$$

Thus if we write

$$
(6.9)\qquad \varrho_1(x)=\frac{1}{n}\Big[\beta_0+2\sum_{j=1}^{(s+1)n}\beta_j\cos jx\Big]+O\Big(\frac{1}{n}\Big),
$$

(notice that $\beta_{(s+1)n}=a_{0,0}(1)=\frac{\tilde D_{1,1}(0)}{\tilde D(0)}=0$ from (3.5)), then we see from (6.7) and (6.8) that

$$
(6.10)\qquad \Delta^2\beta_j=\begin{cases} O\Big(\dfrac{1}{n^2}\Big) & \text{if } j \text{ or } j+1 \text{ are multiples of } n\\ O\Big(\dfrac{1}{n}\Big) & \text{otherwise.}\end{cases}
$$

We now apply a double Abel summation to (6.9) to obtain

$$
\begin{aligned}
(6.11)\qquad \varrho_1(x)=\frac{1}{n}\Big[&\sum_{j=1}^{(s+1)n-2}\Delta^2\beta_j\Big(\frac{\sin\frac{jx}{2}}{\sin\frac{x}{2}}\Big)^2+\\
&+\Delta\beta_{(s+1)n-1}\Big(\frac{\sin\frac{(s+1)n-2}{2}x}{\sin\frac{x}{2}}\Big)^2\Big]+O\Big(\frac{1}{n}\Big).
\end{aligned}
$$

Since

$$
\sum_{k=0}^{2n-1}\Big(\frac{\sin\frac{j(x-x_k)}{2}}{\sin\frac{x-x_k}{2}}\Big)^2=2nj,\quad j=1,2,\dots
$$

we see from (6.10) and (6.11) that

$$\sum_{k=0}^{2n-1} |\varrho_1(x-x_k)| = \frac{1}{n}\Big[2n \sum_{j=1}^{(s+1)n-2} j\Delta^2\beta_j + \Delta\beta_{(s+1)n-1} O(n^2)\Big] + O(1) = \\ = O(1). \tag{6.12}$$

Now assume that $2 \leqq \nu \leqq p+q$ and that m_ν is odd (the case where m_ν is even is similar). Again, from (3.12) we have

$$|\varrho_\nu(x)| \leqq n^{-1-m_\nu}\Big|\sum_{\lambda=0}^{s}\sum_{j=0}^{n-1} a_{\lambda j}(\nu)\sin((s+1-\lambda)n-j)x\Big| + O(n^{-1-m_\nu}) =$$

$$= n^{-1-m_\nu}\Big|\sum_{j=1}^{(s+1)n-1} \gamma_j(\nu)\sin jx\Big| + O(n^{-1-m_\nu})$$

where, as in (6.7), we have

$$\Delta_{\gamma j} := \gamma_{j+1} - \gamma_j = \begin{cases} O\left(\frac{1}{n}\right) & \text{if } j \text{ or } j+1 \text{ are not multiples of } n, \\ O(1) & \text{otherwise.} \end{cases}$$

Using Abel transform once, we have

$$|\varrho_\nu(x)| \leqq n^{-1-m_\nu} \sum_{j=1}^{(s+1)n-1} |\Delta\gamma_j| \left|\frac{\sin jx \sin(j+1)x}{\sin\frac{x}{2}}\right| +$$

$$+ |\gamma_{s+1)n}| \left|\frac{\sin(s+1)nx\sin(s+2)nx}{\sin\frac{x}{2}}\right|.$$

Since

$$\sum_{k=1}^{2n-1} \left|\frac{\sin j(x-x_k)\sin(j+1)(x-x_k)}{\sin\frac{x-x_k}{2}}\right| = O(n\log n),$$

we obtain

$$\sum_{k=0}^{2n-1} |\varrho(x-x_k)| = O(n^{-1-m_\nu})O(n\log n) \sum_{j=1}^{(s+1)n-1} |\Delta(\gamma_j)| = \\ = O(\frac{\log n}{n^{m_\nu}}). \tag{6.13}$$

The result follows from (6.12) and (6.13).

Now we are able to prove our main result on the convergence of some 2-periodic interpolation operators.

THEOREM 2. *Under the conditions of Lemma 2 above, if we set*

$$L_n(f,x) := \sum_{k=0}^{n-1} f(x_{2k})\varrho_1(x - x_{2k}), \tag{6.14}$$

then we have

$$\|f(x) - L_n(f,x)\| = O\Big(\frac{1}{m^\mu}\Big) \sum_{k=0}^{m} (k+1)^{\mu-1} E_k(f), \tag{6.15}$$

where

$$\mu = \min(m_2, m_{p+1}), \quad m = \Big[\frac{n}{(\log n)^{1/\mu}}\Big]$$

for all continuous functions $f(x)$, where $E_k(f)$ is the best trigonometric approximation of order k to $f(x)$.

REMARK. (6.15) implies that $\lim_{n\to\infty} \|f(x) - L_n(f,x)\| = 0$ and, in particular, if $f(x) \in \mathrm{Lip}\ \alpha$, then

$$\|f(x) - L_n(f;x)\| = O\Big(\frac{(\log n)^{\alpha/\mu}}{n^\mu}\Big).$$

The proof is the same as that of Theorem 1, but with a reference to Lemma 2 instead of Lemma 1.

In connection with Theorem A case (a), we do not have a general convergence theorem similar to that of case (c), because condition (6.6) is not always guaranteed. Nevertheless, in some special cases, it is still possible as in the following example.

EXAMPLE. Let $p = 2$, $q = 1$ in the case of $(0, m_2; m_3)$ interpolation where m_2 is odd and $m_3 \geqq 2$ is even. From the general formula (2.12), we obtain

$$\varrho_1(x) = \frac{1}{n}\Big[a_{1,r}(1) + 2\sum_{j=1}^{r} a_{1,r-j}(1)\cos jx + 2\sum_{j=0}^{n-1} a_{0,j}(1)\cos(n+r-j)x\Big].$$

From (2.5), we see that

$$D_{1,1}\Big(\frac{1}{2}\Big) = \frac{3^{m_3} + 3^{m_2}}{2^{m_1+m_2}} = D_{1,2}\Big(-\frac{1}{2}\Big)$$

so that $a_{1,0}(1) = a_{0,n-1}(0)$ and the method of proof of Lemma 2 works, yielding the following estimate for the operator (5.1):

$$\|f(x) - L_n(f,x)\| = O\Big(\frac{\log n}{n^\mu}\Big) \sum_{k=0}^{m} (k+1)^{\mu-1} E_k(f),$$

where $\mu = \min(m_2, m_3)$ and $m = [\frac{n}{(\log n)^\mu}]$.

The condition $m_3 > 0$ in this example cannot be dropped, as the example of (0,1;0) shows. Here the fundamental polynomials are easy to calculate. Indeed we have

$$\varrho_1(x) = \frac{\sin nx \sin\frac{nx}{2} \cos\frac{x}{2}}{2n^2 \sin^2\frac{x}{2}}, \quad \varrho_2(x) = \frac{\sin nx \sin\frac{nx}{2}}{n \sin\frac{x}{2}},$$

$$\varrho_3(x) = -\frac{\sin^2\frac{nx}{2} \cos\frac{nx}{2}}{n \sin\frac{x-x_1}{2}}.$$

We see that

$$\Big\| \sum_{k=0}^{n-1} |\varrho_1(x - x_{2k})| \Big\| = O(1), \quad \Big\| \sum_{k=0}^{n-1} |\varrho_2(x - x_{2k})| \Big\| = O\Big(\frac{\log n}{n}\Big)$$

but

$$\begin{aligned}\Big\| \sum_{k=0}^{n-1} |\varrho_3(x - x_{2k})| \Big\| &\geqq \sum_{k=0}^{n-1} \Big|\varrho_3\big(\frac{\pi}{2n} - x_{2k}\big)\Big| \\ &\geqq \frac{1}{n} \sin^2\frac{\pi}{4} \cos\frac{\pi}{4} \sum_{k=0}^{n-1} \frac{1}{\sin\frac{4k+1}{2n}\pi} \\ &\geqq c \log n.\end{aligned}$$

Thus in the case of (0,1;0), the interpolant

$$L_n(f,x) := \sum_{k=0}^{n-1} f(x_{2k}) \varrho_1(x - x_{2k}) + \sum_{k=0}^{n-1} f(x_{2k+1}) \varrho_3(x - x_{2k})$$

cannot converge uniformly for all continuous 2π-periodic functions.

7. Conclusion

Unfortunately, we do not have estimates for the fundamental polynomials in cases III to VI in Theorem A when n is odd.

Finally, we mention that there is an error in the formulation of Theorem 3 in [4]. In the notation of the present paper there we considered the problem of $(0; m_1)$-interpolation with $p = 1$, $q = 1$. It was proved there that if m_1 is odd, then the problem of $(0; m_1)$-interpolation on the nodes $x_k = \frac{k\pi}{n}, k = 0, 1, \ldots, 2n-1$ is regular if and only if n is odd and $\varepsilon = 1$. Furthermore if m_1 is even, then the necessary and sufficient condition for regularity is $\varepsilon = 0$ and where n could be even or odd. The correct statement of Theorem 3 in [4] is then as follows:

Let m_1, n and ε satisfy either of the conditions listed above. Then for any $f(x) \in C_{2\pi}$, we have

$$\Big\| f(x) - \sum_{j=0}^{n-1} f(x_{2j}) \varrho_1(x - x_{2j}) \Big\| =$$

$$= O\Big(n^{\frac{1-(-1)^{m_1}}{2}} E_{[m/4]}(f) + n^{-m_1} \sum_{k=0}^{m} (k+1)^{m_1-1} E_k(f) \Big).$$

We mention that this statement when m_1 is odd is a special case of Theorem A case IV, i.e., in this case the relation

$$\Big\| \sum_{k=0}^{2n-1} |\varrho_1(x - x_k)| \Big\| = O(n)$$

is proved. We suspect that this latter relation holds in all cases IV to VI (n odd).

REFERENCES

[1] Lorentz, G. G., Jetter, K. and Riemenschneider, S. D., *Birkhoff interpolation*, Encyclopedia of Mathematics and its Applications, **19** (Chapter 2, Sec. 2.4 and Chapter 11), Addison-Wesley Publ. Co., Reading, Mass., 1983. *MR* **84g**:41002

[2] Sharma, A., Szabados, J. and Varga, R. S., Some 2-periodic trigonometric interpolation problems on equidistant nodes, *Analysis* **11** (1991), 165–190.

[3] Sharma, A. and Varga, R. S., On a particular 2-periodic lacunary trigonometric interpolation problem on equidistant nodes, *Resultate Math.* **16** (1989), 383–404. *MR* **91c**:42005

[4] Sharma, A., Szabados, J. and Varga, R. S., 2-periodic lacunary trigonometric interpolation: the $(0; M)$ case, *Constructive Theory of Functions '87* (Varna, 1987), Publishing House of the Bulgarian Academy of Sciences, Sofia, 1988, 420–427. *MR* **90e**:42010

[5] Szabados, J., On the rate of convergence of a lacunary trigonometric interpolation process, *Acta Math. Hungar.* **47** (1986), 361–370. *MR* **87k**:42006

(Received June 14, 1991)

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ALBERTA
EDMONTON, ALBERTA
T6G 2G1
CANADA

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

DEPARTMENT OF MATHEMATICS
KENT STATE UNIVERSITY
KENT, OH 44242
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 433–435

A NOTE ON WEAK SYMMETRY PROPERTIES OF QUASI-UNIFORMITIES

J. DEÁK

Abstract

Two notions of weak quasi-uniform symmetry can be replaced by a common generalization when proving that certain spaces are quiet.

Recall that a quasi-uniform space $(X,\mathcal{U})$ is *semi-symmetric* [2] provided that if C is closed, H is open, and $U^{-1}[C]\subset H$ for some $U\in\mathcal{U}$ then there is a $V\in\mathcal{U}$ such that $V[C]\subset H$; *open-symmetric* [6] provided that $A\,\delta_{\mathcal{U}}\,B$ iff $B\,\delta_{\mathcal{U}}\,A$ whenever A and B are open; *uniformly regular* [1] if for any $U\in\mathcal{U}$ there is a $V\in\mathcal{U}$ such that $\overline{Vx}\subset Ux$ $(x\in X)$; *quiet* [4] it for any $U\in\mathcal{U}$ there is a $V\in\mathcal{U}$ (called *quiet for* U) such that $Vx\in\mathfrak{g}$, $V^{-1}y\in\mathfrak{f}$ imply $x\,Uy$ whenever $(\mathfrak{f},\mathfrak{g})$ is a Cauchy filter pair (*Cauchy* means that for any $U\in\mathcal{U}$ there are $F\in\mathfrak{f}$ and $G\in\mathfrak{g}$ with $F\times G\subset U$). A simple rewording of the definition gives that $\mathcal{U}$ is semi-symmetric iff $A\,\delta_{\mathcal{U}}\,B$ is equivalent to $B\,\delta_{\mathcal{U}}\,A$ for closed sets A and B ([6] 3.3); therefore we change over to the more consistent terminology *closed-symmetric.*

Quiet quasi-uniformities are uniformly regular [5]. Conversely , closed-symmetric or open-symmetric uniformly regular spaces are quiet ([8] Theorem 4 and Proposition 8); we are going to show that essentially the same proof yields a more general result.

DEFINITION. A quasi-uniformity is *mixed-symmetric* provided that if H is open, C is closed and $H\,\delta_{\mathcal{U}}\,C$ then $C\,\delta_{\mathcal{U}}\,H$. □

Closed-symmetry implies mixed-symmetry: let H be open, C closed, $C\,\bar{\delta}_{\mathcal{U}}\,H$, and take $U\in\mathcal{U}$ such that $U^2[C]\cap H=\emptyset$; then $U[C]\cap\overline{H}=\emptyset$, thus $C\,\bar{\delta}_{\mathcal{U}}\,\overline{H}$, therefore $\overline{H}\,\bar{\delta}_{\mathcal{U}}\,C$, $H\,\bar{\delta}_{\mathcal{U}}\,C$. Similarly, open-symmetry also implies mixed-symmetry: if we assume in addition that Ux is open $(x\in X)$ then $U[C]$ is open and $U[C]\,\bar{\delta}_{\mathcal{U}}\,H$, thus $H\,\bar{\delta}_{\mathcal{U}}\,C$ again.

PROPOSITION. *Each mixed-symmetric uniformly regular quasi-uniformity is quiet.*

1991 *Mathematics Subject Classification.* Primary 54E15.

Key words and phrases. Quasi-uniformity, uniformly regular, quiet, open-symmetric, closed-symmetric = semi-symmetric, mixed-symmetric.

Research supported by Hungarian National Foundation for Scientific Research Grant No. 2114.

PROOF. Let $U \in \mathcal{U}$. Choose $U_0, V, W, Z \in \mathcal{U}$ such that $U_0^2 \subset U$, $\overline{Vx} \subset U_0 x$ $(x \in X)$, $W^2 \subset V$, $\overline{Zx} \subset Wx$ $(x \in X)$. We show that Z is quiet for U. Assume the contrary, and pick a Cauchy filter pair $(\mathfrak{f}, \mathfrak{g})$ and $x, y \in X$ such that $Zx \in \mathfrak{g}$, $Z^{-1}y \in \mathfrak{f}$, but $x\, U y$ does not hold. Then $U_0 x \cap U_0^{-1} y = \emptyset$. From $W[\overline{Zx}] \subset \overline{Vx}$ we have $\overline{Zx}\, \delta_{\mathcal{U}}\, X \setminus \overline{Vx}$, thus $X \setminus \overline{Vx}\, \overline{\delta}_{\mathcal{U}}\, \overline{Zx}$ by the mixed-symmetry. Now $X \setminus \overline{Vx} \supset X \setminus U_0 x \supset U_0^{-1} y \supset Z^{-1} y$, so $Z^{-1} y\, \overline{\delta}_{\mathcal{U}}\, Zx$. Hence there is a $Q \in \mathcal{U}$ with $Q[Z^{-1}y] \cap Zx = \emptyset$, a contradiction, since $Z^{-1}y$ is in the first and Zx in the second member of a Cauchy filter pair. □

There exist mixed-symmetric uniformly regular spaces that are neither open-symmetric nor closed-symmetric: take the disjoint sum of the following two spaces. (Other examples with the same properties were given in [6] 5.1 and 5.2.)

EXAMPLES. a) *Closed-symmetric, not open-symmetric.* On

$$X = \{(0, 1/n), (1/n, 1/n) : n \in \mathbf{N}\} \cup \{(0,0)\},$$

let d be the trace of the Sorgenfrey quasi-metric of $\mathbf{R}^2$, i.e.

$$d((x', x''), (y', y'')) = \begin{cases} \max\{y' - x', y'' - x''\} & \text{if } x' \leqq y', x'' \leqq y'', \\ 1 & \text{otherwise.} \end{cases}$$

$\mathcal{U} = \mathcal{U}(d)$ is uniformly regular. The sets

$$(1) \qquad \{(0, 1/n) : n \in \mathbf{N}\}, \quad \{(1/n, 1/n) : n \in \mathbf{N}\}$$

show that $\mathcal{U}$ is not open-symmetric. But $\mathcal{U}$ is closed-symmetric, because disjoint closed sets are always far (one of them is finite and does not contain (0,0); such a set is far from its complement in both directions).

b) *Open-symmetric, not closed-symmetric.* Let now

$$X = \{(0, 1/n) : n \in \mathbf{N}\} \cup \{(1/k, 1/n) : k, n \in \mathbf{N}, k \geqq n\},$$

and define $\mathcal{U} = \mathcal{U}(d)$ just as above. $\mathcal{U}$ is again uniformly regular. The sets in (1) show that $\mathcal{U}$ is not closed-symmetric. We are going to check that $\mathcal{U}$ is open-symmetric. Let A and B be disjoint open sets, $A\, \delta_{\mathcal{U}}\, B$. For $i \in \mathbf{N}$, choose $x_i = (x_i', x_i'') \in A$ and $y_i = (y_i', y_i'') \in B$ such that $d(x_i, y_i) < 1/i$. Assume first that $x_i'' \to 0$; then $y_i'' \to 0$, too. For $\varepsilon > 0$ fixed, take $j \in \mathbf{N}$ such that $x_j'' < \varepsilon$. As A is open, we may assume that $x_j' \neq 0$. Pick $i \in \mathbf{N}$ such that $y_i'' < x_j'$; then $0 \leqq y_i' \leqq y_i'' < x_j' \leqq x_j'' < \varepsilon$, so $d(y_i, x_j) < \varepsilon$. Hence $B\, \delta_{\mathcal{U}}\, A$. If $x_i'' \not\to 0$ then choose $n \in \mathbf{N}$ such that $x_i'' = 1/n$ $(i \in I)$ with an infinite $I \subset \mathbf{N}$. Now $y_i'' = 1/n$ for $i \in I$ large enough, $x_i' \to 0$ $(i \in I)$, $y_i' \to 0$ $(i \in I)$, thus $(0, 1/n) \notin A$ (because A is open and $A \cap B = \emptyset$), i.e. $x_i' \neq 0$ $(i \in I)$, and clearly $B\, \delta_{\mathcal{U}}\, A$ again. □

REMARKS. a) The following modification of the definition of mixed-symmetry may also seem to be reasonable: if H is open, C is closed and $C\, \delta_{\mathcal{U}}\, H$

then $H\,\delta_{\mathcal{U}}\,C$. This definition, however, does not give anything new, since it is equivalent to the symmetry of $\delta_{\mathcal{U}}$. (The proof is the same as that of the mixed-symmetry of open-symmetric spaces.) Cf. also the first definition in [6] §2.

b) If $\mathcal{U}$ is mixed-symmetric and regular then it is locally symmetric: given $x \in X$ and $U \in \mathcal{U}$, take $V, V_0, W \in \mathcal{U}$ such that $\overline{Vx} \subset Ux$, $V_0^2 \subset V$, $\overline{Wx} \subset V_0 x$, $W \subset V_0$. Now $W[\overline{Wx}] \subset \overline{Vx}$, so $\overline{Wx}\,\overline{\delta}_{\mathcal{U}}\,X \setminus \overline{Vx}$, i.e. $X \setminus \overline{Vx}\,\overline{\delta}_{\mathcal{U}}\,\overline{Wx}$ by the mixed-symmetry, hence there is a $Z \in \mathcal{U}$ with $Z[X \setminus Ux] \cap Wx = \emptyset$; assuming also that $Z \subset W$, this implies $Z^{-1}[Zx] \subset Ux$.

c) Mixed-symmetry cannot be replaced by local symmetry in the proposition: [3] Example 1.3 b) is doubly uniformly regular, doubly locally symmetric (since both topologies are discrete), but not quiet.

d) A uniformly regular space can be open-symmetric as well as closed-symmetric without $\delta_{\mathcal{U}}$ being symmetric: take the trace on $\{0\} \cup \{1/n : n \in \mathbb{N}\}$ of the Sorgenfrey quasi-uniformity of $\mathbb{R}$. (Or, what is almost the same, $\mathcal{U}^{-1}$ from [8] Example 7.)

e) If $\mathcal{U}$ is mixed-symmetric and $\mathcal{U}^{-1}$ is point-symmetric then $\delta_{\mathcal{U}}$ is symmetric (a generalization of [6] 3.4 and 4.2, cf. [8] Lemma 4). Indeed, assume that $A\,\overline{\delta}_{\mathcal{U}}\,B$ and take $V \in \mathcal{U}$ with $V^3[A] \cap B = \emptyset$. According to [8] Lemma 4, the point-symmetry of $\mathcal{U}^{-1}$ is equivalent to the statement that $\overline{S} \subset V[S]$ whenever $S \subset X$ and $V \in \mathcal{U}$. Thus $V[\overline{A}] \subset V^2[A] \subset \overline{V^2[A]} \subset V^3[A] \subset X \setminus B$. From $\overline{A}\,\overline{\delta}_{\mathcal{U}}\,X \setminus \overline{V^2[A]}$ we obtain by the mixed symmetry that $X \setminus \overline{V^2[A]}\,\overline{\delta}_{\mathcal{U}}\,\overline{A}$, hence $B\,\overline{\delta}_{\mathcal{U}}\,A$.

REFERENCES

[1] Császár, Á., Extensions of quasi-uniformities, *Acta Math. Acad. Sci. Hungar.* **27** (1981), No. 1–3, 121–145. *MR* **82f**:54039

[2] Deák, J., On the coincidence of some notions of quasi-uniform completeness defined by filter pairs, *Studia Sci. Math. Hungar.* **26** (1991), 411–413.

[3] Deák, J., Extending and completing quiet quasi-uniformities, *Studia Sci. Math. Hungar.* **29** (1994), 351–364.

[4] Doitchinov, D., On completeness of quasi-uniform spaces, *C. R. Acad. Bulgar. Sci.* **41** (1988), No. 7, 5–8. *MR* **89j**:54028

[5] Fletcher, P. and Hunsaker, W., Uniformly regular quasi-uniformities, *Topology Appl.* **37** (1990), 285–291.

[6] Fletcher, P. and Hunsaker, W., Symmetry conditions in terms of open sets, *Topology Appl.* **45** (1992), 39–47.

[7] Fletcher, P. and Lindgren, W. F., *Quasi-uniform spaces*, Lecture Notes in Pure Appl. Math. **77**, Marcel Dekker, New York, 1982. *MR* **84h**:54026

[8] Künzi, H. P. A., Mrševič, M., Reilly, R. and Vamanamurthy, M. K., Convergence, precompactness and symmetry in quasi-uniform spaces, *Math. Japon.* **38** (1993), 239–253.

(Received June 25, 1991)

MTA MATEMATIKAI KUTATÓINTÉZETE
POSTAFIÓK 127
H–1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 437–442

ON SEQUENCES OF ZEROS AND ONES

P. KISS and B. ZAY

Let k and n be positive integers. In this paper we investigate the binary sequences of length n consisting of zeros and ones such that there are at least $k-1$ zeros between any two ones. We denote by $F_k(n)$ the number of such sequences, and by $w_k(n)$ the total number of the ones in these sequences.

The ratio

$$G_k(n)=\frac{w_k(n)}{nF_k(n)} \tag{1}$$

shows the mean value of the number of ones in a sequence of length n. Let

$$G_k=\lim_{n\to\infty} G_k(n) \tag{2}$$

if the limit exists. It is easy to see that limit (2) exists for $k=1$, and $G_1=\frac{1}{2}$. Furthermore P. H. St. John [3] proved that $G_2=\frac{5-\sqrt{5}}{10}$.

In this paper we show that limit (2) exists for any positive integer k. We prove:

THEOREM. *For any positive integer k, the sequence $G_k(n)$, $n=1,2,\dots,$ is convergent and*

$$G_k=\lim_{n\to\infty} G_k(n)=\frac{1}{\alpha_1^k+k-1}, \tag{3}$$

where α_1 is the greatest positive root of the polynomial $f(x)=x^k-x^{k-1}-1$.

PROOF. We may suppose that $k\geqq 2$ since in the case $k=1$ we have $\alpha_1=2$ and so by (3) also $G_1=1/2$ follows as we have seen above.

If there are at least 2 ones in a sequence then, by the conditions, $n\geqq k+1$. From this, using the definitions of $F_k(n)$ and $w_k(n)$,

$$F_k(n)=n+1 \quad \text{for } n=1,2,\dots,k \tag{4}$$

1991 *Mathematics Subject Classification.* Primary 11B83; Secondary 11B37.
Key words and phrases. Sequence, linear recurrence, length of sequences.

Research (partially) supported by Hungarian National Foundation for Scientific Research Grant No. 1641.

and

$$(5)\qquad w_k(n)=n \quad \text{for } n=1,2,\dots,k$$

follow. If $n>k$, then a sequence of length n can be constructed by writing a zero at the end of sequences of length $n-1$ or by writing $k-1$ zeros and a one at the end of sequences of length $n-k$. From this we obtain that

$$(6)\qquad F_k(n)=F_k(n-1)+F_k(n-k)$$

and

$$(7)\qquad w_k(n)=w_k(n-1)+w_k(n-k)+F_k(n-k)$$

for $n>k$. (6) shows that the sequence $F_k(n)$, $n=1,2,\dots$, is a linear recurrence of order k with characteristic polynomial

$$f(x)=x^k-x^{k-1}-1.$$

But by (6) and (7)

$$\begin{aligned} &w_k(n)-w_k(n-1)-w_k(n-k)=\\ &=w_k(n-1)+w_k(n-k)+F_k(n-k)-\\ &\quad -w_k(n-2)-w_k(n-k-1)-F_k(n-k-1)-w_k(n-k-1)-\\ &\quad -w_k(n-2k)-F_k(n-2k)=w_k(n-1)+w_k(n-k)-w_k(n-2)-\\ &\quad -w_k(n-k-1)-w_k(n-k-1)-w_k(n-2k)\end{aligned}$$

and so

$$w_k(n)=2w_k(n-1)-w_k(n-2)+2w_k(n-k)-2w_k(n-k-1)-w_k(n-2k)$$

for $n>2k$. It shows that the sequence $w_k(n)$, $n=1,2,\dots$, is also a linear recurrence of order $2k$ with characteristic polynomial

$$g(x)=x^{2k}-2x^{2k-1}+x^{2k-2}-2x^k+2x^{k-1}+1=(x^k-x^{k-1}-1)^2=f^2(x).$$

The sequences $F_k(n)$ and $w_k(n)$ can be defined also for $n=0,-1,-2,\dots$ by the formulas

$$(8)\qquad F_k(n-k)=F_k(n)-F_k(n-1)$$

and

$$(9)\qquad w_k(n-k)=w_k(n)-w_k(n-1)-F_k(n-k)$$

using the initial terms given in (4) and (5). By (4) and (8) we get

$$(10)\qquad F_k(n)=\begin{cases} 1 & \text{if } 1-k\leqq n\leqq 0,\\ 0 & \text{if } 2-2k\leqq n\leqq -k,\\ 1 & \text{if } n=1-2k,\end{cases}$$

and similarly, by (5), (9) and (10),

$$w_k(n) = \begin{cases} 0 & \text{if } 2-2k \leqq n \leqq 0, \\ -1 & \text{if } n = 1-2k \end{cases} \tag{11}$$

follows.

Let $S(n)$, $n = 0, 1, 2, \ldots$, be a k-th order linear recursive sequence defined by the characteristic polynomial $f(x)$ and by the initial terms

$$S(n) = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{if } 1 \leqq n \leqq k-1. \end{cases} \tag{12}$$

Similarly let $R(n)$, $n = 0, 1, 2, \ldots$, be a $2k$-th order linear recurrence defined by the characteristic polynomial $g(x) = f^2(x)$ and by the initial terms

$$R(n) = \begin{cases} -1 & \text{if } n = 0, \\ 0 & \text{if } 1 \leqq n \leqq 2k-1. \end{cases} \tag{13}$$

The sequences $F_k(n)$, $S(n)$ and $w_k(n)$, $R(n)$ have the same characteristic polynomial, respectively, and from (10), (12), and (11), (13)

$$F_k(n) = S(n+2k-1)$$

and

$$w_k(n) = R(n+2k-1)$$

follow. By (1) these imply that if the limit in (2) exists then

$$G_k = \lim_{n\to\infty} \frac{R(n)}{nS(n)}. \tag{14}$$

Denote by $\alpha_1, \alpha_2, \ldots, \alpha_k$ the roots of the polynomial $f(x) = x^k - x^{k-1} - 1$. H. R. P. Ferguson [1] and later V. E. Hoggatt, Jr. and K. Alladi [2] proved that these roots are distinct and there is a positive root among them which is maximal in absolute value, thus we may assume that

$$|\alpha_i| < \alpha_1 \text{ for } i = 2, 3, \ldots, k. \tag{15}$$

The roots of the characteristic polynomial of the sequence $R(n)$ are also $\alpha_1, \alpha_2, \ldots, \alpha_k$ with multiplicities two. Thus, using the known explicite form of the terms of linear recursive sequences, the terms of our sequences can be expressed by

$$R(n) = \sum_{i=1}^{k} (a_i + a_{i+k} n) \alpha_i^n \tag{16}$$

and

$$S(n)=\sum_{i=1}^{k} b_i\alpha_i^n \tag{17}$$

for any $n \geqq 0$, where the coefficients a_j $(1 \leqq j \leqq 2k)$ and b_i $(1 \leqq i \leqq k)$ depend on α_i's and on the initial terms of the sequences. By (16) and (17) we have

$$\frac{R(n)}{nS(n)}=\frac{\sum_{i=1}^{k}(a_i+a_{i+k}n)\alpha_i^n}{n\sum_{i=1}^{k} b_i\alpha_i^n}=\frac{\sum_{i=1}^{k}\left(\frac{a_i}{n}+a_{i+k}\right)(\alpha_i/\alpha_1)^n}{\sum_{i=1}^{k} b_i(\alpha_i/\alpha_1)^n}.$$

But $a_i/n \to 0$ as $n\to\infty$ $(i=1,2,\dots,k)$ and, by (15), $(\alpha_i/\alpha_1)^n \to 0$ as $n\to\infty$ $(i=2,3,\dots,k)$ and so by (14)

$$G_k=\lim_{n\to\infty}\frac{R(n)}{nS(n)}=\frac{a_{k+1}}{b_1}. \tag{18}$$

Let us consider the system of equations obtained from (17) by substituting $0,1,\dots,k-2$ and $k-1$, respectively. By (12), using Cramer's rule,

$$b_1=\frac{d_1}{d} \tag{19}$$

follows, where

$$d=\begin{vmatrix} 1 & 1 & \dots & 1\\ \alpha_1 & \alpha_2 & \dots & \alpha_k\\ \vdots & \vdots & & \vdots\\ \alpha_1^{k-1} & \alpha_2^{k-1} & \dots & \alpha_k^{k-1}\end{vmatrix} \quad\text{and}\quad d_1=\begin{vmatrix} 1 & 1 & \dots & 1\\ 0 & \alpha_2 & \dots & \alpha_k\\ \vdots & \vdots & & \vdots\\ 0 & \alpha_2^{k-1} & \dots & \alpha_k^{k-1}\end{vmatrix}.$$

d is a $k\times k$ Vandermonde determinant and so we have

$$d=\prod_{1\leqq j<i\leqq k}(\alpha_i-\alpha_j).$$

Similarly we can get

$$d_1=\left(\prod_{i=2}^{k}\alpha_i\right)\left(\prod_{2\leqq j<i\leqq k}(\alpha_i-\alpha_j)\right)$$

and by (19) we obtain

$$b_1=\prod_{i=2}^{k}\frac{\alpha_i}{\alpha_i-\alpha_1}. \tag{20}$$

Since the α_i's $(i = 1, 2, \dots, k)$ are the roots of the polynomial $f(x) = x^k - x^{k-1} - 1$ we have

$$\prod_{i=2}^{k} \alpha_i = \frac{\prod_{i=1}^{k} \alpha_i}{\alpha_1} = \frac{-(-1)^k}{\alpha_1} = \frac{(-1)^{k+1}}{\alpha_1}. \tag{21}$$

Let $P(x)$ be the polynomial defined by

$$P(x) = \prod_{i=2}^{k} (\alpha_i - x).$$

Then, using that $\alpha_1^k - \alpha_1^{k-1} = 1$, we get

$$\begin{aligned} P(x) &= (-1)^{k-1} \prod_{i=2}^{k} (x - \alpha_i) = (-1)^{k-1} \frac{x^k - x^{k-1} - 1}{x - \alpha_1} = \\ &= (-1)^{k-1} \frac{(x^k - \alpha_1^k) - (x^{k-1} - \alpha_1^{k-1})}{x - \alpha_1} = \\ &= (-1)^{k-1} (x^{k-1} + x^{k-2}\alpha_1 + \dots + \alpha_1^{k-1} - x^{k-2} - x^{k-3}\alpha_1 - \dots - \alpha_1^{k-2}). \end{aligned}$$

From this it follows that

$$\prod_{i=2}^{k} (\alpha_i - \alpha_1) = P(\alpha_1) = (-1)^{k-1} (k\alpha_1^{k-1} - (k-1)\alpha_1^{k-2})$$

and so, by (20) and (21),

$$\begin{aligned} b_1 &= \frac{1}{\alpha_1 (k\alpha_1^{k-1} - (k-1)\alpha_1^{k-2})} = \frac{1}{\alpha_1^k + (k-1)(\alpha_1^k - \alpha_1^{k-1})} = \\ &= \frac{1}{\alpha_1^k + k - 1}. \end{aligned} \tag{22}$$

If we consider the system of equations obtained from (16) by substituting $0, 1, \dots, 2k-2$ and $2k-1$, respectively, then similarly as above we get

$$a_{k+1} = \frac{D_1}{D} \tag{23}$$

where

$$D = \begin{vmatrix} 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \\ \alpha_1 & \alpha_2 & \dots & \alpha_k & \alpha_1 & \alpha_2 & \dots & \alpha_k \\ \alpha_1^2 & \alpha_2^2 & \dots & \alpha_k^2 & 2\alpha_1^2 & 2\alpha_2^2 & \dots & 2\alpha_k^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_1^{2k-1} & \alpha_2^{2k-1} & \dots & \alpha_k^{2k-1} & (2k-1)\alpha_1^{2k-1} & (2k-1)\alpha_2^{2k-1} & \dots & (2k-1)\alpha_k^{2k-1} \end{vmatrix}$$

and

$$D_1 = \begin{vmatrix} 1 & 1 & \dots & 1 & -1 & 0 & \dots & 0 \\ \alpha_1 & \alpha_2 & \dots & \alpha_k & 0 & \alpha_2 & \dots & \alpha_k \\ \alpha_1^2 & \alpha_2^2 & \dots & \alpha_k^2 & 0 & 2\alpha_2^2 & \dots & 2\alpha_k^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \alpha_1^{2k-1} & \alpha_2^{2k-1} & \dots & \alpha_k^{2k-1} & 0 & (2k-1)\alpha_2^{2k-1} & \dots & (2k-1)\alpha_k^{2k-1} \end{vmatrix}$$

since, by (13), the first $2k$ initial terms of the sequence $R(n)$ are $-1, 0, 0, \dots, 0$. Using similar arguments as at the treatment of Vandermonde determinants, after some elementary calculation we get

$$D = (-1)^{\frac{k(k-1)}{2}} \left(\prod_{i=1}^{k} \alpha_i\right) \left(\prod_{1 \leqq j < i \leqq k} (\alpha_i - \alpha_j)^4\right) \tag{24}$$

and

$$D_1 = (-1)^{\frac{k(k-1)}{2}} \left(\prod_{i=1}^{k} \alpha_i\right) \left(\prod_{i=2}^{k} \alpha_i^2 (\alpha_i - \alpha_1)^2\right) \left(\prod_{2 \leqq j < i \leqq k} (\alpha_i - \alpha_j)^4\right). \tag{25}$$

By (23), (24), (25), and (20)

$$a_{k+1} = \prod_{i=2}^{k} \frac{\alpha_i^2}{(\alpha_i - \alpha_1)^2} = b_1^2$$

follows. So by (18) and (22)

$$G_k = b_1 = \frac{1}{\alpha_1^k + k - 1}$$

which proves the theorem.

REFERENCES

[1] Ferguson, H. R. P., On a generalization of the Fibonacci numbers useful in memory allocation schema; or all about the zeroes of $z^k - z^{k-1} - 1$, $k > 0$, *Fibonacci Quart.* **14** (1976), 233-243. *MR* **54** #4082

[2] Hoggatt, Jr., V. E. and Alladi, K., Limiting ratios of convolved recursive sequences, *Fibonacci Quart.* **15** (1977), 211-214. *MR* **58** #486

[3] St. John, P. H., On the asymptotic proportions of zeros and ones in Fibonacci sequences, *Fibonacci Quart.* **22** (1984), 144-145. *MR* **85f**:11011

(Received May 30, 1991)

ESZTERHÁZY KÁROLY TANÁRKÉPZŐ FŐISKOLA
MATEMATIKAI TANSZÉK
LEÁNYKA U. 4
H–3300 EGER
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 29 (1994), 443–480

RANDOM INCREMENTS OF A WIENER PROCESS AND THEIR APPLICATIONS

QI-MAN SHAO

Abstract

Let $\{W(t), t \geqq 0\}$ be a Wiener process and $\{\tau_n, n \geqq 1\}$ be a sequence of stopping time with respect to $W(\cdot)$. This paper discusses how big are the random increments $W(\sum_{i=1}^{n+k} \tau_i) - W(\sum_{i=1}^{n} \tau_i)$, which makes it possible that the law of iterated logarithm type problem can be exchanged to one of strong law of large numbers type. As applications, the increments for martingale difference sequences as well as for independent random variables are obtained.

1. Introduction

Let $\{W(t), t \geqq 0\}$ be a standard Wiener process. It is well known that

$$\limsup_{T\to\infty} \frac{|W(t)|}{(2T\log\log T)^{1/2}} = 1 \quad \text{a.s.} \tag{1.1}$$

Csörgő and Révész [4] initiated the study of large increments of a Wiener process and showed that

$$\limsup_{T\to\infty} \sup_{0\leqq t\leqq T-a_T} \sup_{0\leqq s\leqq a_T} \frac{|W(t+s)-W(t)|}{(2a_T(\log(T/a_T)+\log\log T))^{1/2}} = 1 \quad \text{a.s.} \tag{1.2}$$

for a_T satisfying

(i) $0 < a_T \leqq T$;
(ii) a_T is non-decreasing;
(iii) T/a_T is non-decreasing.

1980 *Mathematics Subject Classification.* Primary 60F15, 60G16; Secondary 60G42, 60G50.

Key words and phrases. Random increments, Wiener process, martingale, partial sums.

Research supported by the Fok Yingtung Education Foundation and by National Natural Science Foundation of China.

Clearly, (1.1) is a special case of (1.2) with $a_T = T$. The (1.2) type increment is now called the Csörgő and Révész's increment. Motivated by a statistical problem, Hanson and Russo [7] presented another type increment, so-called lag increment. They proved

$$\limsup_{T\to\infty} \sup_{a_T \leqq t \leqq T} \frac{|W(T)-W(T-t)|}{(2t(\log(T/t)+\log\log t))^{1/2}} = 1 \quad \text{a.s.} \tag{1.3}$$

under suitable conditions on a_T. Since then, a great amount of work has been done on the increments of Wiener process as well as for partial sums of random variables (see, for example, [3], [1], [2], [13]). Csörgő and Révész [5] established how big the increments of partial sums of independent identically distributed random variables are via the strong approximation theorem. Applying the well-known exponential inequality and estimating the probability of partial sums directly, Lin [9], [10], [11] obtained the corresponding results on the increment for sums of independent non-identically distributed random variables. However, Lin's method seems helpless for dependent even for mixing sequences since it is usually difficult to establish an appropriate exponential inequality for dependent random variables. Noting that there is no strong approximation theorem for dependent random variables as sharp as that for iid random variables, one has to explore a new way. While dealing with the increment of independent random variables with the r-th moment generating function in [14], the author found that the well-known Skorohod embedding theorem is very useful although it was out of power to obtain a sharp result on strong approximation for independent random variables. This gives us an idea that we can use the Skorohod–Strassen martingale embedding theorem to discuss the increment of partial sums of dependent random variables. In order to investigate the increment of martingale difference sequences, independent random variables and mixing sequences, this paper aims at studying how big the random increments of a Wiener process are, which makes it possible that the law of the iterated logarithm type problem can be changed to the strong law of large numbers type problem.

We will unify the Csörgő and Révész's increment and lag increment and discuss how big are the random increments of a Wiener process in Section 2 and give applications to the martingale difference sequences and to the partial sums of independent not necessarily identically distributed random variables in Section 3 and Section 4, respectively.

Throughout this paper we will use the following notations: $\log x = \ln\max(x,e)$ for $x \in \mathbf{R}^1$, where ln is the natural logarithm; $[x]$ denotes the integer part of x; $\sum_{i=n}^{m} = 0$ if $m<n$; $a \vee b$ denotes $\max\{a,b\}$. Let $\{a_N, N \geqq 1\}$ be a sequence of positive integer numbers, $\{b_N, N \geqq 1\}$ and $\{c_N, N \geqq 1\}$ sequences of non-negative integer numbers, and $\{\sigma_i, i \geqq 1\}$ a sequence of non-negative

real numbers. Put

(1.4) $$\sigma_{n,k} = \sum_{i=n+1}^{n+k} \sigma_i,$$

(1.5) $$\beta_{n,k} = \{2\sigma_{n,k}(\log(\sigma_{0,n+k}/\sigma_{n,k}) + \log\log \sigma_{n,k})\}^{-1/2},$$

(1.6) $$\alpha_{n,N} = \{2\sigma_{n,a_N}(\log(\sigma_{0,a_N+b_N}/\sigma_{n,a_N}) + \log\log \sigma_{0,a_N+b_N})\}^{-1/2}.$$

2. How big are the random increments of a Wiener process

Let $\{W(t), \mathcal{F}_t, t \geqq 0\}$ be a standard Wiener process and $\{\tau_n, n \geqq 1\}$ be a sequence of stopping times with respect to $\mathcal{F}$. Let $\{a_N, N \geqq 1\}$ be a sequence of positive integer numbers, $\{b_N, N \geqq 1\}$ and $\{c_N, N \geqq 1\}$ sequences of non-negative integer numbers, and $\{\sigma_n, n \geqq 1\}$ a sequence of non-negative real numbers. Let $\sigma_{n,k}, \beta_{n,k}$ and $\alpha_{n,N}$ be defined as in (1.4), (1.5) and (1.6), respectively.

Our main results are as follows.

THEOREM 2.1. *Assume that*

(2.1) $$\max_{0 \leqq n \leqq b_N} \max_{a_N \leqq k \leqq a_N + c_N} \Big(\sum_{i=1+n}^{n+k} \tau_i \Big) / \sigma_{n,k} \leqq 1 \quad a.s. \text{ as } N \to \infty$$

and

(2.2) $$\lim_{N\to\infty} \min_{0 \leqq n \leqq b_N} \sigma_{n,a_N} = \infty.$$

Then we have

(2.3)
$$\limsup_{N\to\infty} \max_{0 \leqq n \leqq b_N} \max_{a_N \leqq k \leqq a_N + c_N} \max_{1 \leqq j \leqq k} \beta_{n,k} \Big| W\Big(\sum_{i=1}^{n+j} \tau_i\Big) - W\Big(\sum_{i=1}^{n} \tau_i\Big) \Big| \leqq 1 \quad a.s.$$

THEOREM 2.2. *Assume that there exists a constant $A \geqq 2$ such that*

(2.4) $$\max_{0 \leqq n \leqq b_N} \Big| \sum_{i=1+b_N}^{n+b_N} (\tau_i - \sigma_i) \Big| / \sigma_{b_N,a_N} \to 0 \quad a.s. \text{ as } N \to \infty,$$

(2.5) $$\sum_{i=1}^{b_N} \tau_i \leqq A\sigma_{0,b_N+a_N} \quad a.s. \text{ as } N \to \infty,$$

(2.6) $$\forall N \geqq 2, \quad \sum_{i=1+b_{N-1}}^{b_N} \sigma_i \leqq A \sum_{i=1+b_N}^{b_N+a_N} \sigma_i,$$

(2.7) $$\forall N \geqq 2, \quad \sum_{i=1}^{b_N+a_N} \sigma_i \leqq A \sum_{i=1}^{b_{N-1}+a_{N-1}} \sigma_i,$$

(2.8) $$\lim_{N\to\infty} \sigma_{b_N, a_N} = \infty.$$

Then we have

(2.9) $$\limsup_{N\to\infty} \alpha_{b_N,N}\Big(W\Big(\sum_{i=1}^{b_N+a_N} \tau_i\Big) - W\Big(\sum_{i=1}^{b_N} \tau_i\Big)\Big) = 1 \quad a.s.$$

THEOREM 2.3. *Assume that* (2.2), (2.6) *and* (2.7) *are satisfied. Moreover, suppose that*

(2.10) $$\max_{0\leqq n\leqq b_N+c_N} \max_{1\leqq k\leqq a_N} \Big|\sum_{i=1+n}^{n+k} (\tau_i - \sigma_i)\Big| / \sigma_{n,a_N} \to 0 \quad a.s. \text{ as } N\to\infty.$$

Then (2.9) *holds true and* (2.3) *remains valid with equality instead of inequality. If, in addition, we assume also that*

(2.11) $$\forall\, 0<\varepsilon<1, \quad \sum_{N=1}^{\infty} \exp\Big(-\sum_{j=0}^{[\frac{b_N}{a_N}]}\Big(\frac{\sigma_{ja_N,a_N}}{\sigma_{0,b_N+a_N}\log \sigma_{0,b_N+a_N}}\Big)^{1-\varepsilon}\Big) < \infty,$$

then

(2.12) $$\lim_{N\to\infty} \max_{0\leqq n\leqq b_N} \alpha_{n,N}\Big(W\Big(\sum_{i=1}^{n+a_N} \tau_i\Big) - W\Big(\sum_{i=1}^{n} \tau_i\Big)\Big) = 1 \quad a.s.,$$

(2.13)
$$\lim_{N\to\infty} \max_{0\leqq n\leqq b_N} \max_{a_N\leqq k\leqq a_N+c_N} \max_{1\leqq j\leqq k} \beta_{n,k}\Big|W\Big(\sum_{i=1}^{n+j} \tau_i\Big) - W\Big(\sum_{i=1}^{n} \tau_i\Big)\Big| = 1 \quad a.s.$$

To prove our theorems, we need the following lemmas.

LEMMA 2.1. *Let* $\{A_n, n\geqq 1\}$, $\{B_n, n\geqq 1\}$ *and* $\{D_n, n\geqq 1\}$ *be sequences of events with* $A_n \subset B_n \cup D_n$ *for* $n\geqq 1$. *If* $P(D_n, i.o.) = 0$, *then*

$$P(A_n, i.o.) \leqq P(B_n, i.o.).$$

The proof is trivial and so is omitted.

LEMMA 2.2. *Let* $\{A_n, n \geqq 1\}$, $\{B_n, n \geqq 1\}$ *and* $\{A_n^*, n \geqq 1\}$ *be sequences of events with* $A_n^* \subset A_n \cup B_n$ *for* $n \geqq 1$. *Assume that* $\bigcap_{i=1}^{n-1} A_i^c \cap B_i^c$ *and* A_n^{*c} *are independent for each* $n \geqq 2$ *and that*

$$\sum_{n=1}^{\infty} P(A_n^*) = \infty, \quad P(B_n, i.o.) = 0.$$

Then

$$P(A_n, i.o.) = 1.$$

PROOF. It suffices to show that

$$\lim_{n\to\infty} \lim_{m\to\infty} P\Big(\bigcap_{k=n}^{n+m} A_k^c \cap B_k^c\Big) = 0 \tag{2.14}$$

by Lemma 2.1 and the hypothesis $P(B_n, i.o.) = 0$. Notice that

$$P\Big(\bigcap_{k=n}^{n+m} A_k^c \cap B_k^c\Big) \leqq P\Big(\bigcap_{k=n}^{n+m-1} (A_k^c B_k^c) \bigcap A_{n+m}^{*c}\Big) = P\Big(\bigcap_{k=n}^{n+m-1} (A_k^c B_k^c)\Big) P(A_{n+m}^{*c})$$

by the assumptions. Recurring the above procedure, we can obtain that

$$P\Big(\bigcap_{k=n}^{n+m} A_k^c \cap B_k^c\Big) \leqq \prod_{k=n}^{n+m} P(A_k^{*c}) \leqq \exp\Big(-\sum_{k=n}^{n+m} P(A_k^*)\Big). \tag{2.15}$$

Now (2.14) follows from (2.15) and the condition $\sum_{k=1}^{\infty} P(A_k^*) = \infty$.

LEMMA 2.3. *Let* $\{W(t), \mathcal{F}_t, t \geqq 0\}$ *be a standard Wiener process and* τ *be a stopping time of* W. *Then* $\sigma\{W(\tau + t) - W(\tau), t \geqq 0\}$ *and* $\mathcal{F}_\tau$ *are independent and* $\{W(\tau + t) - W(\tau), \mathcal{F}_{\tau+t}, t \geqq 0\}$ *is also a standard Wiener process.*

This is the well-known strong Markov property of Wiener process.

LEMMA 2.4. *Let* $\{W(t), \mathcal{F}_t, t \geqq 0\}$ *be a standard Wiener process. Then, for every* $0 < \varepsilon < 1$, *there exists a positive constant* K *depending only on* ε *such that*

$$P\Big(\sup_{0\leqq t\leqq T-h}\ \sup_{0\leqq s\leqq h} |W(t+s) - W(t)| \geqq x\sqrt{h}\Big) \leqq K\frac{T}{h} e^{-\frac{x^2}{2+\varepsilon}}, \tag{2.16}$$

$$P(W(h) \geqq x\sqrt{h}) \geqq \frac{1}{K} e^{-\frac{2+\varepsilon}{2}x^2}, \tag{2.17}$$

for all $T \geqq h > 0$, $x > 0$.

The proof refers to Csörgő and Révész [5], p. 23 and p. 29.

LEMMA 2.5. *Let $\{W(t), t \geqq 0\}$ be a standard Wiener process. Then*

$$\limsup_{a\to\infty} \sup_{t>0} \sup_{s>0} \frac{|W(t+s)-W(t)|}{(2s(\log\frac{s+t}{s}+\log\log(a\vee(s+\frac{1}{s})))^{1/2}} = 1 \quad a.s. \tag{2.18}$$

PROOF. From the well-known law of the iterated logarithm it is obvious that left-hand side of (2.18) is greater than or equal to 1 almost surely. Noting that

$$\sup_{t>0} \sup_{s>0} \frac{|W(t+s)-W(t)|}{(2s(\log\frac{s+t}{s}+\log\log(a\vee(s+\frac{1}{s})))^{1/2}}$$

is a nonincreasing function of a, we only need to show that

$$P\left(\sup_{t>0} \sup_{s>0} \frac{|W(t+s)-W(t)|}{(2s(\log\frac{s+t}{s}+\log\log(a\vee(s+\frac{1}{s})))^{1/2}} \geqq \theta^2\right) \to 0 \tag{2.19}$$

as $a\to\infty$ for every $\theta > 1$. We have

$$\begin{aligned}
&\sup_{t>0} \sup_{s>0} \frac{|W(t+s)-W(t)|}{(2s(\log\frac{s+t}{s}+\log\log(a\vee(s+\frac{1}{s})))^{1/2}} \leqq \\
&\leqq \sup_{-\infty<j<\infty} \sup_{-\infty<i<\infty} \sup_{\theta^{i-1}\leqq t\leqq\theta^i} \sup_{\theta^{j-1}\leqq s\leqq\theta^j} \frac{|W(t+s)-W(t)|}{(2s(\log\frac{s+t}{s}+\log\log(a\vee\theta^{|j|})))^{1/2}} \\
&\leqq \sup_{-\infty<j<\infty} \sup_{j\leqq i<\infty} \sup_{0<t\leqq\theta^i} \sup_{0\leqq s\leqq\theta^j} \theta^{1/2} \frac{|W(t+s)-W(t)|}{(2\theta^j(\log\theta^{i-j}+\log\log(a\vee\theta^{|j|})))^{1/2}}.
\end{aligned} \tag{2.20}$$

Applying (2.16), we get that there is a positive constant K depending only on θ such that for each $-\infty < j \leqq i < \infty$, $a \geqq 1$,

$$\begin{aligned}
&P\left(\sup_{0<t\leqq\theta^i} \sup_{0\leqq s\leqq\theta^j} \frac{|W(t+s)-W(t)|}{(2\theta^j(\log\theta^{i-j}+\log\log(a\vee\theta^{|j|})))^{1/2}} \geqq 0\right) \leqq \\
&\leqq K\theta^{i-j}\exp\left(-\theta(\log\theta^{i-j}+\log\log(a\vee\theta^{|j|}))\right) \leqq K\theta^{(\theta-1)(i-j)}(|j|+a)^{-\theta}.
\end{aligned} \tag{2.21}$$

Now (2.19) follows from (2.20) and (2.21) immediately. This completes the proof of the lemma.

LEMMA 2.6. *We have*

$$\max_{0\leqq n\leqq b_N} \Big(\sum_{i=1}^{n} \tau_i\Big)/\sigma_{0,n+a_N} \leqq \max_{0\leqq n\leqq b_N} \Big(\sum_{i=1+n}^{n+a_N} \tau_i\Big)/\sigma_{n,a_N}. \tag{2.22}$$

PROOF. Note that for $0 \leqq n \leqq b_N$

$$\sum_{i=1}^{n} \tau_i \leqq \sum_{i=1}^{([n/a_N]+1)a_N} \tau_i = \sum_{j=0}^{[n/a_N]} \sigma_{ja_N,a_N} \sum_{i=1+ja_N}^{(j+1)a_N} \tau_i/\sigma_{ja_N,a_N} \leqq$$

$$\leqq \sum_{j=0}^{[n/a_N]} \sigma_{ja_N,a_N} \max_{0\leqq n\leqq b_N} \Big(\sum_{i=1+n}^{n+a_N} \tau_i\Big)/\sigma_{n,a_N} \leqq \sigma_{0,n+a_N} \max_{0\leqq n\leqq b_N} \Big(\sum_{i=1+n}^{n+a_N} \tau_i\Big)/\sigma_{n,a_N}.$$

This proves (2.22).

LEMMA 2.7. *Let $\{\eta_i, i \geqq 1\}$ be a sequence of random variables. Then we have*

$$\max_{0\leqq n\leqq b_N} \max_{a_N\leqq k\leqq a_N+c_N} \max_{1\leqq j\leqq k} \Big|\sum_{i=1+n}^{n+j} (\eta_i - \sigma_i)\Big|/\sigma_{n,k} \leqq$$

(2.23)

$$\leqq 3 \max_{0\leqq n\leqq b_N+c_N} \max_{1\leqq j\leqq a_N} \Big|\sum_{i=1+n}^{n+j} (\eta_i - \sigma_i)\Big|/\sigma_{n,a_N}.$$

PROOF. Write $j = ma_N + l$, where m, l are integers with $0 \leqq l < a_N$ for $a_N \leqq k \leqq a_N + c_N$ and $a_N \leqq j \leqq k$. Then

$$\Big|\sum_{i=1+n}^{n+j} (\eta_i - \sigma_i)\Big| \leqq \sum_{v=0}^{m-1}\Big|\sum_{i=1+n+va_N}^{n+(v+1)a_N} (\eta_i - \sigma_i)\Big| + \Big|\sum_{i=1+n+ma_N}^{n+j} (\eta_i - \sigma_i)\Big| \leqq$$

$$\leqq \sum_{v=0}^{m-1} \sigma_{n+va_N,a_N}\Big|\sum_{i=1+n+va_N}^{n+(v+1)a_N} (\eta_i - \sigma_i)\Big|/\sigma_{n+va_N,a_N} +$$

$$+\sigma_{n+j-a_N,a_N}\Big|\sum_{i=1+n+j-a_N}^{n+j} (\eta_i - \sigma_i)\Big|/\sigma_{n+j-a_N,a_N} +$$

$$+\sigma_{n+j-a_N,a_N}\Big|\sum_{i=1+n+j-a_N}^{1+n+ma_N} (\eta_i - \sigma_i)\Big|/\sigma_{n+j-a_N,a_N} \leqq$$

$$\leqq \Big(2\sigma_{n+j-a_N,a_N} + \sum_{v=0}^{m-1} \sigma_{n+va_N,a_N}\Big) \max_{0\leqq u\leqq b_N+c_N} \max_{1\leqq s\leqq a_N} \Big|\sum_{i=1+u}^{u+s} (\eta_i - \sigma_i)\Big|/\sigma_{u,a_N} \leqq$$

$$\leqq 3\sigma_{n,k} \max_{0\leqq u\leqq b_N+c_N} \max_{1\leqq s\leqq a_N} \Big|\sum_{i=1+u}^{u+s} (\eta_i - \sigma_i)\Big|/\sigma_{u,a_N}.$$

On the other hand, it is easy to see that

$$\Big|\sum_{i=1+a_N}^{n+j}(\eta_i-\sigma_i)\Big|\leqq\sigma_{n,k}\max_{0\leqq u\leqq b_N+c_N}\max_{1\leqq s\leqq a_N}\Big|\sum_{i=1+u}^{u+s}(\eta_i-\sigma_i)\Big|/\sigma_{u,a_N}.$$

Now (2.23) follows from the above inequalities.

PROOF OF THEOREM 2.1. It suffices to show that for every $0<\varepsilon<1$

$$P\Big(\max_{0\leqq n\leqq b_N}\max_{a_N\leqq k\leqq a_N+c_N}\max_{1\leqq j\leqq k}\beta_{n,k}\Big|W\Big(\sum_{i=1}^{n+j}\tau_i\Big)-W\Big(\sum_{i=1}^{n}\tau_i\Big)\Big|\geqq 1+\varepsilon, i.o.\Big)=0. \tag{2.24}$$

Notice that

$$\Big\{\max_{0\leqq n\leqq b_N}\max_{a_N\leqq k\leqq a_N+c_N}\max_{1\leqq j\leqq k}\beta_{n,k}\Big|W\Big(\sum_{i=1}^{n+j}\tau_i\Big)-W\Big(\sum_{i=1}^{n}\tau_i\Big)\Big|\geqq 1+\varepsilon\Big\}\subset$$

$$\subset\Big\{\max_{0\leqq n\leqq b_N}\max_{a_N\leqq k\leqq a_N+c_N}\frac{\sum_{i=1+n}^{n+k}\tau_i}{\sigma_{n,k}}\geqq 1+\frac{\varepsilon^2}{4}\Big\}\cup\Big\{\max_{0\leqq n\leqq b_N}\frac{\sum_{i=1}^{n}\tau_i}{\sigma_{0,n+a_N}}\geqq 4\Big\}\cup$$

$$\cup\Big\{\max_{0\leqq n\leqq b_N}\max_{a_N\leqq k\leqq a_N+c_N}\max_{1\leqq j\leqq k}\beta_{n,k}\Big|W\Big(\sum_{i=1}^{n+j}\tau_i\Big)-W\Big(\sum_{i=1}^{n}\tau_i\Big)\Big|\geqq 1+\varepsilon,$$

$$\max_{0\leqq n\leqq b_N}\max_{a_N\leqq k\leqq a_N+c_N}\frac{\sum_{i=1+n}^{n+k}\tau_i}{\sigma_{n,k}}<1+\frac{\varepsilon^2}{4},\ \max_{0\leqq n\leqq b_N}\frac{\sum_{i=1}^{n}\tau_i}{\sigma_{0,n+a_N}}<4\Big\}\subset$$

$$\subset\Big\{\max_{0\leqq n\leqq b_N}\max_{a_N\leqq k\leqq a_N+c_N}\frac{\sum_{i=1+n}^{n+k}\tau_i}{\sigma_{n,k}}\geqq 1+\frac{\varepsilon^2}{4}\Big\}\cup\Big\{\max_{0\leqq n\leqq b_N}\frac{\sum_{i=1}^{n}\tau_i}{\sigma_{0,n+a_N}}\geqq 4\Big\}\cup$$

$$\cup\Big\{\max_{0\leqq n\leqq b_N}\max_{a_N\leqq k\leqq a_N+c_N}\sup_{0\leqq t\leqq 4\sigma_{0,n+a_N}}\sup_{0\leqq s\leqq(1+\frac{\varepsilon^2}{4})\sigma_{n,k}}\beta_{n,k}|W(t+s)-W(t)|\geqq 1+\varepsilon\Big\}\subset$$

$$\subset\Big\{\max_{0\leqq n\leqq b_N}\max_{a_N\leqq k\leqq a_N+c_N}\frac{\sum_{i=1+n}^{n+k}\tau_i}{\sigma_{n,k}}\geqq 1+\frac{\varepsilon^2}{4}\Big\}\cup\Big\{\max_{0\leqq n\leqq b_N}\frac{\sum_{i=1}^{n}\tau_i}{\sigma_{0,n+a_N}}\geqq 4\Big\}\cup$$

$$\cup\Big\{\sup_{t>0}\sup_{s>0}\frac{|W(t+s)-W(t)|}{(2s(\log\frac{t+s}{s}+\log\log(\sigma_N^*\vee(s+\frac{1}{s}))))^{1/2}}\geqq 1+\frac{\varepsilon}{4}\Big\}$$

for every N sufficiently large, where $\sigma_N^*=\min_{0\leqq n\leqq bN}\sigma_{n,a_N}$. Now (2.24) follows from (2.1) and Lemmas 2.1, 2.5 and 2.6. This completes the proof of Theorem 2.1.

PROOF OF THEOREM 2.2. From the proof of Theorem 2.1 one can see that

$$(2.25)\qquad \limsup_{N\to\infty} \alpha_{b_N,N}\Big(W\Big(\sum_{i=1}^{b_N+a_N}\tau_i\Big)-W\Big(\sum_{i=1}^{b_N}\tau_i\Big)\Big)\leqq 1 \quad \text{a.s.}$$

So it suffices to show that for every $0<\varepsilon<\frac{1}{8}$

$$(2.26)\qquad \limsup_{N\to\infty} \alpha_{b_N,N}\Big(W\Big(\sum_{i=1}^{b_N+a_N}\tau_i\Big)-W\Big(\sum_{i=1}^{b_N}\tau_i\Big)\Big)\geqq 1-6\varepsilon \quad \text{a.s.}$$

Let $N_1=1$. Define

$$(2.27)\qquad N_{k+1}=\min\Big\{n:\ \sum_{i=1}^{b_n}\sigma_i+\varepsilon^2\sum_{i=1+b_n}^{b_n+a_n}\sigma_i\geqq\sum_{i=1}^{b_{N_k}+a_{N_k}}\sigma_i\Big\},\quad k=1,2,\ldots.$$

Then

$$(2.28)\qquad \forall k\geqq 1,\quad \sum_{i=1}^{b_{N_{k+1}}}\sigma_i+\varepsilon^2\sum_{i=1+b_{N_{k+1}}}^{b_{N_{k+1}}+a_{N_{k+1}}}\sigma_i\geqq\sum_{i=1}^{b_{N_k}+a_{N_k}}\sigma_i$$

and

$$(2.29)\qquad \forall n<N_{k+1},\quad \sum_{i=1}^{b_n}\sigma_i+\varepsilon^2\sum_{i=1+b_n}^{b_n+a_n}\sigma_i<\sum_{i=1}^{b_{N_k}+a_{N_k}}\sigma_i.$$

By (2.28) and (2.29), we find that $N_{k+1}>N_k$, $b_{N_{k+1}}+a_{N_{k+1}}>b_{N_k}+a_{N_k}$ for each $k\geqq 1$. We first prove that

$$(2.30)\qquad \sum_{k=1}^{\infty}\frac{\sigma_{b_{N_k},a_{N_k}}}{\sigma_{0,b_{N_k}+a_{N_k}}\log\sigma_{0,b_{N_k}+a_{N_k}}}=\infty.$$

In terms of (2.29), (2.6) and (2.7), we obtain

$$(2.31)\qquad \sum_{i=1}^{b_{N_{k-1}}+a_{N_{k-1}}}\sigma_i\geqq\sum_{i=1}^{b_{N_k}}\sigma_i-\sum_{i=1+b_{N_k-1}}^{b_{N_k}}\sigma_i\geqq\sigma_{0,b_{N_k}+a_{N_k}}-(A+1)\sigma_{b_{N_k},a_{N_k}}$$

and

$$(2.32)\qquad \sum_{i=1}^{b_{N_{k-1}}+a_{N_{k-1}}}\sigma_i\geqq\varepsilon^2\sum_{i=1}^{b_{N_k-1}+a_{N_k-1}}\sigma_i\geqq\frac{\varepsilon^2}{A}\sum_{i=1}^{b_{N_k}+a_{N_k}}\sigma_i$$

by (2.29) and (2.7). Using (2.31) and (2.32), we have

$$\sum_{k=2}^{\infty} \frac{\sigma_{b_{N_k},a_{N_k}}}{\sigma_{0,b_{N_k}+a_{N_k}} \log \sigma_{0,b_{N_k}+a_{N_k}}} \geqq \frac{1}{A+1} \sum_{k=2}^{\infty} \frac{\sigma_{0,b_{N_k}+a_{N_k}} - \sigma_{0,b_{N_{k-1}}+a_{N_{k-1}}}}{\sigma_{0,b_{N_k}+a_{N_k}} \log \sigma_{0,b_{N_k}+a_{N_k}}} \geqq$$

$$\geqq \frac{\varepsilon^2}{A(A+1)} \sum_{k=2}^{\infty} \frac{\sigma_{0,b_{N_k}+a_{N_k}} - \sigma_{0,b_{N_{k-1}}+a_{N_{k-1}}}}{\sigma_{0,b_{N_{k-1}}+a_{N_{k-1}}} \log(\frac{A}{\varepsilon^2}\sigma_{0,b_{N_{k-1}}+a_{N_{k-1}}})} \geqq$$

$$\geqq \frac{\varepsilon^2}{A(A+1)} \sum_{k=2}^{\infty} \int_{\sigma_{0,b_{N_{k-1}}+a_{N_{k-1}}}}^{\sigma_{0,b_{N_k}+a_{N_k}}} \frac{1}{x \log(\frac{xA}{\varepsilon^2})}\, dx = \infty$$

as desired. Put

$$\mathcal{G} = \{k\colon b_{N_k} \geqq b_{N_{k-1}} + a_{N_{k-1}}\}, \quad \mathcal{H} = \{k\colon b_{N_k} < b_{N_{k-1}} + a_{N_{k-1}}\}.$$

We formulate the proof of (2.26) in two cases.

Case I. Assume that

$$\sum_{k\in\mathcal{G}} \frac{\sigma_{b_{N_k},a_{N_k}}}{\sigma_{0,b_{N_k}+a_{N_k}} \log \sigma_{0,b_{N_k}+a_{N_k}}} = \infty. \tag{2.33}$$

For each $k \in \mathcal{G}$, let

$$\tau(k) = \sum_{i=1}^{b_{N_k}} \tau_i, \quad \theta(k) = \sum_{i=1+b_{N_k}}^{b_{N_k}+a_{N_k}} \tau_i,$$

$$\lambda(k) = \sigma_{b_{N_k},a_{N_k}}, \qquad \alpha(k) = \alpha_{b_{N_k},N_k}, \qquad \gamma(k) = \sigma_{0,b_{N_k}+a_{N_k}},$$

$$A_k = \{\alpha(k)(W(\tau(k)+\theta(k)) - W(\tau(k))) \geqq 1-4\varepsilon\},$$

$$B_k = \{|\theta(k) - \lambda(k)| \geqq 2\varepsilon^2\lambda(k)\},$$

$$A_k^* = \Big\{\inf_{(1-2\varepsilon^2)\lambda(k)\leqq s\leqq(1+2\varepsilon^2)\lambda(k)} \alpha(k)(W(s+\tau(k)) - W(\tau(k))) \geqq 1-4\varepsilon\Big\}.$$

Clearly, $A_k^* \subset A_k \cup B_k$, $\tau(k') + \theta(k') \leqq \tau(k'')$ and $\mathcal{F}_{\tau(k')+\theta(k')} \subset \mathcal{F}_{\tau(k'')}$ for every $k' < k''$, $k, k', k'' \in \mathcal{G}$. Hence A_k^* and $\mathcal{F}_{\tau(k)}$ are independent and so are A_k^* and $\{A_{k'}, b_{k'}; k' < k, k' \in \mathcal{G}\}$ by Lemma 2.3. To arrive at (2.26), we only need to verify that

$$\sum_{k\in\mathcal{G}} P(A_k^*) = \infty \tag{2.34}$$

by (2.4) and Lemma 2.2.

Applying Lemmas 2.3 and 2.4, we have

$$P(A_k^*) = P\Big(\inf_{(1-2\varepsilon^2)\lambda(k)\leqq s\leqq(1+2\varepsilon^2)\lambda(k)} \alpha(k)W(s) \geqq 1-4\varepsilon\Big) \geqq$$

$$\geqq P(\alpha(k)W(\lambda(k)) \geqq 1-\varepsilon) - P\Big(\sup_{0\leqq t\leqq 4\lambda(k)} \sup_{0\leqq s\leqq 4\varepsilon^2\lambda(k)} \alpha(k)|W(t+s)-W(t)| \geqq 3\varepsilon\Big) \geqq$$

$$\geqq \frac{1}{K}\exp\left\{-(1-\varepsilon)\Big(\log\frac{\sigma_{0,b_{N_k}+a_{N_k}}}{\sigma_{b_{N_k},a_{N_k}}} + \log\log\sigma_{0,b_{N_k}+a_{N_k}}\Big)\right\} -$$

$$-\frac{32K}{\varepsilon^2}\exp\left\{-2\Big(\log\frac{\sigma_{0,b_{N_k}+a_{N_k}}}{\sigma_{b_{N_k},a_{N_k}}} + \log\log\sigma_{0,b_{N_k}+a_{N_k}}\Big)\right\} \geqq$$

$$\geqq \frac{\sigma_{b_{N_k},a_{N_k}}}{2K\sigma_{0,b_{N_k}+a_{N_k}}\log\sigma_{0,b_{N_k}+a_{N_k}}}$$

for every $k\in\mathcal{G}$ sufficiently large. This proves that (2.34) holds true by the hypothesis (2.33) and so does (2.26) in this case.

Case II. Assume that

$$\sum_{k\in\mathcal{G}} \frac{\sigma_{b_{N_k},a_{N_k}}}{\sigma_{0,b_{N_k}+a_{N_k}}\log\sigma_{0,b_{N_k}+a_{N_k}}} < \infty. \tag{2.35}$$

Then, by (2.30)

$$\sum_{k\in\mathcal{H}} \frac{\sigma_{b_{N_k},a_{N_k}}}{\sigma_{0,b_{N_k}+a_{N_k}}\log\sigma_{0,b_{N_k}+a_{N_k}}} = \infty. \tag{2.36}$$

From (2.28) it follows that for each $k\in\mathcal{H}$

$$0 \leqq \sigma_{0,b_{N_{k-1}}+a_{N_{k-1}}} - \sigma_{0,b_{N_k}} \leqq \varepsilon^2\sigma_{b_{N_k},a_{N_k}} \tag{2.37}$$

and hence

$$(1-\varepsilon^2)\sigma_{b_{N_k},a_{N_k}} \leqq \sigma_{0,b_{N_k}+a_{N_k}} - \sigma_{0,b_{N_{k-1}}+a_{N_{k-1}}} \leqq \sigma_{b_{N_k},a_{N_k}}. \tag{2.38}$$

For each $k\in\mathcal{H}$, write

$$\alpha(k)\Big(W\Big(\sum_{i=1}^{b_{N_k}+a_{N_k}}\tau_i\Big) - W\Big(\sum_{i=1}^{b_{N_k}}\tau_i\Big)\Big) =$$

$$= \alpha(k)\Big(W\Big(\sum_{i=1}^{b_{N_k}+a_{N_k}}\tau_i\Big) - W\Big(\sum_{i=1}^{b_{N_{k-1}}+a_{N_{k-1}}}\tau_i\Big)\Big) + \tag{2.39}$$

$$+\alpha(k)\Big(W\Big(\sum_{i=1}^{b_{N_{k-1}}+a_{N_{k-1}}}\tau_i\Big) - W\Big(\sum_{i=1}^{b_{N_k}}\tau_i\Big)\Big).$$

By (2.4), we get

$$\limsup_{k\in\mathcal{H},k\to\infty}\Big|\sum_{i=1+b_{N_k}}^{b_{N_{k-1}}+a_{N_{k-1}}}(\tau_i-\sigma_i)\Big|/\sigma_{b_{N_k},a_{N_k}}=0,\quad \text{a.s.}$$

Therefore

$$\limsup_{k\in\mathcal{H},k\to\infty}\sum_{i=1+b_{N_k}}^{b_{N_{k-1}}+a_{N_{k-1}}}\tau_i/\sigma_{b_{N_k},a_{N_k}}\leqq\varepsilon^2\quad\text{a.s.}\tag{2.40}$$

by (2.37). It is easy to see that

$$\Big\{\alpha(k)\Big|W\Big(\sum_{i=1}^{b_{N_{k-1}}+a_{N_{k-1}}}\tau_i\Big)-W\Big(\sum_{i=1}^{b_{N_k}}\tau_i\Big)\Big|\geqq2\varepsilon\Big\}\subset$$

$$\subset\Big\{\sum_{i=1+b_{N_k}}^{b_{N_{k-1}}+a_{N_{k-1}}}\tau_i\geqq2\varepsilon^2\sigma_{b_{N_k},a_{N_k}}\Big\}\cup\Big\{\sum_{i=1}^{b_{N_k}}\tau_i\geqq2A\sigma_{0,b_{N_k}+a_{N_k}}\Big\}\cup$$

$$\cup\Big\{\sup_{0\leqq t\leqq2A\sigma_{0,b_{N_k}+a_{N_k}}}\ \sup_{0\leqq s\leqq2\varepsilon^2\sigma_{b_{N_k},a_{N_k}}}\alpha(k)|W(t+s)-W(t)|\geqq2\varepsilon\Big\}.$$

Similarly to the proof of (2.24), we can obtain that

$$\limsup_{k\in\mathcal{H},k\to\infty}\ \sup_{0\leqq t\leqq2A\sigma_{0,b_{N_k}+a_{N_k}}}\ \sup_{0\leqq s\leqq2\varepsilon^2\sigma_{b_{N_k},a_{N_k}}}\alpha(k)|W(t+s)-W(t)|\leqq2\varepsilon\quad\text{a.s.}\tag{2.41}$$

and therefore

$$\limsup_{k\in\mathcal{H},k\to\infty}\alpha(k)\Big|W\Big(\sum_{i=1}^{b_{N_{k-1}}+a_{N_{k-1}}}\tau_i\Big)-W\Big(\sum_{i=1}^{b_{N_k}}\tau_i\Big)\Big|\leqq2\varepsilon,\quad\text{a.s.}\tag{2.42}$$

by (2.40), (2.5), (2.41) and Lemma 2.1.

To finish the proof of (2.26), it suffices to show that

$$\limsup_{k\in\mathcal{H},k\to\infty}\alpha(k)\Big(W\Big(\sum_{i=1}^{b_{N_k}+a_{N_k}}\tau_i\Big)-W\Big(\sum_{i=1}^{b_{N_{k-1}}+a_{N_{k-1}}}\tau_i\Big)\Big)\geqq1-4\varepsilon,\quad\text{a.s.}\tag{2.43}$$

Let

$$\tau(k)=\sum_{i=1}^{b_{N_{k-1}}+a_{N_{k-1}}}\tau_i,\qquad\theta(k)=\sum_{i=1+b_{N_{k-1}}+a_{N_{k-1}}}^{b_{N_k}+a_{N_k}}\tau_i,$$

$$\lambda(k)=\sum_{i=1+b_{N_{k-1}}+a_{N_{k-1}}}^{b_{N_k}+a_{N_k}}\sigma_i,\qquad \gamma(k)=\sigma_{0,b_{N_k}+a_{N_k}}.$$

Using (2.38) and (2.40) and proceeding the same lines of the proof in the first case, one can obtain that (2.43) holds true.

Now the proof of Theorem 2.2 is completed.

PROOF OF THEOREM 2.3. The first part of our conclusion follows easily from the assumption (1.6), Lemmas 2.6 and 2.7 and Theorems 2.1 and 2.2. Assuming (2.11), we prove that (2.12) is true. It suffices to show that for every $0<\varepsilon<1/2$

$$P\Big(\max_{0\leqq j\leqq[\frac{b_N}{a_N}]}\alpha_{ja_N,N}\Big(W\Big(\sum_{i=1}^{(j+1)a_N}\tau_i\Big)-W\Big(\sum_{i=1}^{ja_N}\tau_i\Big)<1-\varepsilon,i.o.\Big)=0 \tag{2.44}$$

by (2.3). Let

$$\tau_N(j)=\sum_{i=1}^{ja_N}\tau_i,\quad \theta_N(j)=\sum_{i=1+ja_N}^{(j+1)a_N}\tau_i,\quad \lambda_N(j)=\sigma_{ja_N,N}.$$

Clearly, we have

$$\Big\{\max_{0\leqq j\leqq[\frac{b_N}{a_N}]}\alpha_{ja_N,N}\Big(W\Big(\sum_{i=1}^{(j+1)a_N}\tau_i\Big)-W\Big(\sum_{i=1}^{ja_N}\tau_i\Big)<1-\varepsilon\Big\}\subset$$

$$\subset\Big\{\max_{0\leqq j\leqq[\frac{b_N}{a_N}]}\frac{|\theta_N(j)-\lambda_N(j)|}{\lambda_N(j)}\geqq\frac{\varepsilon^2}{32}\Big\}\cup$$

$$\cup\Big\{\max_{0\leqq j\leqq[\frac{b_N}{a_N}]}\alpha_{ja_N,N}(W(\tau_N(j)+\theta_N(j))-W(\tau_N(j))<$$

$$<1-\varepsilon,\ \max_{0\leqq j\leqq[\frac{b_N}{a_N}]}\frac{|\theta_N(j)-\lambda_N(j)|}{\lambda_N(j)}<\frac{\varepsilon^2}{32}\Big\}.$$

Therefore, we only need to prove

$$P\Big(\max_{0\leqq j\leqq[\frac{b_N}{a_N}]}\alpha_{ja_N,N}(W(\tau_N(j)+\theta_N(j))-W(\tau_N(j)))<1-\varepsilon,$$

$$\max_{0\leqq j\leqq[\frac{b_N}{a_N}]}\frac{|\theta_N(j)-\lambda_N(j)|}{\lambda_N(j)}<\frac{\varepsilon^2}{32},i.o.\Big)=0 \tag{2.45}$$

by (2.10) and Lemma 2.1.

Noting that $\{W(\tau_N(j)+s)-W(\tau_N(j)), s \geqq 0\}$ and $\{\mathcal{F}_{\tau_N(l)+\theta_N(l)}, l<j\}$ are independent by Lemma 2.3, we find that

$$P\Big(\max_{0\leqq j\leqq[\frac{b_N}{a_N}]} \alpha_{ja_N,N}(W(\tau_N(j)+\theta_N(j)))-W(\tau_N(j))<1-\varepsilon,$$

$$\max_{0\leqq j\leqq[\frac{b_N}{a_N}]}\frac{|\theta_N(j)-\lambda_N(j)|}{\lambda_N(j)}<\frac{\varepsilon^2}{32}\Big)\leqq$$

$$\leqq P\Big(\max_{0\leqq j<[\frac{b_N}{a_N}]} \alpha_{ja_N,N}(W(\tau_N(j)+\theta_N(j)))-W(\tau_N(j))<1-\varepsilon,$$

$$\max_{0\leqq j<[\frac{b_N}{a_N}]}\frac{|\theta_N(j)-\lambda_N(j)|}{\lambda_N(j)}<\frac{\varepsilon^2}{32},$$

$$\inf_{(1-\frac{\varepsilon^2}{32})\lambda_N([\frac{b_N}{a_N}])\leqq s\leqq(1+\frac{\varepsilon^2}{32})\lambda_N([\frac{b_N}{a_N}])} \alpha_{[\frac{b_N}{a_N}]a_N,N}\Big(W\Big(\tau\Big(\Big[\frac{b_N}{a_N}\Big]\Big)+s\Big)-$$

$$-W\Big(\tau_N\Big(\Big[\frac{b_N}{a_N}\Big]\Big)\Big)<1-\varepsilon\Big)=$$

$$=P\Big(\inf_{(1-\frac{\varepsilon^2}{32})\lambda_N([\frac{b_N}{a_N}])\leqq s\leqq(1+\frac{\varepsilon^2}{32})\lambda_N([\frac{b_N}{a_N}])} \alpha_{[\frac{b_N}{a_N}]a_N,N}\Big(W\Big(\tau\Big(\Big[\frac{b_N}{a_N}\Big]\Big)+s\Big)-$$

$$-W\Big(\tau_N\Big(\Big[\frac{b_N}{a_N}\Big]\Big)\Big)<1-\varepsilon\Big)$$

$$P\Big(\max_{0\leqq j<[\frac{b_N}{a_N}]} \alpha_{ja_N,N}(W(\tau_N(j)+\theta_N(j)))-W(\tau_N(j))<1-\varepsilon,$$

$$\max_{0\leqq j<[\frac{b_N}{a_N}]}\frac{|\theta_N(j)-\lambda_N(j)|}{\lambda_N(j)}<\frac{\varepsilon^2}{32}\Big)$$

and hence

$$P\Big(\max_{0\leqq j<[\frac{b_N}{a_N}]} \alpha_{ja_N,N}(W(\tau_N(j)+\theta_N(j))-W(\tau_N(j))<1-\varepsilon,$$

$$\max_{0\leqq j\leqq[\frac{b_N}{a_N}]} \frac{|\theta_N(j)-\lambda_N(j)|}{\lambda_N(j)}<\frac{\varepsilon^2}{32}\Big)\leqq$$

(2.46)

$$\leqq \prod_{j=0}^{[\frac{b_N}{a_N}]} P\Big(\inf_{(1-\frac{\varepsilon^2}{32})\lambda_N(j)\leqq s\leqq(1+\frac{\varepsilon^2}{32})\lambda_N(j)} \alpha_{ja_N,N}(W(\tau(j)+s)-W(\tau_N(j))<1-\varepsilon\Big)=$$

$$=\prod_{j=0}^{[\frac{b_N}{a_N}]} P\Big(\inf_{(1-\frac{\varepsilon^2}{32})\lambda_N(j)\leqq s\leqq(1+\frac{\varepsilon^2}{32})\lambda_N(j)} \alpha_{ja_N,N}W(s)<1-\varepsilon\Big)$$

by Lemma 2.3. Applying Lemma 2.4 yields for each $0\leqq j\leqq[\frac{b_N}{a_N}]$ that

$$P\Big(\inf_{(1-\frac{\varepsilon^2}{32})\lambda_N(j)\leqq s\leqq(1+\frac{\varepsilon^2}{32})\lambda_N(j)} \alpha_{ja_N,N}W(s)<1-\varepsilon\Big)\leqq$$

$$\leqq P(\alpha_{ja_N,N}W\Big(\lambda_N(j))<1-\frac{\varepsilon}{2}\Big)+$$

$$(2.47)\qquad +P\Big(\sup_{0\leqq t\leqq 4\lambda_N(j)}\ \sup_{0\leqq s\leqq \varepsilon^2\lambda_N(j)/16} \alpha_{ja_N,N}|W(t+s)-W(t)|>\frac{\varepsilon}{2}\Big)\leqq$$

$$\leqq 1-K^{-1}(\sigma_{ja_N,N}/(\sigma_{0,b_N+a_N}\log\sigma_{0,b_N+a_N}))^{1-\varepsilon/3}+$$

$$+K(\sigma_{ja_N,N}/(\sigma_{0,b_N+a_N}\log\sigma_{0,b_N+a_N}))^{3/2}\leqq$$

$$\leqq 1-(\sigma_{ja_N,N}/(\sigma_{0,b_N+a_N}\log\sigma_{0,b_N+a_N}))^{1-\varepsilon/4}\leqq$$

$$\leqq \exp\left(-(\sigma_{ja_N,N}/(\sigma_{0,b_N+a_N}\log\sigma_{0,b_N+a_N}))^{1-\varepsilon/4}\right)$$

for every N sufficiently large.

Now (2.45) follows from (2.46), (2.47), (2.11) and the Borel–Cantelli lemma. This completes the proof of Theorem 2.3.

Due to the requirement in the proof of our future work, we give the following remark.

REMARK 2.1. *Let* $\{a(j,N)\}$ *be an array of positive integers with* $a(0,N)=0$ *and* $a_N\leqq a(j,N)-a(j-1,N)\leqq a_N+c_N$ *for each* $j\geqq 1$. *Assume* (2.2), (2.6), (2.7), (2.10) *and*

$$(2.48)\quad \forall\, 0<\varepsilon<1,\qquad \sum_{N=1}^{\infty}\exp\left(-\sum_{i=0}^{jN}\Big(\frac{\sigma_{a(i,N),a(i+1,N)-a(i,N)}}{\sigma_{0,b_N+a_N}\log\sigma_{0,b_N+a_N}}\Big)^{1-\varepsilon}\right)<\infty,$$

where $j_N := \max\{i : a(i,N) \leqq b_N\}$. *Then, we have*

(2.49) $$\lim_{N\to\infty} \max_{0\leqq n\leqq j_N} \alpha'_{n,N}\Big(W\Big(\sum_{i=1}^{a(n,N)} \tau_i\Big) - W\Big(\sum_{i=1}^{a(n-1,N)} \tau_i\Big)\Big) = 1 \quad a.s.,$$

where
(2.50)
$$\alpha'_{n,N} = \Big\{2\Big(\sum_{i=1+a(n-1,N)}^{a(n,N)} \sigma_i\Big)\Big(\log \frac{\sum_{i=1}^{a_N+b_N} \sigma_i}{\sum_{i=1+a(n-1,N)}^{a(n,N)} \sigma_i} + \log\log \sum_{i=1}^{a_N+b_N} \sigma_i\Big)\Big\}^{-1/2}.$$

PROOF. Let
$$\tau_N(j) = \sum_{i=1}^{a(j,N)} \tau_i, \quad \theta_N(j) = \sum_{i=1+a(j,N)}^{a(j+1,N)} \tau_i, \quad \lambda_N(j) = \sum_{i=1+a(j,N)}^{a(j+1,N)} \sigma_i.$$

It is easy to see that

$$\Big\{\max_{1\leqq n\leqq j_N} \alpha'_{n,N}\Big(W\Big(\sum_{i=1}^{a(n,N)} \tau_i\Big) - W\Big(\sum_{i=1}^{a(n-1,N)} \tau_i\Big) < 1-\varepsilon\Big\} \subset$$
$$\subset \Big\{\max_{1\leqq n\leqq j_N} \alpha'_{n,N}(W(\tau_N(n-1)+\theta_N(n-1)) - W(\tau_N(n-1)) < 1-\varepsilon,$$
$$\max_{1\leqq j\leqq j_N} \frac{|\theta_N(n)-\lambda_N(n)|}{\lambda_N(n)} < \frac{\varepsilon^2}{32}\Big\} \cup \Big\{\max_{1\leqq n\leqq j_N} \frac{|\theta_N(n)-\lambda_N(n)|}{\lambda_N(n)} \geqq \frac{\varepsilon^2}{32}\Big\}.$$

By (2.10), Lemma 2.7 and the assumption that $a_N \leqq a(n,N) - a(n-1,N) \leqq$ $\leqq a_N + c_N$, we have

$$P\Big(\max_{1\leqq n\leqq j_N} |\theta_N(n)-\lambda_N(n)|/\lambda_N(n) > \frac{\varepsilon^2}{32}, i.o.\Big) = 0.$$

Along the same lines of the proof of (2.45), one can obtain that

$$P\Big(\max_{1\leqq n\leqq j_N} \alpha'_{n,N}(W(\tau_N(n-1)+\theta_N(n-1)) - W(\tau_N(n-1)) < 1-\varepsilon,$$
$$\max_{1\leqq n\leqq j_N} \frac{|\theta_N(n)-\lambda_N(n)|}{\lambda_N(n)} < \frac{\varepsilon^2}{32}, i.o.\Big) = 0.$$

Hence

(2.51) $$P(\max_{1\leqq n\leqq j_N} \alpha'_{n,N}(W(\tau_N(n-1)+\theta_N(n-1)) - W(\tau_N(n-1)) < 1-\varepsilon, i.o.) = 0$$

for every $0<\varepsilon<1$. This proves (2.49) by (2.3) and (2.51).

3. How big are the increments for martingale difference sequences

The results on the random increments of a Wiener process in Section 2 and the Strassen's embedding theorem enable us to establish the increments for martingale difference sequences. The latter plays also a key role in discussing the increments for dependent random variables. Throughout this section let $\{Y_n, \mathcal{F}_n, n \geqq 1\}$ be a sequence of square integrable martingale differences and $\{a_N, N \geqq 1\}$, $\{b_N, N \geqq 1\}$ and $\{c_N, N \geqq 1\}$ be sequences of integer numbers with $a_N \to \infty$ as in Section 2. Put

$$\sigma_n = EY_n^2, \quad T_n = \sum_{i=1}^{n} Y_i, \quad T_n(k) = \sum_{i=n+1}^{n+k} Y_i, \quad n \geqq 0, \ k \geqq 1.$$

Define

$$\mathcal{H}_k = \{N : k \leqq a_N < k+1\}, \ M_k = \max\{b_N + c_N : N \in \mathcal{H}_k\}, \ M_k = -1 \text{ if } \mathcal{H}_k = \emptyset,$$

and let $\sigma_{n,k}$, $\beta_{n,k}$ and $\alpha_{n,N}$ be as in (1.4), (1.5) and (1.6), respectively.

THEOREM 3.1. *Assume that there exist a non-decreasing sequence* $\{D_n, n \geqq 1\}$ *and a constant* $A \geqq 2$ *such that*

(3.1) $$|Y_n| \leqq D_n \quad \text{for each } \ n \geqq 1,$$

(3.2) $$\max_{0 \leqq n \leqq M_k} \max_{1 \leqq j \leqq k+1} \frac{\left| \sum_{i=1+n}^{n+j} E(Y_i^2|\mathcal{F}_{i-1}) - EY_i^2 \right|}{\sigma_{n,k+1}} \to 0 \quad \text{a.s. as } k \to \infty,$$

(3.3) $$\sigma_{n,k+1} \leqq A\sigma_{n,k} \quad \text{for all } \ 0 \leqq n \leqq M_k \ \text{ and } \ k \geqq 1,$$

(3.4) $$\sum_{k=1}^{\infty} \sum_{n=0}^{M_k} \exp\left(-\varepsilon \sigma_{n,k+1} / D_{n+k+1}^2\right) < \infty.$$

Then, we have

(3.5) $$\limsup_{N \to \infty} \max_{0 \leqq n \leqq b_N} \max_{a_N \leqq k \leqq a_N + c_N} \max_{1 \leqq j \leqq k} \beta_{n,k} |T_n(j)| \leqq 1 \quad \text{a.s.}$$

THEOREM 3.2. *Assume* (2.6), (2.7) *and the conditions in Theorem 3.1 are satisfied. Then*

(3.6) $$\limsup_{N \to \infty} \alpha_{b_N,N} T_{b_N}(a_N) = 1 \quad \text{a.s.}$$

(3.7) $$\limsup_{N\to\infty} \max_{0\leqq n\leqq b_N} \max_{a_N\leqq k\leqq a_N+c_N} \max_{1\leqq j\leqq k} \beta_{n,k}|T_n(j)| = 1 \quad a.s.$$

In addition, if (2.11) *is satisfied, then*

(3.8) $$\lim_{N\to\infty} \max_{0\leqq n\leqq b_N} \alpha_{n,N} T_n(a_N) = 1 \quad a.s.$$

(3.9) $$\lim_{N\to\infty} \max_{0\leqq n\leqq b_N} \max_{a_N\leqq k\leqq a_N+c_N} \max_{1\leqq j\leqq k} \beta_{n,k}|T_n(j)| = 1. \quad a.s.$$

THEOREM 3.3. *Assume that* (2.6), (2.7) *and* (3.1) *are satisfied and suppose also that*

(3.10) $$\max_{0\leqq n\leqq a_N} \frac{\left|\sum_{i=1+b_N}^{n+b_N} E(Y_i^2|\mathcal{F}_{i-1}) - EY_i^2\right|}{\sigma_{b_N,a_N}} \to 0 \quad a.s. \text{ as } N\to\infty,$$

(3.11) $$\sum_{i=1}^{b_N} (E(Y_i^2 \mid \mathcal{F}_{i-1}) - EY_i^2) \leqq A\sigma_{0,b_N+a_N} \quad a.s. \text{ as } N\to\infty$$

and

(3.12) $$\sum_{N=1}^{\infty} \exp\left(-\varepsilon\sigma_{b_N,a_N}/D^2_{b_N+a_N}\right) < \infty.$$

Then (3.6) holds true.

The following theorem makes the assumptions in Theorems 3.1 and 3.2 to be verified easily.

THEOREM 3.4. *Let* $H(x): [0,\infty)\to(0,\infty)$ *be a nondecreasing function. Assume there exist constants* $\delta > 0$, $C_1 > 0$, $C_2 > 0$ *and* $\theta \geqq 0$ *such that*

(3.13) $$P(|Y_n| \geqq \varepsilon H(n), i.o.) = 0 \quad \text{for every } \varepsilon > 0,$$

(3.14) $$H^2(x)/x^\theta \quad \text{is nondecreasing,}$$

(3.15) $$C_1 n^\theta \leqq \sigma_n \leqq (E|Y_n|^{2+\delta})^{\frac{2}{2+\delta}} \leqq C_2 n^\theta \quad \text{for each } n \geqq 1,$$

(3.16) $$E(Y_n^2 \mid \mathcal{F}_{n-1}) - EY_n^2 = o(n^\theta) \quad a.s.,$$

(3.17) $$E(|Y_n|^{2+\delta} \mid \mathcal{F}_{n-1}) \leqq C_2 n^{\frac{\theta(2+\delta)}{2}} \quad a.s. \text{ as } n\to\infty$$

$$\lim_{N\to\infty} \frac{a_N}{\log(a_N+b_N+C_N)} = \infty, \tag{3.18}$$

$$\frac{a_N(a_N+b_N+c_N)^\theta \log^2(a_N/\log(a_N+b_N+c_N))}{H^2(30(a_N+b_N+c_N))\log(a_N+b_N+c_N)} \geqq C_1 \text{ for each } N \geqq 1. \tag{3.19}$$

Then (3.5) *holds. If, in addition, we also have*

$$b_N - b_{N-1} \leqq C_2 a_N \tag{3.20}$$

and

$$a_N \leqq C_2 a_{N-1} \tag{3.21}$$

for every $N \geqq 2$, *then* (3.6) *and* (3.7) *are true. Furthermore, if we assume that for every* $0 < \varepsilon < 1$

$$\sum_{N=1}^{\infty} \exp\Big(-\Big(\frac{a_N+b_N}{a_N}\Big)^{\varepsilon} \log^{-1+\varepsilon}(a_N+b_N)\Big) < \infty, \tag{3.22}$$

then (3.8) *and* (3.9) *also hold true.*

An immediate consequence of Theorem 3.4 is

COROLLARY 3.1. *Let* $\{a_N, N \geqq 1\}$ *be a sequence of nondecreasing integers with* $1 \leqq a_N \leqq N$ *and* $\{Y_n, \mathcal{F}_n, n \geqq 1\}$ *be a martingale difference sequence satisfying* (3.15), (3.16) *and* (3.17) *for some* $\delta > 0$, $C_1 > 0$, $C_2 > 0$ *and* $\theta \geqq 0$. *Assume that*

$$P(|Y_n| \geqq \varepsilon n^{p/2}, i.o.) = 0 \quad \text{for some } \theta < p \leqq 1+\theta \text{ and for every } \varepsilon > 0, \tag{3.23}$$

and

$$a_N \geqq C_1 \frac{N^{p-\theta}}{\log N} \quad \text{for every } N \geqq 1. \tag{3.24}$$

Then we have

$$\limsup_{N\to\infty} \frac{\sum_{i=1}^{N} Y_i}{(2\sigma_{0,N}\log\log\sigma_{0,N})^{1/2}} = 1 \quad a.s.$$

$$\limsup_{N\to\infty} \frac{\sum_{i=1+N}^{N+a_N} Y_i}{(2\sigma_{N,a_N}(\log(\sigma_{0,N}/\sigma_{N,a_N}) + \log\log\sigma_{0,N}))^{1/2}} = 1 \quad a.s.$$

$$\limsup_{N\to\infty} \max_{0\leqq n\leqq N} \max_{a_N\leqq k\leqq a_N+N} \max_{1\leqq j\leqq k} \beta_{n,k}|T_n(j)| = 1 \quad a.s.$$

If, in addition, we also assume

$$\lim_{N\to\infty} \frac{\log(N/a_N)}{\log\log N} = \infty,$$

then

$$\lim_{N\to\infty} \max_{0\leqq n\leqq N} \frac{\sum_{i=n+1}^{n+a_N} Y_i}{(2\sigma_{n,a_N} \log(\sigma_{0,N}/\sigma_{n,a_N}))^{1/2}} = 1 \quad a.s.$$

and

$$\lim_{N\to\infty} \max_{0\leqq n\leqq N} \max_{a_N\leqq k\leqq a_N+N} \max_{1\leqq j\leqq k} \beta_{n,k}|T_n(j)| = 1 \quad a.s.$$

It is clear that conditions (3.2), (3.10), (3.11), (3.16) and (3.17) are superfluous if $\{Y_n, n \geqq 1\}$ are independent random variables. Hence one only needs to verify the condition (3.4) when studying the increments of partial sums of independent random variables.

We start with two preliminary lemmas.

LEMMA 3.1. *Let $\{\xi_n, \mathcal{F}_n, n \geqq 1\}$ be a square integrable martingale difference sequence. Then there exists without loss of generality (in the sense of Strassen [18], p. 333) a standard Wiener process $\{W(t), t \geqq 0\}$ and a sequence $\{\tau_i, i \geqq 1\}$ of stopping times with respect to $W(\cdot)$ such that*

$$\sum_{i=1}^{n} \xi_i = W\Big(\sum_{i=1}^{n} \tau_i\Big), \quad n \geqq 1, \tag{3.25}$$

$$E(\tau_n \mid \mathcal{F}_{n-1}) = E(\xi_n^2 \mid \mathcal{F}_{n-1}), \quad n \geqq 1, \tag{3.26}$$

and

$$E(\tau_n^k \mid \mathcal{F}_{n-1}) \leqq 2(8/\pi^2)^{k-1} k! E(|\xi_n|^{2k} \mid \mathcal{F}_{n-1}), \quad k \geqq 1. \tag{3.27}$$

This is the well-known Skorohod–Strassen embedding theorem (cf. Hall and Heyde [6], pp. 269–273; see also Philipp and Stout [12], p. 255).

LEMMA 3.2. *Let $\{\xi_n, \mathcal{F}_n, n \geqq 1\}$ be a square integrable martingale difference sequence with $|\xi_n| \leqq d_n$ for a sequence $\{d_n, n \geqq 1\}$ of non-decreasing numbers. Put*

(3.28)

$$M_n^{(i)}(t) = \exp\Big(t \sum_{j=1+n}^{i+n} (\tau_j - E(\tau_j^2|\mathcal{F}_{j-1})) - 4t^2 \sum_{j=1+n}^{i+n} d_j^2 E(\xi_j^2|\mathcal{F}_{j-1})\Big), \quad i \geqq 0.$$

Then, for every n, m and $|t| \leqq \frac{1}{4}d_{n+m}^{-2}$ but fixed, $\{M_n^{(i)}(t), \mathcal{F}_{i+n}; 0 \leqq i \leqq \leqq n+m\}$ is a non-negative supermartingale with first element $M_n^{(0)} = 1$ and for all $x > 0$

$$P\Big(\max_{0 \leqq i \leqq m} M_n^{(i)}(t) \geqq x\Big) \leqq 1/x. \tag{3.29}$$

PROOF. It follows from (3.26) and (3.27) that for every j with $1+n \leqq \leqq j \leqq n+m$

$$E(e^{t\tau_j}|\mathcal{F}_{j-1}) = 1 + tE(\tau_j|\mathcal{F}_{j-1}) + \sum_{k=2}^{\infty} \frac{t^k E(\tau_j^k|\mathcal{F}_{j-1})}{k!} \leqq$$
$$\leqq 1 + tE(\xi_j^2|\mathcal{F}_{j-1}) + 2\sum_{k=2}^{\infty} t^k \Big(\frac{8}{\pi^2}\Big)^{k-1} E(|\xi_j|^{2k}|\mathcal{F}_{j-1}) \leqq$$
$$\leqq 1 + tE(\xi_j^2|\mathcal{F}_{j-1}) + 2E\Big(\frac{t^2\xi_j^4}{1-|t|\xi_j^2}|\mathcal{F}_{j-1}\Big) \leqq$$
$$\leqq 1 + tE(\xi_j^2|\mathcal{F}_{j-1}) + 4t^2 d_j^2 E(\xi_j^2|\mathcal{F}_{j-1})$$

by the conditions $|\xi_j| \leqq d_j$ and $|t| \leqq \frac{1}{4}d_{n+m}^{-2}$. Therefore

$$E(e^{t\tau_j}|\mathcal{F}_{j-1}) \leqq \exp\big(tE(\xi_j^2|\mathcal{F}_{j-1}) + 4t^2 d_j^2 E(\xi_j^2|\mathcal{F}_{j-1})\big) \tag{3.30}$$

and

$$\begin{aligned} &E(M_n^{(i)}(t)|\mathcal{F}_{i+n-1}) = \\ &\quad = E(\exp(t(\tau_{i+n} - E(\xi_{i+n}^2|\mathcal{F}_{n+i-1})))|\mathcal{F}_{i+n-1}) \times \\ &\quad \times \exp\Big(t\sum_{j=1+n}^{n+i-1}(\tau_j - E(\xi_j^2|\mathcal{F}_{j-1})) - 4t^2\sum_{j=1+n}^{i+n} d_j^2 E(\xi_j^2|\mathcal{F}_{j-1})\Big) \leqq \\ &\quad \leqq M_n^{(i-1)}(t) \end{aligned} \tag{3.31}$$

by (3.30) for each $2 \leqq i \leqq m$, as desired. The relation (3.29) follows from Corollary 5.4.1 of Stout [17].

PROOF OF THEOREM 3.1. By Lemma 3.1, it suffices to show that

$$\limsup_{N\to\infty} \max_{0\leqq n\leqq b_N} \max_{a_N\leqq k\leqq a_N+c_N} \max_{1\leqq j\leqq k} \beta_{n,k}\Big|W\Big(\sum_{i=1}^{n+j}\tau_i\Big) - W\Big(\sum_{i=1}^{n}\tau_i\Big)\Big| \leqq 1 \text{ a.s.} \tag{3.32}$$

where $\{\tau_i, i \geqq 1\}$ is a sequence of stopping times satisfying (3.25), (3.26) and (3.27), and $\{\xi_n, n \geqq 1\} = \{Y_n, n \geqq 1\}$.

From (3.4) and $a_N \to \infty$ we have

$$\lim_{N\to\infty} \min_{0\leqq n\leqq b_N} \sigma_{n,a_N} = \infty.$$

Hence we only need to prove

$$\max_{0\leqq n\leqq b_N+c_N}\max_{1\leqq j\leqq a_N}\Big|\sum_{i=1+n}^{n+j}(\tau_i-\sigma_i)\Big|/\sigma_{n,a_N}\to 0 \text{ a.s.} \tag{3.33}$$

as $N\to\infty$, by Theorem 2.1 and Lemma 2.7. Put

$$m_k=\inf_{N\geqq k}a_N.$$

Then $m_k\to\infty$ as $k\to\infty$. In terms of (3.3) and the definition of $\mathcal{H}_k$, we have

$$\begin{aligned}
&\max_{N\geqq k}\max_{0\leqq n\leqq b_N+c_N}\max_{1\leqq j\leqq a_N}\Big|\sum_{i=1+n}^{n+j}(\tau_i-\sigma_i)\Big|/\sigma_{n,a_N}\leqq\\
&\leqq\max_{l\geqq m_k}\max_{N\in\mathcal{H}_l}\max_{0\leqq n\leqq b_N+c_N}\max_{1\leqq j\leqq a_N}\Big|\sum_{i=1+n}^{n+j}(\tau_i-\sigma_i)\Big|/\sigma_{n,a_N}\leqq\\
&\leqq\max_{l\geqq m_k}\max_{0\leqq n\leqq M_l}\max_{1\leqq j\leqq l+1}\Big|\sum_{i=1+n}^{n+j}(\tau_i-\sigma_i)\Big|/\sigma_{n,l}\leqq\\
&\leqq A\max_{l\geqq m_k}\max_{0\leqq n\leqq M_l}\max_{1\leqq j\leqq l+1}\Big|\sum_{i=1+n}^{n+j}(\tau_i-\sigma_i)\Big|/\sigma_{n,l+1}.
\end{aligned}$$

Hence (3.33) is implied by

$$\max_{0\leqq n\leqq M_l}\max_{1\leqq j\leqq l+1}\Big|\sum_{i=1+n}^{n+j}(\tau_i-\sigma_i)\Big|/\sigma_{n,l+1}\to 0 \quad \text{a.s.} \tag{3.34}$$

as $l\to\infty$. Noting that for every $0<\varepsilon<1$, by (3.26)

$$\begin{aligned}
&\Big\{\max_{0\leqq n\leqq M_l}\max_{1\leqq j\leqq l+1}\Big|\sum_{i=1+n}^{n+j}(\tau_i-\sigma_i)\Big|/\sigma_{n,l+1}\geqq 2\varepsilon\Big\}\subset\\
&\subset\Big\{\max_{0\leqq n\leqq M_l}\max_{1\leqq j\leqq l+1}\Big|\sum_{i=1+n}^{n+j}(E(Y_i^2\big|\mathcal{F}_{i-1})-\sigma_i)\Big|/\sigma_{n,l+1}\geqq\varepsilon\Big\}\cup\\
&\cup\Bigg\{\max_{0\leqq n\leqq M_l}\max_{1\leqq j\leqq l+1}\frac{\Big|\sum_{i=1+n}^{n+j}(\tau_i-E(\tau_i|\mathcal{F}_{i-1}))\Big|}{\sigma_{n,l+1}}\geqq\varepsilon,\\
&\qquad\max_{0\leqq n\leqq M_l}\frac{\sum_{i=1+n}^{n+l+1}E(Y_i^2|\mathcal{F}_{i-1})}{\sigma_{n,l+1}}\leqq 2\Bigg\}
\end{aligned}$$

and using (3.2), Lemma 2.1 and the Borel–Cantelli lemma, one can find that (3.34) will follow from

(3.35)
$$\sum_{k=1}^{\infty}\sum_{n=0}^{M_k} P\left(\max_{1\leqq j\leqq k+1}\frac{\left|\sum_{i=1+n}^{n+j}(\tau_i - E(\tau_i|\mathcal{F}_{i-1}))\right|}{\sigma_{n,k+1}} \geqq \varepsilon, \frac{\sum_{i=1+n}^{n+k+1} E(Y_i^2|\mathcal{F}_{i-1})}{\sigma_{n,k+1}} \leqq 2\right) < \infty$$

for each $0 < \varepsilon < 1$. Take

$$t = \frac{\varepsilon}{16 D_{n+k+1}^2}$$

in Lemma 3.2. Then, by (3.29) we obtain

$$P\left(\max_{1\leqq j\leqq k+1}\frac{\sum_{i=1+n}^{n+j}(\tau_i - E(\tau_i|\mathcal{F}_{i-1}))}{\sigma_{n,k+1}} \geqq \varepsilon, \frac{\sum_{i=1+n}^{n+k+1} E(Y_i^2|\mathcal{F}_{i-1})}{\sigma_{n,k+1}} \leqq 2\right) \leqq \tag{3.36}$$

(3.36)
$$\leqq P\left(\max_{1\leqq j\leqq k+1} \exp\left(t\sum_{i=1+n}^{n+j}(\tau_i - E(\tau_i|\mathcal{F}_{i-1})) - 4t^2\sum_{i=1+n}^{n+j} D_{n+i}^2 E(Y_i^2|\mathcal{F}_{i-1})\right) \geqq\right.$$
$$\left.\geqq \exp\left(t\varepsilon\sigma_{n,k+1} - 4t^2\sum_{i=1+n}^{n+k+1} D_{n+i}^2 E(Y_i^2|\mathcal{F}_{i-1})\right), \sum_{i=1+n}^{n+k+1} E(Y_i^2|\mathcal{F}_{i-1}) \leqq 2\sigma_{n,k+1}\right) \leqq$$
$$\leqq P\left(\max_{1\leqq j\leqq k+1} \exp\left(t\sum_{i=1+n}^{n+j}(\tau_i - E(\tau_i|\mathcal{F}_{i-1})) - 4t^2\sum_{i=1+n}^{n+j} D_{n+i}^2 E(Y_i^2|\mathcal{F}_{i-1})\right) \geqq\right.$$
$$\left.\geqq \exp\left(t\varepsilon\sigma_{n,k+1} + 8t^2 D_{n+k+1}^2 \sigma_{n,k+1}\right)\right) \leqq$$
$$\leqq \exp\left(-t\varepsilon\sigma_{n,k+1} + 8t^2 D_{n+k+1}^2\sigma_{n,k+1}\right) = \exp\left(-\frac{\varepsilon^2\sigma_{n,k+1}}{32 D_{n+k+1}^2}\right).$$

Similarly, we have

$$P\left(\max_{1\leqq j\leqq k+1}\frac{\sum_{i=1+n}^{n+j}(\tau_i - E(\tau_i|\mathcal{F}_{i-1}))}{\sigma_{n,k+1}} \leqq -\varepsilon, \frac{\sum_{i=1+n}^{n+k+1} E(Y_i^2|\mathcal{F}_{i-1})}{\sigma_{n,k+1}} \leqq 2\right) \leqq \tag{3.37}$$
$$\leqq \exp\left(-\frac{\varepsilon^2\sigma_{n,k+1}}{32 D_{n+k+1}^2}\right).$$

We finally conclude from (3.36) and (3.37)

$$P\left(\max_{1\leqq j\leqq k+1}\frac{|\sum_{i=1+n}^{n+j}(\tau_i - E(\tau_i|\mathcal{F}_{i-1}))|}{\sigma_{n,k+1}} \geqq \varepsilon, \frac{\sum_{i=1+n}^{n+k+1}E(Y_i^2|\mathcal{F}_{i-1})}{\sigma_{n,k+1}} \leqq 2\right) \leqq \leqq 2\exp\left(-\frac{\varepsilon^2\sigma_{n,k+1}}{32D^2_{n+k+1}}\right). \tag{3.38}$$

This proves (3.35) by (3.38) and (3.4). Now the proof of Theorem 3.1 is completed.

PROOF OF THEOREM 3.2. The conclusion is an immediate consequence of (3.33), Lemmas 2.7 and 3.1 and Theorem 2.3.

PROOF OF THEOREM 3.3. By Lemma 3.1 and Theorem 2.2, it suffices to show that

$$\max_{1\leqq j\leqq a_N}\left|\sum_{i=1+b_N}^{j+b_N}(\tau_i - \sigma_i)\right|/\sigma_{b_N,a_N} \to 0 \quad \text{a.s.} \tag{3.39}$$

and

$$\sum_{i=1}^{b_N}\tau_i \leqq 7A\sigma_{0,a_N+b_N} \quad \text{a.s.} \tag{3.40}$$

as $N\to\infty$, where $\{\tau_i, i\geqq 1\}$ is a sequence of stopping times satisfying (3.25), (3.26) and (3.27), and $\{\xi_n, n\geqq 1\} = \{Y_n, n\geqq 1\}$.

Using Lemma 2.1, (3.10), (3.11) and the Borel–Cantelli lemma, one can see that (3.39) and (3.40) follow from

$$\sum_{N=1}^{\infty}P\left(\max_{1\leqq j\leqq a_N}\frac{\left|\sum_{i=1+b_N}^{j+b_N}(\tau_i - E(\tau_i|\mathcal{F}_{i-1}))\right|}{\sigma_{b_N,a_N}} \geqq \varepsilon, \frac{\sum_{i=1+b_N}^{a_N+b_N}E(Y_i^2|\mathcal{F}_{i-1})}{\sigma_{b_N,a_N}} \leqq 2\right) < \infty \tag{3.41}$$

and

$$\sum_{N=1}^{\infty}P\left(\frac{\sum_{i=1}^{b_N}(\tau_i - E(\tau_i|\mathcal{F}_{i-1}))}{\sigma_{0,a_N+b_N}} \geqq 4A, \frac{\sum_{i=1}^{b_N}E(Y_i^2|\mathcal{F}_{i-1})}{\sigma_{0,a_N+b_N}} \leqq 2A\right) < \infty \tag{3.42}$$

for each $0<\varepsilon<1$.

Taking $|t| = \frac{\varepsilon}{16D^2_{a_N+b_N}}$ and $|t| = \frac{1}{4D^2_{b_N}}$ in Lemma 3.2, respectively, similarly to (3.38), we obtain

(3.43)
$$P\Big(\max_{1\leq j\leq a_N}\Big|\sum_{i=1+b_N}^{j+b_N}(\tau_i - E(\tau_i|\mathcal{F}_{i-1}))\Big| \geq \varepsilon\sigma_{b_N,a_N}, \sum_{i=1+b_N}^{a_N+b_N} E(Y_i^2|\mathcal{F}_{i-1}) \leq 2\sigma_{b_N,a_N}\Big) \leq 2\exp\Big(-\frac{\varepsilon^2\sigma_{b_N,a_N}}{32D^2_{a_N+b_N}}\Big).$$

and

(3.44)
$$P\Big(\sum_{i=1}^{b_N}(\tau_i - E(\tau_i|\mathcal{F}_{i-1})) \geq 4A\sigma_{0,a_N+b_N}, \sum_{i=1}^{b_N} E(Y_i^2|\mathcal{F}_{i-1}) \leq 2A\sigma_{0,a_N+b_N}\Big) \leq$$
$$\leq 2\exp(-A\sigma_{0,a_N+b_N}/(2D^2_{b_N})) \leq 2\exp(-A\sigma_{0,a_N+b_N}/(2D^2_{a_N+b_N})).$$

Now (3.43), (3.44) and (3.12) yield (3.41) and (3.42). This completes the proof of Theorem 3.3.

To finish the proof of Theorem 3.4, we need another Lemma.

LEMMA 3.3. *Let ξ be a random variable and $\mathcal{A}$ be a σ-field. Assume $E(\xi|\mathcal{A}) = 0$, a.s.. Then*

(3.45)
$$E(e^{t\xi I\{\xi\leq C\}}|\mathcal{A}) \leq \exp(t^2E(\xi^2|\mathcal{A}) + t^{2+\delta}e^{2tC}E(|\xi|^{2+\delta}|\mathcal{A}))$$

for every $0 < \delta \leq 1$, $t \geq 0$, and $C \geq 0$, where I, as usually, is the indicator function.

PROOF. We have

(3.46)
$$E(e^{t\xi I\{\xi\leq C\}}|\mathcal{A}) = 1 - tE(\xi I\{\xi > C\}|\mathcal{A}) + \frac{t^2}{2}E(\xi^2 I\{\xi \leq C\}|\mathcal{A}) + $$
$$+ E\Big(\sum_{i=3}^{\infty}\frac{(t\xi)^i I\{\xi\leq C\}}{i!}|\mathcal{A}\Big) \leq$$
$$\leq 1 + \frac{t^2}{2}E(\xi^2 I\{\xi\leq C\}|\mathcal{A}) + E\Big(\sum_{i=3}^{\infty}\frac{(t\xi)^i I\{\xi\leq C\}}{i!}|\mathcal{A}\Big).$$

Noting that $\sum_{i=3}^{\infty}\frac{x^i}{i!}$ is a monotone increasing function on $(-\infty,\infty)$, we find

(3.47)
$$E\Big(\sum_{i=3}^{\infty}\frac{(t\xi)^i I\{\xi\leq C\}}{i!}|\mathcal{A}\Big) \leq E\Big(\sum_{i=3}^{\infty}\frac{(t\xi)^i I\{0\leq\xi\leq C\}}{i!}|\mathcal{A}\Big)$$
$$\leq E((t\xi)^3 e^{t\xi} I\{0\leq\xi\leq C\}|\mathcal{A}) \leq$$
$$\leq E((t\xi)^{2+\delta}e^{2t\xi}I\{0\leq\xi\leq C\}|\mathcal{A}) \leq$$
$$\leq t^{2+\delta}e^{2tC}E(|\xi|^{2+\delta}|\mathcal{A}).$$

This proves (3.45) by (3.46) and (3.47).

PROOF OF THEOREM 3.4. Assume, without loss of generality, $0<\delta\leqq 1$. We first work with (3.5). It suffices to show that for every $0<\varepsilon<1/8$

$$\limsup_{N\to\infty}\max_{0\leqq n\leqq b_N}\max_{a_N\leqq k\leqq a_N+c_N}\max_{1\leqq j\leqq k}\beta_{n,k}|T_n(j)|\leqq 1+7\varepsilon \quad \text{a.s.} \tag{3.48}$$

Let $B>0$, $\eta=\eta(\varepsilon)>0$ with $\eta(\varepsilon)\downarrow 0$ as $\varepsilon\to 0$. Define

$$\begin{aligned}
Y_{j,1}&=Y_jI\{|Y_j|\leqq Bj^{\frac{\theta}{2}}\}-E(Y_jI\{|Y_j|\leqq Bj^{\frac{\theta}{2}}\}|\mathcal{F}_{j-1}),\\
Y_{j,2}&=Y_jI\{|Y_j|> Bj^{\frac{\theta}{2}}\}-E(Y_jI\{|Y_j|> Bj^{\frac{\theta}{2}}\}|\mathcal{F}_{j-1}),\\
Y_{j,3}&=Y_{j,2}I\{Y_{j,2}\leqq 2\eta H(j)\},\\
T_j^{(i)}(k)&=\sum_{l=1+j}^{j+k}Y_{l,i},\quad T^{(i)}(k):=T_0^{(i)}(k),\quad i=1,2,3.\\
\sigma_j^{(i)}&:=\sigma_j^{(i)}(B)=EY_{j,i}^2,\quad i=1,2.\\
\sigma_{n,k}^{(1)}&=\sum_{j=1+n}^{n+k}\sigma_j^{(1)},\\
\beta_{n,k}^{(1)}&=\{2\sigma_{n,k}^{(1)}(\log(\sigma_{0,n+k}^{(1)}/\sigma_{n,k}^{(1)})+\log\log\sigma_{n,k}^{(1)})\}^{-\frac{1}{2}}.
\end{aligned}$$

Clearly, (3.15) implies

$$\sum_{j=1+n}^{n+k}\sigma_j\asymp(n+k)^{\theta}k, \text{ uniformly on } n \text{ and } k. \tag{3.49}$$

Hence

$$\beta_{n,k}^{-2}\asymp(n+k)^{\theta}k\Big(\log\frac{n+k}{k}+\log\log((n+k)^{\theta}k)\Big), \text{ uniformly on } n \text{ and } k. \tag{3.50}$$

That is, there exists a constant $c>0$ depending only on C_1, C_2 and θ such that for any $0\leqq n<n+k$

$$\begin{aligned}
&c\Big\{(n+k)^{\theta}k\Big(\log\frac{n+k}{k}+\log\log((n+k)^{\theta}k)\Big)\Big\}^{-\frac{1}{2}}\leqq\beta_{n,k}\leqq\\
&\leqq\frac{1}{c}\Big\{(n+k)^{\theta}k\Big(\log\frac{n+k}{k}+\log\log((n+k)^{\theta}k)\Big)\Big\}^{-\frac{1}{2}}
\end{aligned} \tag{3.51}$$

and

$$c(n+k)^{\theta}k\leqq\sigma_{n,k}\frac{1}{c}(n+k)^{\theta}k. \tag{3.52}$$

Using (3.15) again, we can get

$$\sigma_j^{(2)}/\sigma_j \to 0, \tag{3.53}$$

$$\sigma_j^{(1)}/\sigma_j \to 1, \tag{3.54}$$

as $B \to \infty$, uniformly on $j \geqq 1$.

On the other hand, (3.16) and (3.17) yield

$$\lim_{B\to\infty} \lim_{j\to\infty} |E(Y_{j,1}^2|\mathcal{F}_{j-1}) - \sigma_j^{(1)}|/j^\theta = 0 \quad \text{a.s..} \tag{3.55}$$

Therefore, we can choose a constant B such that

$$(1+\varepsilon)^{-4} \leqq \sigma_j^{(1)}/\sigma_j \leqq 1, \text{ for all } j \geqq 1, \tag{3.56}$$

$$\sigma_j^{(2)}/\sigma_j \leqq c^3\varepsilon^2/(72^2 2^{\theta+1}), \text{ for all } j \geqq 1, \tag{3.57}$$

$$B > \max(C_2/(\eta H(1)), (2C_2/\eta)^{1/\delta}), \tag{3.58}$$

$$\lim_{j\to\infty} |E(Y_{j,1}^2|\mathcal{F}_{j-1}) - \sigma_j^{(1)}|/j^\theta \leqq \frac{1}{2}\varepsilon C_1 \quad \text{a.s..} \tag{3.59}$$

By (3.18) and the definition of M_k

$$\lim_{k\to\infty} \log M_k/k = 0. \tag{3.60}$$

Hence we have

$$\forall\gamma > 0, \quad \sum_{k=1}^{\infty}\sum_{n=0}^{M_k} \exp\Big(-\gamma\frac{\sigma_{n,k+1}}{B^2(k+n+1)^\theta}\Big) < \infty \tag{3.61}$$

by (3.52). According to (3.59), (3.15) and (3.56), we obtain

$$\lim_{k\to\infty} \max_{0\leqq n\leqq M_k} \max_{1\leqq j\leqq k+1} \frac{\Big|\sum_{i=1+n}^{n+j} E(Y_{i,1}^2|\mathcal{F}_{i-1})\sigma_i^{(1)}\Big|}{\sigma_{n,k+1}^{(1)}} \leqq \frac{\varepsilon}{2} \quad \text{a.s..} \tag{3.62}$$

Using (3.61), (3.62) and (3.56), along the same line of the proof of Theorem 3.1, we arrive at

$$\limsup_{N\to\infty} \max_{0\leqq n\leqq b_N} \max_{a_N\leqq k\leqq a_N+c_N} \max_{1\leqq j\leqq k} \beta_{n,k}|T_n^{(1)}(j)| \leqq 1+\varepsilon \quad \text{a.s..} \tag{3.63}$$

To finish the proof of (3.48), the rest we need to prove is

$$\limsup_{N\to\infty} \max_{0\leqq n\leqq b_N} \max_{a_N\leqq k\leqq a_N+c_N} \max_{1\leqq j\leqq k} \beta_{n,k}|T_n^{(2)}(j)| \leqq 6\varepsilon \quad \text{a.s..} \tag{3.64}$$

We first show that

$$\limsup_{N\to\infty} \max_{0\leqq n\leqq b_N} \max_{a_N\leqq k\leqq a_N+c_N} \max_{1\leqq j\leqq k} \beta^*_{n,k} T_n^{(2)}(j) \leqq 6\varepsilon c \quad \text{a.s..} \tag{3.65}$$

where $\beta^*_{n,k} = \{k(n+k)^\theta(\log\frac{n+k}{k} + \log\log(k(n+k)^\theta))\}^{-\frac{1}{2}}$.

Since

$$\begin{aligned}
&\{Y_{j,2} \geqq 2\eta H(j)\} \subset \\
&\subset \{Y_j I\{|Y_j| > Bj^{\theta/2}\} \geqq \eta H(j)\} \cup \{E(Y_j I\{|Y_j| > Bj^{\theta/2}\}|\mathcal{F}_{j-1}) \geqq \eta H(j)\} \subset \\
&\subset \{Y_j \geqq \eta H(j)\} \cup \{E(Y_j^2|\mathcal{F}_{j-1}) \geqq Bj^{\theta/2}\eta H(j)\} \subset \\
&\subset \{|Y_j| \geqq \eta H(j)\} \cup \{E(Y_j^2|\mathcal{F}_{j-1}) \geqq Bj^{\theta/2}\eta H(1)\} \subset \\
&\subset \{|Y_j| \geqq \eta H(j)\} \cup \{E(Y_j^2|\mathcal{F}_{j-1}) \geqq 2\sigma_j\}
\end{aligned}$$

by (3.58), we obtain from (3.13), (3.16) and Lemma 2.1 that

$$P(Y_{j,2} \geqq 2\eta H(j), i.o.) = 0. \tag{3.66}$$

To prove (3.65), applying (3.66) and Lemma 2.1 again, we only need to show

$$\limsup_{N\to\infty} \max_{0\leqq n\leqq b_N} \max_{a_N\leqq k\leqq a_N+c_N} \max_{1\leqq j\leqq k} \beta^*_{n,k} T_n^{(3)}(j) \leqq 6\varepsilon c \quad \text{a.s..} \tag{3.67}$$

From Lemma 2.4 it is not difficult to find that

$$\left\{\exp\left(tT_n^{(3)}(j) - \frac{t^2}{2}\sum_{i=1+n}^{n+j} E(Y_{i,2}^2|\mathcal{F}_{i-1}) - 9t^{2+\delta}\sum_{i=1+n}^{n+j} e^{4t\eta H(i)} E(|Y_i|^{2+\delta}|\mathcal{F}_{i-1})\right),\right.$$

$$\left.\mathcal{F}_{n+j}; j \geqq 1\right\}$$

is a non-negative supermartingale for every $t \geqq 0$ and $n \geqq 0$, but fixed. Hence

$$\begin{aligned}
&P\Bigg(\max_{1\leqq j\leqq k} \exp\left(tT_n^{(3)}(j) - \frac{t^2}{2}\sum_{i=1+n}^{n+j} E(Y_{i,2}^2|\mathcal{F}_{i-1}) - \right. \\
&\left.\left. -9t^{2+\delta}\sum_{i=1+n}^{n+j} e^{4t\eta H(i)} E(|Y_i|^{2+\delta}|\mathcal{F}_{i-1})\right) \geqq x\right) \leqq \frac{1}{x}
\end{aligned} \tag{3.68}$$

for every $x>0$ and $k \geqq 1$, by the well-known maximal inequality (cf. Stout [17], p. 299). We now use (3.68) and (3.58) and obtain that for every $x>0$ and $t \geqq 0$

$$P\Big(\max_{1\leqq j\leqq n} T_n^{(3)}(j) \geqq \varepsilon cx,\ \max_{n\leqq i\leqq n+k} E(|Y_i|^{2+\delta}|\mathcal{F}_{i-1}) \leqq 2C_2 i^{\frac{\theta(2+\delta)}{2}}\Big) \leqq$$

$$\leqq P\Big(\max_{1\leqq j\leqq n} T_n^{(3)}(j) \geqq \varepsilon cx,\ \sum_{i=1+n}^{n+k} E(|Y_i|^{2+\delta}|\mathcal{F}_{i-1}) \leqq 2C_2 \sum_{i=1+n}^{n+k} i^{\frac{\theta(2+\delta)}{2}},$$

$$\sum_{i=1+n}^{n+k} E(Y_{i,2}^2|\mathcal{F}_{i-1}) \leqq \eta \sum_{i=1+n}^{n+k} i^{\theta}\Big) \leqq$$

$$\leqq P\Big(\max_{1\leqq j\leqq k} \exp\Big(tT_n^{(3)}(j) - \frac{t^2}{2}\sum_{i=1+n}^{n+j} E(Y_{i,2}^2|\mathcal{F}_{i-1}) - \tag{3.69}$$

$$-9t^{2+\delta}\sum_{i=1+n}^{n+j} e^{4t\eta H(i)} E(|Y_i|^{2+\delta}|\mathcal{F}_{i-1})\Big) \geqq$$

$$\geqq \exp\Big(t\varepsilon cx - \frac{t^2}{2}\eta\sum_{i=1+n}^{n+k} i^{\theta} - 18C_2 t^{2+\delta} e^{4t\eta H(n+k)} \sum_{i=1+n}^{n+k} i^{\frac{\theta(2+\delta)}{2}}\Big)\Big) \leqq$$

$$\leqq \exp\Big(-t\varepsilon cx + \frac{t^2}{2}\eta\sum_{i=1+n}^{n+k} i^{\theta} + 18C_2 t^{2+\delta} e^{4t\eta H(n+k)} \sum_{i=1+n}^{n+k} i^{\frac{\theta(2+\delta)}{2}}\Big).$$

Let

$$\mathcal{G}_i = \{N : e^i \leqq a_N < e^{i+1}\}, \quad B_i = \max\{a_N + b_N + c_N : N \in \mathcal{G}_i\}$$

$$K_i = \log(\inf_{j\geqq i} a_j), \quad p_i = 1 + [\log(B_i/e^i)], \quad q_{i,v} = 1 + [B_i/e^{i+v}]$$

$$\mathcal{A} = \{i : \mathcal{G}_i \neq \emptyset\}.$$

Clearly, by (3.19)

$$\frac{e^{i+1} B_i^{\theta} \log^2(e^{i+1}/\log(B_i))}{H^2(30B_i)\log B_i} \geqq C_1 \tag{3.70}$$

for each $i \in \mathcal{A}$. Note that for each $l \geqq 1$

$$\max_{N\geqq l}\ \max_{0\leqq n\leqq b_N}\ \max_{a_N\leqq k\leqq a_N+c_N}\ \max_{1\leqq j\leqq k} \beta_{n,k}^* T_n^{(3)}(j) \leqq$$

$$\leqq \max_{i\geqq K_l, i\in\mathcal{A}}\ \max_{N\in\mathcal{G}_i}\ \max_{0\leqq n\leqq b_N}\ \max_{a_N\leqq k\leqq a_N+c_N}\ \max_{1\leqq j\leqq k} \beta_{n,k}^* T_n^{(3)}(j) \leqq$$

$$\leqq \max_{i \geqq K_l, i \in \mathcal{A}} \max_{0 \leqq n \leqq B_i} \max_{e^i \leqq k \leqq e^i + B_i} \max_{1 \leqq j \leqq k} \beta^*_{n,k} T_n^{(3)}(j) \leqq$$

$$\leqq \max_{i \geqq K_l, i \in \mathcal{A}} \max_{0 \leqq n \leqq B_i} \max_{0 \leqq v \leqq p_i} \max_{e^{i+v} \leqq k \leqq e^{i+v+1}} \max_{1 \leqq j \leqq k} \beta^*_{n,k} T_n^{(3)}(j) \leqq$$

$$\leqq \max_{i \geqq K_l, i \in \mathcal{A}} \max_{0 \leqq n \leqq B_i} \max_{0 \leqq v \leqq p_i} \max_{1 \leqq j \leqq e^{i+v+1}} \beta^*_{n,e^{i+v}} T_n^{(3)}(j) \leqq$$

$$\leqq \max_{i \geqq K_l, i \in \mathcal{A}} \max_{0 \leqq v \leqq p_i} \max_{0 \leqq u \leqq q_{i,v}} \max_{ue^{i+v+1} \leqq n \leqq (u+1)e^{i+v+1}} \max_{1 \leqq j \leqq e^{i+v+1}} \beta_{u,i,v} T_n^{(3)}(j) \leqq$$

$$\leqq 6 \max_{i \geqq K_l, i \in \mathcal{A}} \max_{0 \leqq v \leqq p_i} \max_{0 \leqq u \leqq 1 + q_{i,v}} \max_{1 \leqq j \leqq e^{i+v+1}} \beta_{u,i,v} T^{(3)}_{ue^{i+v+1}}(j),$$

where $\beta_{u,i,v} = (((u+1)e^{i+v})^\theta e^{i+v}(\log((u+1)(i+v)))^{-\frac{1}{2}}$. Therefore, (3.67) will follow from

(3.71) $$\limsup_{i \in \mathcal{A}, i \to \infty} \max_{0 \leqq v \leqq p_i} \max_{0 \leqq u \leqq 1 + q_{i,v}} \max_{1 \leqq j \leqq e^{i+v+1}} \beta_{u,i,v} T^{(3)}_{ue^{i+v+1}}(j) \leqq \varepsilon c \quad \text{a.s.}$$

The latter is implied by, according to (3.17), Lemma 2.1 and the Borel–Cantelli lemma

(3.72) $$\sum_{i \in \mathcal{A}} \sum_{0 \leqq v \leqq p_i} \sum_{0 \leqq u \leqq 1 + q_{i,v}} P_{i,v,u} < \infty$$

where

$$P_{i,v,u} = P\Bigg(\max_{1 \leqq j \leqq e^{i+v+1}} \beta_{u,i,v} T^{(3)}_{ue^{i+v+1}}(j) \geqq \varepsilon c,$$

$$\max_{ue^{i+v+1} \leqq l \leqq (u+1)e^{i+v+1}} \frac{E(|Y_l|^{2+\delta} \mid \mathcal{F}_{l-1})}{l^{\theta(2+\delta)/2}} \leqq 2C_2 \Bigg).$$

Recalling the definitions of p_i and $q_{i,v}$, we have

(3.73) $$(u+1)e^{i+v+1} \leqq 30 B_i$$

for every $0 \leqq u \leqq 1 + q_{i,v}$, $0 \leqq v \leqq p_i$. Let

(3.74) $$t = \frac{72}{\varepsilon c} \Big(\frac{\log((u+1)(i+v))}{((u+1)e^{i+v})^\theta e^{i+v}} \Big)^{1/2},$$

(3.75) $$\eta = \min\Big(\frac{\varepsilon c C_1^{\frac{1}{2}} \delta}{1200}, \frac{(\varepsilon c)^2 e^{-\theta}}{150} \Big).$$

Then

$$4t\eta H((u+1)e^{i+v+1}) \leqq$$

$$\leqq \frac{300e^{\theta}\eta H((u+1)e^{i+v+1})}{\varepsilon c((u+1)e^{i+v+1})^{\theta/2}}\Big(\frac{\log((u+1)(i+v))}{e^{i+v}}\Big)^{1/2} \leqq$$

$$\leqq \frac{300\eta H(30B_i)}{\varepsilon c B_i^{\theta/2}}\Big(\frac{\log B_i}{e^i}\Big)^{1/2} \leqq \tag{3.76}$$

$$\leqq \frac{300\eta \log(e^i/\log B_i)}{\varepsilon c C_1^{1/2}} \leqq \frac{1}{4}\delta \log(e^i/\log B_i)$$

by (3.70), (3.74) and (3.75).

From (3.69), (3.76), (3.75) and (3.60), we conclude

$$P_{i,v,u} \leqq$$

$$\leqq \exp\Big(-\frac{t\varepsilon c}{\beta^*_{u,i,v}} + \frac{t^2}{2}\eta \sum_{j=1+ue^{i+v+1}}^{(u+1)e^{i+v+1}} j^{\theta} +$$

$$+18C_2 t^{2+\delta} e^{4t\eta H((u+1)e^{i+v+1})} \sum_{j=1+ue^{i+v+1}}^{(u+1)e^{i+v+1}} j^{\frac{\theta(2+\delta)}{2}}\Big) \leqq \tag{3.77}$$

$$\leqq \exp\Big(-t\varepsilon c(((u+1)e^{i+v})^{\theta} e^{i+v} \log((u+1)(i+v)))^{1/2} +$$

$$+t^2\eta((u+1)e^{i+v+1})^{\theta} e^{i+v} +$$

$$+ 18C_2 t^{2+\delta} e^{4t\eta H((u+1)e^{i+v+1})}((u+1)e^{i+v+1})^{\theta(2+\delta)/2} e^{i+v+1}\Big) \leqq$$

$$\leqq \exp\Big(-72\Big(1-\frac{72\eta e^{\theta}}{(\varepsilon c)^2}\Big)\log((u+1)(i+v)) +$$

$$+18C_2 t^2((u+1)e^{i+v+1})^{\theta} e^{i+v+1}\frac{72}{\varepsilon c}\Big(\frac{\log B_i}{e^{i+v}}\Big)^{\delta/4}\Big) \leqq$$

$$\leqq \exp(-10\log((u+1)(i+v))) \leqq (u+1)^{-10}(i+v)^{-10}$$

for every $i \in \mathcal{A}$ sufficiently large.

This proves that (3.72) holds. We now conclude that (3.65) is true. Similarly, we have

$$\limsup_{N\to\infty} \max_{0\leqq n\leqq b_N} \max_{a_N\leqq k\leqq a_N+c_N} \max_{1\leqq j\leqq k} \beta^*_{n,k}(-T_n^{(2)}(j)) \leqq 6\varepsilon c \quad \text{a.s..} \tag{3.78}$$

Now (3.64) is proved by (3.65), (3.78) and (3.52) and so is (3.5).

We next prove (3.6) and (3.7). It is easy to see that (3.15), (3.20) and (3.21) imply conditions (2.6) and (2.7). Applying Theorem 3.2, (3.56) and (3.61), one can get

$$\limsup_{N\to\infty} \alpha_{b_n,N} T^{(1)}_{b_N}(a_N) \geqq (1+\varepsilon)^{-1} \quad \text{a.s.}$$

which together with (3.64) yields

$$\limsup_{N\to\infty} \alpha_{b_N,N} T_{b_N}(a_N) \geqq (1+\varepsilon)^{-1} - 7\varepsilon \quad \text{a.s.} \tag{3.79}$$

Now (3.6) and (3.7) follow from (3.5), (3.79) and the arbitrariness of ε.

In addition, if (3.22) is satisfied, then for every $0<\gamma<1$

$$\begin{aligned} \sum_{j=0}^{[b_N/a_N]} \Big(\frac{\sigma_{ja_N,a_N}}{\sigma_{0,b_N+a_N}\log \sigma_{0,b_N+a_N}}\Big)^{1-\gamma} &\geqq \\ \geqq K \sum_{j=0}^{[b_N/a_N]} \Big(\frac{(j+1)^\theta a_N^{1+\theta}}{(b_N+a_N)^{\theta+1}\log(b_N+a_N)}\Big)^{1-\gamma} &\geqq \\ \geqq K\Big(\frac{b_N+a_N}{a_N}\Big)^\gamma \log^{-1+\gamma}(b_N+a_N)& \end{aligned}$$

for every $N \geqq 1$ and for some positive K.

Applying Theorem 3.2 to $\{Y_{j,1}, j\geqq 1\}$, in combination with (3.64), yields (3.8) and (3.9). This completes the proof of Theorem 3.4.

REMARK 3.1. If Conditions (3.2), (3.3) and (3.4) are replaced by

$$\max_{0\leqq n\leqq b_N+c_N}\ \max_{1\leqq j\leqq a_N} \frac{\Big|\sum_{i=1+n}^{n+j} E(Y_i^2|\mathcal{F}_{i-1}) - EY_i^2\Big|}{\sigma_{n,a_N}} \to 0 \quad \text{a.s.} \tag{3.80}$$

$$\sum_{N=1}^{\infty}\sum_{n=0}^{b_N+c_N} \exp\big(-\varepsilon\sigma_{n,a_N}/D^2_{n+a_N}\big) < \infty. \tag{3.81}$$

Then, the conclusion of Theorems 3.1 and 3.2 remain valid.

REMARK 3.2. According to Remark 2.1, if Condition (2.11) in Theorem 3.2 is replaced by (2.48), then

$$\lim_{N\to\infty}\max_{1\leqq n\leqq j_N} \alpha'_{n,N}(T(a(n,N)) - T(a(n-1,N))) = 1 \quad \text{a.s.} \tag{3.82}$$

where $\alpha'_{n,N}$ is defined as in Section 2.

REMARK 3.3. Let $\{a(j,N)\}$ be an array of integers satisfying $a(0,N) = 0$, $a(1,N) = a_N$ and $a_N \leqq a(j,N) - a(j-1,N) \leqq a_N + b_N$ for every $j \geqq 1$. If Condition (3.22) in Theorem 3.4 is replaced by

$$\sum_{N=1}^{\infty} \exp\Big(-\sum_{i=0}^{j_N} \frac{(a(i+1,N)^{\theta}(a(i+1,N)-a(i,N)))^{1-\varepsilon}}{(a_N+b_N)^{1+\theta}\log^{1-\varepsilon}(a_N+b_N)}\Big) < \infty \tag{3.83}$$

then (3.82) holds true.

4. How big are the increments of partial sums of independent random variables

There has been a great amount of work on the increments of partial sums of independent random variables. One can refer to Hanson and Russo [8] and Lin [10], [11]. The general result by now is due to Shao [16] via the strong approximation theorem. As a direct application of our theorems in Section 3, we arrange the corresponding results for independent random variables in this section.

Assume throughout this section $\{a_N, N \geqq 1\}$, $\{b_N, n \geqq 1\}$ and $\{c_N, N \geqq 1\}$ are sequences of integer numbers and $\{X_n, n \geqq 1\}$ are independent random variables with $EX_n = 0$ and $EX_n^2 < \infty$ for each $n \geqq 1$. Put $\sigma_n = EX_n^2$, $S_n(k) = \sum_{i=1+n}^{n+k} X_i$. Let $\sigma_{n,k}$, $\beta_{n,k}$ and $\alpha_{n,N}$ be as in (1.4), (1.5) and (1.6), respectively.

THEOREM 4.1. *Let $H(x) : [0,\infty) \to (0,\infty)$ be a nondecreasing function. Assume there exist constants $\delta > 0$, $C_1 > 0$, $C_2 > 0$ and $\theta \geqq 0$ such that*

$$\sum_{n=1}^{\infty} P(|X_n| \geqq \varepsilon H(n)) < \infty \text{ for every } \varepsilon > 0, \tag{4.1}$$

$$H^2(x)/x^{\theta} \text{ is nondecreasing,} \tag{4.2}$$

$$C_1 n^{\theta} \leqq \sigma_n \leqq (E|X_n|^{2+\delta})^{\frac{2}{2+\delta}} \leqq C_2 n^{\theta} \text{ for each } n \geqq 1, \tag{4.3}$$

$$\lim_{n\to\infty} a_N = \infty, \tag{4.4}$$

$$H^2(x) \geqq C_1 x^{\theta} \log x \text{ for each } x \geqq 1, \tag{4.5}$$

$$\frac{a_N(a_N+b_N+c_N)^{\theta}\log^2 a_N}{H^2(30(a_N+b_N+c_N))\log(a_N+b_N+c_N)} \geqq C_1 \text{ for each } N \geqq 1. \tag{4.6}$$

Then

$$\limsup_{N\to\infty} \max_{0\leqq n\leqq b_N} \max_{a_N\leqq k\leqq a_N+c_N} \max_{1\leqq j\leqq k} \beta_{n,k}|S_n(j)| \leqq 1 \ a.s.. \tag{4.7}$$

If we also assume (3.20) *and* (3.21) *are satisfied, then*

$$\limsup_{N\to\infty} \alpha_{b_N,N} S_{b_N}(a_N) = 1 \ a.s. \tag{4.8}$$

$$\limsup_{N\to\infty} \max_{0\leqq n\leqq b_N} \max_{a_N\leqq k\leqq a_N+c_N} \max_{1\leqq j\leqq k} \beta_{n,k}|S_n(j)| = 1 \ a.s.. \tag{4.9}$$

Furthermore, if (3.22) *is satisfied, then*

$$\lim_{N\to\infty} \max_{0\leqq n\leqq b_N} \alpha_{n,N} S_n(a_N) = 1 \ a.s. \tag{4.10}$$

$$\lim_{N\to\infty} \max_{0\leqq n\leqq b_N} \max_{a_N\leqq k\leqq a_N+c_N} \max_{1\leqq j\leqq k} \beta_{n,k}|S_n(j)| = 1 \ a.s.. \tag{4.11}$$

When $\theta = 0$, we have the following precise one.

THEOREM 4.2. *Let $H(x):[0,\infty)\to(0,\infty)$ be a nondecreasing function. Assume there exist constants $\delta>0$, $C_1>0$, $C_2>0$ and $0<\alpha<1$ such that*

$$\sum_{n=1}^{\infty} P(|X_n| \geqq H(n)) < \infty \tag{4.12}$$

$$C_1 \leqq \sigma_n \leqq (E|X_n|^{2+\delta})^{\frac{2}{2+\delta}} \leqq C_2 \ \textit{for each } n \geqq 1, \tag{4.13}$$

$$E(invH(|X_n|))^{\alpha} \leqq C_2 \ \textit{for every } n \geqq 1, \tag{4.14}$$

$$x/\log(invH(x)) \ \textit{is nondecreasing on } [1,\infty), \tag{4.15}$$

$$\lim_{N\to\infty} a_N/\log(a_N+b_N+c_N) = \infty, \tag{4.16}$$

$$\lim_{N\to\infty} \frac{a_N \log(a_N+b_N+c_N)}{H^2(a_N+b_N+c_N)} = \infty. \tag{4.17}$$

Then, (4.7) *holds. If, in addition,* (3.20) *and* (3.21) *are satisfied, then* (4.8) *and* (4.9) *remain valid. Furthermore, if* (3.22) *is satisfied, then* (4.10) *and* (4.11) *are true.*

From Theorem 4.2 we can get a series of existing and new results.

COROLLARY 4.1. *Let* $0 < r \leqq 1$. *Assume*

$$\forall n \geqq 1, \quad \sigma_n \geqq \sigma > 0, \tag{4.18}$$

$$\forall n \geqq 1. \quad Ee^{t_0|X_n|^r} \leqq M < \infty \text{ for some } t_0 > 0 \tag{4.19}$$

$$\forall N \geqq 1, \quad a_N \leqq C_2 a_{N-1} \text{ for some } C_2 > 0 \tag{4.20}$$

$$\lim_{N\to\infty} a_N = \infty \text{ and } \lim_{N\to\infty} \varepsilon_N = 0. \tag{4.21}$$

Then

$$\limsup_{N\to\infty} \max_{0 \leqq n \leqq \exp(\varepsilon_N a_N^{\frac{r}{2-r}})} \max_{a_N \leqq k \leqq \exp(\varepsilon_N a_N^{\frac{r}{2-r}})} \max_{1 \leqq j \leqq k} \beta_{n,k}|S_n(j)| = 1 \text{ a.s.}. \tag{4.22}$$

COROLLARY 4.2. *Let* $0 < r \leqq 1$. *Assume that* (4.18) *and* (4.19) *are satisfied. Moreover, suppose that*

$$\forall N \geqq 2, \quad C_1 a_{N-1} \leqq a_N \leqq C_2 a_{N-1} \text{ for some } C_1 > 0, C_2 < \infty \tag{4.23}$$

$$\lim_{N\to\infty} \frac{a_N}{\log^{\frac{2}{r}-1} N} = \infty. \tag{4.24}$$

Then

$$\limsup_{N\to\infty} \alpha'_N S_N(a_N) = 1 \text{ a.s.} \tag{4.25}$$

$$\limsup_{N\to\infty} \alpha''_N S_{N-a_N}(a_N) = 1 \text{ a.s.} \tag{4.26}$$

$$\limsup_{N\to\infty} \max_{0 \leqq n \leqq N^p} \max_{a_N \leqq k \leqq a_N + N^p} \max_{1 \leqq j \leqq k} \beta_{n,k}|S_n(j)| = 1 \text{ a.s.} \tag{4.27}$$

for every $p > 0$, *where*

$$\alpha'_N = (2\sigma_{N,a_N}(\log(N/a_N) + \log\log(N + a_N)))^{-1/2},$$

$$\alpha''_N = (2\sigma_{N-a_N,a_N}(\log(N/a_N) + \log\log N))^{-1/2},$$

here $S_j(n) := S_0(n)$ *and* $\sigma_{j,n} := \sigma_{0,n}$ *if* $j < 0$.

COROLLARY 4.3. *Let* $H(x)$ *be a nondecreasing function. Assume that there exist constants* $C_1 > 0$, $C_2 > 0$, $\delta > 0$ *such that* (4.1), (4.13), (4.23),

$$\forall x \geqq 1, \quad H(x) \geqq C_1 x^\delta \tag{4.28}$$

and

$$\text{(4.29)} \qquad \forall N \geqq 2, \quad a_N \geqq C_1 \frac{H^2(N)}{\log N}$$

are satisfied. Then (4.25) *and* (4.26) *as well as*

$$\text{(4.30)} \qquad \limsup_{N\to\infty} \max_{0\leqq n\leqq N} \max_{a_N\leqq k\leqq a_N+N} \max_{1\leqq j\leqq k} \beta_{n,k}|S_n(j)| = 1 \ a.s.$$

hold true. If, in addition, we assume

$$\text{(4.31)} \qquad \lim_{N\to\infty} \frac{\log(N/a_N)}{\log\log N} = \infty$$

then

$$\text{(4.32)} \qquad \lim_{N\to\infty} \max_{0\leqq n\leqq N} \alpha'_{n,N} S_n(a_N) = 1 \ a.s.$$

$$\text{(4.33)} \qquad \lim_{N\to\infty} \max_{0\leqq n\leqq N} \max_{a_N\leqq k\leqq a_N+N} \max_{1\leqq j\leqq k} \beta_{n,k}|S_n(j)| = 1 \ a.s.$$

where $\alpha'_{n,N} = (2\sigma_{n,a_N}\log(n/a_N))^{-1/2}$.

PROOF OF THEOREM 4.1. Clearly, (4.5) and (4.6) imply

$$\text{(4.34)} \qquad a_N^{1/2}\log a_N \geqq C_1 \log(a_N + b_n + c_N).$$

Hence (3.18) is satisfied by (4.4). A combination of (4.34) and (4.6) yields (3.19). Now the conclusion follows from Theorem 3.4 immediately.

PROOF OF THEOREM 4.2. The proof is along the same lines of that of Theorem 3.4 with the next Lemma 4.1 instead of Lemma 3.3.

LEMMA 4.1. *Let X be a random variable with $EX = 0$. Assume $a > 0$, $0 < \alpha \leqq 1$ and $\log(invH(x))/x$ is nondecreasing on $(0,\infty)$. Then*

$$Ee^{tXI\{X\leqq a\}} \leqq \exp\left(\frac{t^2}{2}EX^2 + t^{2+\alpha/2}(E|X|^{2+\alpha})^{\frac{4+\alpha}{4+2\alpha}}(E(invH(|X|))^\alpha)^{\frac{\alpha}{4+4\alpha}}\right)$$

for all $0 < ta \leqq \frac{\alpha^2}{10}\log(invH(a))$.

PROOF. The proof is completely similar to that Lemma 3.3 and so is omitted.

PROOF OF COROLLARY 4.1. Let $H(x) = (\frac{2}{t_0}\log x)^{1/r}$, $b_N = c_N = [\exp(\varepsilon_N a_N^{r/(2-r)})] + 1$ Then the left-hand side of (4.22)$\leqq 1$ a.s. by Theorem 4.2. Put $b_N = c_N = 0$. Then the left-hand side of (4.22)$\geqq 1$ a.s. by Theorem 4.2 again, as desired.

PROOF OF COROLLARY 4.2. This is a sequence of Corollary 4.1 and Theorem 4.2.

PROOF OF COROLLARY 4.3. The result follows from Theorem 4.1 easily.

ACKNOWLEDGEMENTS. This work was one part of my Ph.D dissertation at University of Science and Technology of China, Hefei, Anhui, P.R. China. The author wishes to express his deeply gratitude to his supervisor Professor Xiru Chen for his guidance and encouragement. Thanks also to Professor Chun Su and Lincheng Zhao for their constant encouragement and help in preparing the dissertation. I am also grateful to Professor Chuanrong Lu and Zhengyan Lin of Hangzhou University for their guidance.

REFERENCES

[1] BOOK, S. A. and SHORE, T. R., On large intervals in the Csörgő–Révész theorem on increments of a Wiener process, *Z. Wahrsch. Verw. Gebiete* **46** (1978/79), 1–11. *MR* **80h**:60052

[2] CHEN, G. J., KONG, F. C. and LIN, Z. Y., Answers to some questions about increments of a Wiener process, *Ann. Probab.* **14** (1986), 1252–1261. *MR* **88c**:60067

[3] CSÁKI, E. and RÉVÉSZ, P., How big must be the increments of a Wiener process?, *Acta Math. Acad. Sci. Hungar.* **33** (1979), 37–49. *MR* **80d**:60102

[4] CSÖRGŐ, M. and RÉVÉSZ, P., How big are the increments of a Wiener process ?, *Ann. Probab.* **7** (1979), 731–737. *MR* **80g**:60025

[5] CSÖRGŐ M. and RÉVÉSZ, P., *Strong approximations in probability and statistics*, Probability and Mathematical Statistics, Academic Press, New York, 1981. *MR* **84d**:60050

[6] HALL, P. and HEYDE, C.C., *Martingale limit theory and its application*, Probability and Mathematical Statistics, Academic Press, New York, 1980. *MR* **83a**:60001

[7] HANSON, D.L. and RUSSO, P., Some results on increments of the Wiener process with applications to lag sums of iid random variables, *Ann. Probab.* **11** (1983), 609–623. *MR* **85c**:60127

[8] HANSON, D. L. and RUSSO, P., Some limit results for lag sums of independent, non-iid, random variables, *Z. Wahrsch. Verw. Gebiete* **68** (1985), 425–445. *MR* **86c**:60043

[9] LIN, Z. Y., A result on the increment of partial sums, *Appl. Math. J. Chinese Univ.* **1** (1986), 7–16.

[10] LIN, Z. Y., On Csörgő–Révész's increments of sums of non-iid random variables, *Sci. Sinica Ser. A* **30** (1987), 921–931. *MR* **89i**:60069

[11] LIN, Z. Y., On increments of sums of independent non-identically distributed random variables, *Sci. Sinica Ser. A* **31** (1988), 927–937. *MR* **89m**:60072

[12] PHILIPP, W. and STOUT, W. F., Invariance principles for martingales and sums of independent random variables, *Math. Z.* **192** (1986), 253–264. *MR* **88c**:60094

[13] SHAO, Q. M., A remark on increments of a Wiener process, *J. Math. (Wuhan)* **6** (1986), 175–182. *MR* **88d**:60099

[14] SHAO, Q. M., On the increments of sums of independent random variables, *Chinese J. Appl. Probab. Statist.* **5** (1989), 117–126.

[15] SHAO, Q. M., Limit theorems for sums of dependent and independent random variables, Ph.D Dissertation, Univ. Sci. and Tech. of China, P. R. China, 1989.

[16] SHAO, Q. M., Strong approximation for independent random variables and its applications, 1991 (unpublished manuscript).

[17] STOUT, W. F., *Almost sure convergence*, Probability and Mathematical Statistics, Vol. 24, Academic Press, New York, 1974. *MR* **56**#13334

[18] Strassen, V. A., Almost sure behaviour of sums of independent random variables and martingales, *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability* (Berkeley, Calif., 1965/66), Vol. II: Contributions to Probability Theory, Part 1, Univ. California Press, Berkeley, Calif., 1967, 315–343. *MR* **35** #4969

(Received August 6, 1991)

DEPARTMENT OF MATHEMATICS
HANGZHOU UNIVERSITY
HANGZHOU, ZHEJIANG
PEOPLE'S REPUBLIC OF CHINA

Current address:

DEPARTMENT OF MATHEMATICS
NATIONAL UNIVERSITY OF SINGAPORE
LOWER KENT RIDGE ROAD
SINGAPORE 0511
REPUBLIC OF SINGAPORE

e-mail: matsqm@ leonis.nus.sg

Studia Scientiarum Mathematicarum Hungarica

Editor-in-Chief

D. Szász

Volume 29

Akadémiai Kiadó, Budapest
1994

STUDIA SCIENTIARUM MATHEMATICARUM HUNGARICA

Volume 29

CONTENTS

RECENTLY ACCEPTED PAPERS

WORACEK, H., A remark on affine complete algebras

NATHANSON, M. B. and SÁRKÖZY, A., Metric theorems on minimal bases and maximal nonbases

SANH, N. V., Chain conditions on direct summands

HIAI, F., PETZ, D. and TÓTH, G., Curvature in the geometry of canonical correlation

DELLAVECCHIA, B., MASTROIANNI, G. and VÉRTESI, P., Mean convergence of extended Hermite interpolation of higher order

KOMJÁTH, P., A note on set mappings with meager images

DEÁK, J., Simultaneous extensions

SZABÓ, L. and TALATA, I., An illumination problem for convex polyhedra

Typeset by TypoTEX Ltd., Budapest
PRINTED IN HUNGARY
Akadémiai Kiadó és Nyomda Vállalat, Budapest

CONTENTS

HU ISSN 0081-6906