ANNALES
UNIVERSITATIS SCIENTIARUM BUDAPESTINENSIS DE ROLANDO EÖTVÖS NOMINATAE

SECTIO MATHEMATICA

TOMUS XXVIII.
1985

1986

ANNALES
UNIVERSITATIS SCIENTIARUM BUDAPESTINENSIS DE ROLANDO EÖTVÖS NOMINATAE

SECTIO BIOLOGICA
incepit anno MCMLVII

SECTIO CHIMICA
incepit anno MCMLIX

SECTIO CLASSICA
incepit anno MCMXXIV

SECTIO COMPUTATORICA
incepit anno MCMLXXVIII

SECTIO GEOGRAPHICA
incepit anno MCMLXVI

SECTIO GEOLOGICA
incepit anno MCMLVII

SECTIO HISTORICA
incepit anno MCMLVII

SECTIO IURIDICA
incepit anno MCMLIX

SECTIO LINGUISTICA
incepit anno MCMLXX

SECTIO MATHEMATICA
incepit anno MCMLVIII

SECTIO PAEDAGOGICA ET PSYCHOLOGICA
incepit anno MCMLXX

SECTIO PHILOLOGICA
incepit anno MCMLVII

SECTIO PHILOLOGICA HUNGARICA
incepit anno MCMLXX

SECTIO PHILOLOGICA MODERNA
incepit anno MCMLXX

SECTIO PHILOSOPHICA ET SOCIOLOGICA
incepit anno MCMLXII

SPLINE FUNCTIONS AND CAUCHY PROBLEMS, VI.

APPROXIMATE SOLUTION OF THE DIFFERENTIAL EQUATION $y^{(n)} = f(x, y)$ WITH SPLINE FUNCTIONS

By

THARWAT FAWZY

Suez Canal University, Ismailia, Egypt

(Received September 13, 1979)

1. Introduction and description of the method

In the recent papers [1]–[7] the approximate solution by spline functions of differential equations with given initial value conditions has been studied. In this paper a method to approximate the solution of the initial value problem $y^{(n)} = f(x, y)$ by spline function is given, which generalizes the results of [7].

Consider

$$y^{(n)} = f(x, y) \tag{1}$$

where $f \in C^r([0, b] \times \mathbf{R})$ and $r \in I^+$. We assign to equation (1) the initial conditions:

$$y(0) = y_0,\ y(0) = y_0,\ \ldots,\ y^{(n-1)} = y_0^{(n-1)}. \tag{2}$$

We can define the total q-th derivative of f w.r.t. x, expressed as a function of x and y only, as

$$\frac{d^q}{dx^q} f(x, y) = f^{(q)}(x, y), \qquad q = 0, \quad 1, \ldots, r \tag{3}$$

which could be obtained from the algorithm:

$$f^{(0)}(x, y) = f(x, y)$$

for $\alpha = 0, 1, \ldots, r-1$

$$f^{(\alpha+1)}(x, y) = \frac{\partial}{\partial x} f^{(\alpha)}(x, y) + f(x, y) \frac{\partial}{\partial y} f^{(\alpha)}(x, y)$$

where

$$\frac{d^\alpha}{dx^\alpha} f(x, y(x)) = f^\alpha(x, y(x)).$$

We assume that $f: \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ is defined and continuous with its first, second, ..., r-th derivatives in

$$D: |x - x_0| < \alpha^*, \quad |y^{(i)} - y_0^{(i)}| < \beta_i, \quad i = 0, 1, \ldots, (n-1)$$

and in what follows the interval $[0, b]$ is in D.

We also assume for (x, y), (x, y_1) and (x, y_2) in D

$$|f^{(q)}(x, y)| \leq M, \quad q = 0, 1, \ldots, r$$

and the Lipschitz condition

$$|f^{(q)}(x, y_1) - f^{(q)}(x, y_2)| \leq L|y_1 - y_2|, \quad q = 0, 1, \ldots, r. \tag{4}$$

Let $y: [0, b] \to \mathbf{R}$ be the unique solution of (1)–(2). Our purpose is to construct a polynomial spline function of degree $m \leq 2n + 2r + 1$ approximating the solution y on the interval $[0, b]$. This spline function will be denoted by $s_\Delta(x)$ where Δ is the mesh points

$$\Delta: 0 = x_0 < x_1 < \ldots < x_k < x_{k+1} < \ldots < x_N = b$$

and $x_{k+1} - x_k = h \qquad (k = 0, 1, \ldots, N-1)$.

If we integrate equation (1) $(n-i)$ times, $i = 0, 1, \ldots, (n-1)$ from x_k to x where $x_k \leq x \leq x_{k+1}$ and then putting $x = x_{k+1}$ we get

$$y_{k+1}^{(i)} = y^{(i)}(x_{k+1}) = \sum_{j=0}^{n-i-1} \frac{y_k^{(i+j)}}{j!} h^j + \tag{5}$$

$$+ \underbrace{\int_{x_k}^{x_{k+1}} \cdots \int_{x_k}^{t_{n-i-1}}}_{(n-i)\text{ times}} f(t_{n-i}, y(t_{n-i}))\, dt_{n-i} \ldots dt_1.$$

The relation (5) gives the values of the exact solution and its i-th derivatives at $x = x_{k+1}$ where $i = 0, 1, \ldots, (n-1)$. If $i = n, n+1, \ldots, n+r$ then the exact values of the higher derivatives, using the definition (3), will be

$$y_{k+1}^{(n+q)} = f^{(q)}(x_{k+1}, y_{k+1}), \quad q = 0, 1, \ldots, r, \tag{6}$$

and the corresponding approximate values are defined as

$$\bar{y}_{k+1}^{(n+q)} = f^{(q)}(x_{k+1}, \bar{y}_{k+1}), \quad q = 0, 1, \ldots, r \tag{7}$$

where $\bar{y}_{k+1}$ is the approximate value of y_{k+1} and it can be obtained from the following defintion:

We define the approximate value of $y_{k+1}^{(i)}$ as $\bar{y}_{k+1}^{(i)}$ and it will be given from the relation:

$$\bar{y}_{k+1}^{(i)} = \sum_{j=0}^{n-i-1} \frac{\bar{y}_k^{(i+j)}}{j!} h^j +$$

$$+ \underbrace{\int_{x_k}^{x_{k+1}} \cdots \int_{x_k}^{t_{n-i-1}}}_{(n-i)\text{ times}} f(t_{n-i}, y_k^*(t_{n-i}))\, dt_{n-i} \ldots dt_1 \tag{8}$$

where $i = 0, 1, \ldots, (n-1)$ and

$$y_k^*(t) = \sum_{j=0}^{n+r} \frac{\bar{y}_k^{(j)}}{j!} (t-x_k)^j, \quad x_k \leq t \leq x_{k+1}. \tag{9}$$

Here, it is convenient to write down the Taylor polynomial of the exact solution for $x_k \leq t \leq x_{k+1}$ as

$$y(t) = \sum_{j=0}^{n+r-1} \frac{y_k^{(j)}}{j!} \cdot (t-x_k)^j + \frac{y^{(n+r)}(\xi_k)}{(n+r)!} (t-x_k)^{n+r}, \tag{10}$$

$x_k < \xi_k < x_{k+1}$, which will be used later.

We start the calculations by using the substitutions

$$\bar{y}_0^{(j)} = y_0^{(j)}, \quad j = 0, 1, \ldots, n+r.$$

2. Error estimations

Let $e_{k+1}^{(j)}$ denote the error at any point x_{k+1} of Δ. i. e.:

$$e_{k+1}^{(i)} = |y_{k+1}^{(i)} - \bar{y}_{k+1}^{(i)}| \tag{11}$$

$i = 0, 1, \ldots, n+r$ and $k = 0, 1, \ldots, N-1$, then we can introduce the following lemma:

LEMMA 1. For $n \leq i \leq n+r$ the inequality

$$e_{k+1}^{(i)} \leq L e_{k+1} \tag{12}$$

is true for all $k = 0, 1, \ldots, (N-1)$.

PROOF. From (11), using (6), (7) and the Lipschitz condition (4) it is easy to get the result.

Now, our main task is to seek about the errors $e_{k+1}^{(i)}$ for $0 \leq i \leq n-1$ only, and for this purpose we introduce the following definition regarding the inequality of matrices:

DEFINITION 1. Let $A \equiv [a_{ij}]$ and $B = [b_{ij}]$ be two matrices of the same order. Then we say that

$$A \leq B$$

iff:

i) a_{ij} and b_{ij} are non negative numbers.
ii) $a_{ij} \leq b_{ij}$ for all i, j.

THEOREM 1. *Let $y_{k+1}^{(i)}$ $(i = 0, 1, \ldots, n-1)$ be the exact values of the solution of (1)–(2) and its derivatives at x_{k+1}. If the corresponding approximate values $\bar{y}_{k+1}^{(i)}$ are given by the formula (8), then the error is bounded by the inequality*

$$e_{k+1}^{(i)} = |y_{k+1}^{(i)} - \bar{y}_{k+1}^{(i)}| \leq c_i w_r(h) h^{n+r}$$

which holds for all $k = 0, 1, \ldots, N-1$. Here c_i are constants independent of h and $w_r(h)$ is the modulus of continuity of $y^{(n+r)}(x)$.

PROOF. By using (11), (5), (8), the Lipschitz condition (4), the expansion (9) and (10) it is easy to get

$$e_{k+1}^{(i)} \leq \sum_{j=0}^{n-i-1} \frac{e_k^{(i+j)}}{j!} h^j + L \sum_{j=0}^{n+r} \frac{e_k^{(j)}}{(j+n-i)!} h^{j+n-i} +$$

$$+ \frac{L}{(2n+r-i)!} w_r(h) h^{2n+r-i}. \tag{13}$$

If we used (12), then last inequality can be written in the form

$$\begin{aligned} e_{k+1}^{(i)} \leq e_k^{(i)}(1+a_{ii}h) + a_{i0}e_k h^{n-i} + a_{i1}e_k h^{n-i+1} + \ldots + a_{ii-1}e_k^{(i-1)}h^{n-1} + 0 + \\ + a_{ii+1}e_k^{(i+1)} + a_{ii+2}e_k^{(i+2)}h^2 + \ldots + a_{ii+s}e_k^{(i+s)}h^s + \ldots + \\ + a_{in-1}e_k^{(n-1)}h^{n-i-1} + b_i w_r(h) h^{2n+r-i}. \end{aligned} \tag{14}$$

Using definition 1, the inequality (14) for $i = 0, 1, \ldots, n-1$ takes the matrix form

$$\begin{bmatrix} e_{k+1} \\ e'_{k+1} \\ \vdots \\ e_{k+1}^{(n-1)} \end{bmatrix} \leq \begin{bmatrix} (1+a_{00}h) & a_{01}h & a_{02}h^2 & \cdots & a_{0\,n-1}h^{n-1} \\ a_{10}h^{n-1} & (1+a_{11}h) & a_{12}h & \cdots & a_{1\,n-1}h^{n-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n-1\,0}h & a_{n-1\,1}h^2 & \cdots & \cdots & (1+a_{n-1n-1}h) \end{bmatrix} \begin{bmatrix} e_k \\ e'_k \\ \vdots \\ e_k^{(n-1)} \end{bmatrix} +$$

$$+ b^* \omega_r(h) h^{n+r+1} \begin{bmatrix} h^{n-1} \\ h^{n-2} \\ \vdots \\ h \\ 1 \end{bmatrix}.$$

Otherwise,

$$E_{k+1} \leq (I_n + hA) E_k + \omega_r(h)\, h^{n+r+1} B \tag{15}$$

where I_n is the n-th order unite matrix, $A = [a_{ij}]$, $i, j = 0, 1, \ldots, n-1$ and B is the $n \times 1$ matrix $B = (b^* b^* \ldots b^*)^T$. Obviously $b^* = \max\limits_{0 \leq i \leq n-1} b_i$ Applying the principle of successive substitution, (15) reduces to

$$E_{k+1} \leq (I_n + hA)^{k+1} E_0 + \omega_r(h) h^{n+r+1} \cdot \sum_{j=0}^{k} (I_n + hA)^j \cdot B \tag{16}$$

and with $E_0 = 0$ we easily get

$$E_{k+1} \leq \omega_r(h) h^{n+r} C$$

where C is an $n \times 1$ matrix whose elements are constants independent of h. Otherwise,

$$C = b e^{bA} \cdot B$$

uad this completes the proof.

THEOREM 2. *Let* $y^{(n+q)}(x_{k+1})$, *given by* (6), *be the higher derivatives of the exact solution of* (1)–(2) *for* $q = 0, 1, \ldots, r$. *If the corresponding approximate values* $\bar{y}_{k+1}^{(n+q)}$ *are given by* (7), *then the error is bounded by the inequality*

$$e_{k+1}^{(n+q)} \leq c_{n+q} w_r(h) h^{n+r}$$

which holds for all $k = 0, 1, \ldots, N-1$.

PROOF. Using the results of theorem 1 and the inequality (12), it will be easy to get the proof.

As a conclusion of Theorems 1 and 2, we have proved that the inequality

$$e_{k+1}^{(i)} \leq c_i w_r(h) h^{n+r} \tag{17}$$

holds for all $i = 0, 1, \ldots, n+r$ and all $k = 0, 1, \ldots, N-1$.

3. Spline function approximating the solution

In this paragraph we construct the spline function approximating the solution of (1)–(2) and we prove that this spline function is unique. Thus we introduce the following theorem:

THEOREM 3. *For the given mesh of points*

$$\Delta : 0 = x_0 < x_1 < \ldots < x_k < x_{k+1} < \ldots < x_N = b$$

where

$$x_{k+1} - x_k = h, \qquad k = 0, 1, \ldots, N-1$$

and for given sets of approximate values

$$\bar{y}^{(q)} : \bar{y}_0^{(q)}, \bar{y}_1^{(q)}, \ldots, \bar{y}_N^{(q)}, \qquad q = 0, 1, \ldots, n+r$$

there is a unique spline function $S_\Delta(x) \equiv S_\Delta(\bar{y}; x)$ *interpolated on* Δ *to the set* $\bar{y}$ *and satisfying the following conditions:*

$$S(\bar{y}; x) = S_\Delta(x) \in C^{n+r}[0, b], \tag{18}$$

$$S_k^{(q)}(x_k) = \bar{y}_k^{(q)}, \qquad S_{N-1}^{(q)}(x_N) = \bar{y}_N^{(q)} \tag{19}$$

where $q = 0, 1, \ldots, n+r$ *and* $k = 0, 1, \ldots, (N-1)$, *Also for*

$$x_k \leq x \leq x_{k+1}$$

$$S_\Delta(x) = S_k(x) = \sum_{j=1}^{n+r} \frac{\bar{y}_k^{(j)}}{j!} (x - x_k)^j + \sum_{p=1}^{n+r+1} a_p^{(k)} (x - x_k)^{p+n+r}. \tag{20}$$

PROOF. From the continuity condition (18) using (19) and (20) it is easy to get

$$\sum_{p=1}^{n+r+1} t! \binom{p+n+r}{t} a_p^{(k)} h^{p-1} = F_t^{(k)} \tag{21}$$

where

$$F_t^{(k)} = h^{t-n-r-1}\left(\bar{y}_{k+1}^{(t)} - \sum_{j=0}^{n+r-t} \frac{\bar{y}_k^{(j+t)}}{j!} h^j\right), \tag{22}$$

$t = 0, 1, \ldots, n+r$ and $k = 0, 1, \ldots, (N-1)$. Here $a_p^{(k)}$ $(p = 1, 2, \ldots, n+r+1)$ are the unknowns to be determined. The system of linear equations (21) in the unknowns $a_p^{(k)}$ has a unique solution since its determinant $D_r \neq 0$.

Here,

$$D_r = |d_{ij}|, \qquad i\&j = 1, 2, \ldots, n+r+1,$$

$$d_{ij} = \binom{n+r+j}{i-1}(i-1)!\, h^{j-1}$$

and it is easy to prove that

$$D_r = h^{1/2(n+r)(n+r+1)} \prod_{t=0}^{n+r} t!$$

and this does not equal to zero since $h \neq 0$.

If we replace the p-th column in D_r by the column

$$(F_0^{(k)}, F_1^{(k)}, \ldots, F_{n+r}^{(k)})^T$$

and denote the resulting determinant by D_r^p, then the solution of the system (21) will be

$$a_p^{(k)} = \frac{D_r^p}{D_r}, \quad p = 1, 2, \ldots, n+r+1 \tag{23}$$

and after factorizing D_r^p in terms of $F_0^{(k)}, F_1^{(k)}, \ldots, F_{n+r}^{(k)}$, the solution (23) will take the form

$$a_p^{(k)} = \frac{1}{h^{p-1}} \sum_{t=0}^{n+r} c_{pt} F_t^{(k)} \tag{24}$$

and this completes the proof.

4. Convergence of the spline function to the solution

Before we prove the convergence of the constructed spline function to the solution of the differential equation (1)–(2), we first prove the following lemma:

LEMMA 2. The following inequalities are true

$$|a_p^{(k)}| \leq \frac{A_p}{h^p} w_r(h)$$

where A_p are constants independent of h and $p = 1, 2, \ldots, n+r+1$.

PROOF. Using (24) we get

$$|a_p^{(k)}| \leq \frac{1}{h^{p-1}} \sum_{t=0}^{n+r} C_{pt} |F_t^{(k)}|. \tag{25}$$

Now, using the Taylor expansion of $y^{(t)}(x)$, i.e.:

$$y^{(t)}(x) = \sum_{j=0}^{n+r-t-1} \frac{y_k^{(j+t)}}{j!} (x-x_k)^j + \frac{y^{(n+r)}(\xi_{kt})}{(n+r-t)!} (x-x_k)^{n+r-t} \tag{26}$$

where $x_k < \xi_{kt} < x_{k-1}$ for all $t = 0, 1, \ldots, n+r$ and all $k = 0, 1, \ldots, N-1$. For all $x = x_{k+1}$ we get

$$y_{k+1}^{(t)} = \sum_{j=0}^{n+r-t-1} \frac{y_k^{(j+t)}}{j!} h^j + \frac{y^{(n+r)}(\xi_{kt})}{(n+r-t)!} h^{n+r-t}$$

and if we used this result with (22) we can get

$$|F_t^{(k)}| \leq h^{t-n-r-1} \left(e_{k+1}^{(t)} + \sum_{j=0}^{n+r-t} \frac{e_k^{(j+t)}}{j!} h^j + \right.$$

$$\left. + \frac{1}{(n+r-t)!} |y^{(n+r)}(\xi_{kt}) - y_k^{(n+r)}| h^{n+rpt} \right). \tag{27}$$

Using the Theorems 1 and 2 and the definition of the modulus of continuity of $y^{(n+r)}(x)$, then (27) becomes

$$|F_t^{(k)}| \leq c_t^* \frac{w_r(h)}{h}, \quad t = 0, 1, \ldots, n+r \tag{28}$$

where c_t^* are constants independent of h. Using the help of this result (28) in (25), the proof of this lemma will be complete.

THEOREM 4. *Let $y(x)$ be the solution of* (1)–(2) *and let $f \in C^r([0, b] \times \mathbf{R}$, where $r \in I^+$. If $S_\Delta(x)$ is the spline function constructed in Theorem* 3, *then there exists a constant K, independent of h, such that*

$$|y^{(q)}(x) - S_\Delta^{(q)}(x)| \leq K w_r(h) h^{n+r-q}$$

for all $x \in [0, b]$ and all $q = 0, 1, \ldots, n+r$.

PROOF. Using (26) and (20), it is easy to get

$$|y^{(q)}(x) - S_\Delta^{(q)}(x)| \leq \sum_{j=0}^{n+r-q} \frac{e_k^{(j+q)}}{j!} h^j + \frac{|y^{(n+r)}(\xi_{kt}) - y_k^{(n+r)}|}{(n+r-q)!} h^{n+r-q} +$$

$$+ \sum_{p=1}^{n+r+1} q! \binom{n+r+p}{q} |a_p^{(k)}| h^{p+n+r-q}.$$

Taking the help of Theorem 1, Theorem 2, the definition of the modulus of continuity of $y^{(n+r)}(x)$ and the lemma 2, the above inequality becomes

$$|y^{(q)}(x) - S_\Delta^{(q)}(x)| \leq c_q^{**} w_r(h) h^{n+r-q}$$

where c_q^{**} $(q = 0, 1, \ldots, n+r)$ are constants independent of h. Taking $K = \max c_q^{**}$ $(q = 0, 1, \ldots, n+r)$, we get

$$|y^{(q)}(x) - S_\Delta^{(q)}(x)| \leq K w_r(h) h^{n+r-q}$$

and the proof is complete.

THEOREM 5. *If $\bar{S}_\Delta^{(n)}(x)$ denotes the function*

$$\bar{S}_\Delta^{(n)}(x) = f(x, S_\Delta(x))$$

where $S_\Delta(x)$ is the spline function approximating the solution of (1)–(2) *and constructed in Theorem* 3, *then for any $x \in [0, b]$ we have*

$$|\bar{S}_\Delta^{(n)}(x) - S_\Delta^{(n)}(x)| \leq M^* w_r(h) h^r$$

where M^* is a constant independent of h.

Otherwise,

$$\lim_{n \to \infty} S_\Delta^{(n)}(x) = f(x, S_\Delta(x)).$$

PROOF. We have

$$|\bar{S}_\Delta^{(n)}(x) - S_\Delta^{(n)}(x)| \leq |\bar{S}_\Delta^{(n)}(x) - y^{(n)}(x)| + |y^{(n)}(x) - S_\Delta^{(n)}(x)| =$$

$$= |f(x, S_\Delta(x)) - f(x, y(x))| + |y^{(n)}(x) - S_\Delta^{(n)}(x)|.$$

Taking the help of the Lipschitz condition and Theorem 4, this becomes

$$|\bar{S}_\Delta^{(n)}(x) - S_\Delta^{(n)}(x)| \leq L K w_r(h) h^{n+r} + K w_r(h) h^r \leq M^* w_r(h) h^r.$$

Hence the proposition.

REMARK. We have proved that the method is stable in [8].

References

[1] GH. MICULA: Numerical Integration of Diferential Equation $y^{(n)} = f(x, y)$ by Spline Functions, *Rev. Roum. Math. Pures et Appl.*, **17** (1972), 1385–1389.

[2] THARWAT FAWZY: Spline Functions and Cauchy Problems, I. Approximate solution of $y'' = f(x, y, y')$ with spline functions, *Annales Univ. Sci., Budapest*, Sectio Comp., **1** (1978), 81–98.

[3] THARWAT FAWZY: Spline Functions and Cauchy Problems, II. Approximate solution of $y'' = f(x, y, y')$ with spline functions, *Acta Math. Acad. Sci. Hungar.*, **29** (1977), 259–271.

[4] THARWAT FAWZY: Spline Functions and Cauchy Problems, III. Approximate solution of $y' = f(x, y)$ with spline functions, *Annales Univ. Sci. Budapest*, Sectio Comp., 1 (1978), 35–45.

[5] THARWAT FAWZY: Spline Functions and Cauchy Problems, IV. On the stability of the method, *Acta Math. Acad. Sci. Hungar.*, **30** (1977), 219–226.

[6] THARWAT FAWZY, KŐHEGYI JÁNOS and FEKETE ISTVÁN: Spline Functions and Cauchy Problems, V. Applications programs to the method, *Annales Univ. Sci. Budapest*, Sectio Comp. **1** (1978), 109–125.

[7] THARWAT FAWZY: Spline Functions and Cauchy Problems, Ph. D. Thesis. The Hungarian Academy of Science. Institute of Math. Researches. Budapest. 1976. D/6906+T.

[8] THARWAT FAWZY: Spline functions and Cauchy Problems, VII, *Annales Univ. Sci. Budapest*, Sectio Mathematica, **24** (1981), 57–62.

EINE BEMERKUNG ÜBER GEWISSE NULLMENGEN VON KETTENBRÜCHEN

von

G. RAMHARTER

Institut für Analysis, Technische Universität, Wien

(Eingegangen am 30.9.1981)

Ist a eine beliebige natürliche Zahl, dann ist bekanntlich für fast alle reellen Zahlen $x \in (0, 1)$ die mittlere Häufigkeit, mit der a als Teilnenner in der regulären Kettenbruchentwicklung $x = \frac{1}{a_1+}\frac{1}{a_2+}\dots =: [a_1, a_2, \dots]$ auftritt, vorhanden und gleich ${}_2\log(1+1/(a^2+2a))$. Im Gegensatz dazu ist bei der semiregulären Entwicklung $\frac{1}{a_1-}\frac{1}{a_2-}\dots$ mit Teilnennern $a_n \geq 2$ die Häufigkeit des Auftretens jeder Zahl $a \geq 3$ fast überall gleich Null. Seit dem Beweis des zuerst erwähnten schon von GAUSS vermuteten Satzes durch LÉVY und (unabhängig) KUZMIN sind zahlreiche weitere Mittelwerteigenschaften der Teilnenner gefunden worden. Man hat aber auch die dabei auftretenden Ausnahme-Nullmengen genauer untersucht, wobei sich insbesondere die Hausdorffdimension als feineres Unterscheidungsmerkmal eignet (vgl. etwa [1]–[4], [7], [8]). Speziell I. J. GOOD hat Zusammenhänge zwischen Wachstumseigenschaften der Teilnenner und den Dimensionszahlen systematisch studiert. P. ERDŐS hat angeregt, die Menge E der Kettenbrüche mit paarweise verschiedenen Teilnennern zu untersuchen. Wir werden sehen, daß deren Dimension entgegen den Erwartungen groß ist, sogar dann noch, wenn man eine beliebige (endliche) Anzahl von Werten für die Teilnenner ausschließt oder darüber hinausgehend nur wachsende Teilnennerfolgen zuläßt. Es hat sich gezeigt, daß die semireguläre Version des Problems eine Anwendung auf das in [5], [6] untersuchte asymmetrische Lagrangespektrum erlaubt. Die im folgenden betrachteten Mengen $G_q \cap E$ enthalten nämlich gerade die am schlechtesten einseitig approximierbaren reellen Zahlen.

Es bezeichne für festes $q \in \mathbb{N}$ F_q die Menge der Zahlen $x = [a_1, a_2, \dots]$, deren Teilnenner alle $\geq q$ sind, und $G_q (\subset F_q)$ die Menge der Zahlen x, für die überdies $(q \leq)$ $a_1 \leq a_2 \leq \dots$ gilt.

SATZ. *Es gilt*

$$\dim G_q = \dim (G_q \cap E) = \frac{1}{2} \quad (q \in \mathbf{N})$$

sowie

$$\dim (F_q \cap E) = \frac{1}{2} + O\left(\frac{\log \log q}{\log q}\right) \quad (q \to \infty).$$

$\dim F_q$ hat nach ([2], Th. 2) das gleiche asymptotische Verhalten wie $\dim (F_q \cap E)$. Die Zusatzbedingung $a_n \neq a_s (n \neq s)$ hat also auf die Dimension in beiden Fällen überraschenderweise keinen Einfluß.

Gleichlautende Aussagen gelten in der semiregulären Version, wenn man $q \geq 3$ voraussetzt.

BEWEIS. Für jedes höchstens abzählbare System $\mathfrak{J}$ von Intervallen I_i mit Längen $|I_i|$ setze man $L_s(\mathfrak{J}) = \Sigma |I_i|^s$. Für eine Menge $H \subset [0, 1]$ und ein $\varepsilon > 0$ bezeichne $\Lambda_{s,\varepsilon}(H) = \inf L_s(\mathfrak{J})$, wobei das Infimum über alle Systeme von Intervallen zu erstrecken ist, die H überdecken und deren Längen sämtlich durch ε beschränkt sind. Dann existiert $h_s(H) = \lim_{\varepsilon \to 0} \Lambda_{s,\varepsilon}(H)$, das sogenannte s-dimensionale Hausdorffsche Maß von H bezüglich der Maßfunktion t^s. Ferner gibt es (wie man zeigt) eine eindeutig bestimmte Zahl $d \in [0, 1]$ (die sogenannte Hausdorffsche Dimension $\dim H$ von H) derart, daß $h_s(H) = \infty$ für $s < d$ und $h_s(H) = 0$ für $s > d$ (Im Fall $d = 0$, bzw. $d = 1$, entfällt die erste, bzw. zweite, dieser Aussagen). Wir nennen ein abgeschlossenes Intervall $I^{(n)} = I(a_1, \ldots, a_n)$ mit Endpunkten $[a_1, \ldots, a_n]$, $[a_1, \ldots, a_n + 1]$ $(n, a_1, \ldots, a_n \in \mathbf{N})$ fundamental (bezüglich H) von n-ter Ordung genau dann, wenn es ein Element $x \in H$ mit $x = [a_1, \ldots, a_n, \ldots]$ gibt. Offenbar wird H für jedes feste $n \in \mathbf{N}$ vom System $\mathfrak{J}^{(n)}$ aller bezüglich H fundamentalen Intervalle der Ordnung n überdeckt; wir bezeichnen $\mathfrak{J}^{(n)}$ als das fundamentale Überdeckungssystem n-ter Ordnung von H. Aus der Definition der Dimension folgt unmittelbar, daß $\dim H \leq s$ gilt, wenn $\liminf_{n \to \infty} L_s(\mathfrak{J}^{(n)}) < \infty$ zutrifft. Jedoch impliziert $\liminf_{n \to \infty} L_s(\mathfrak{J}^{(n)}) > 0$ im allgemeinen nicht $\dim H \geq s$. Dadurch ist der Zugang zu den unteren Abschätzungen, mit denen wir hier beginnen wollen, erschwert. Es ist klar, daß es wegen der Monotonie der Dimension bezüglich Inklusion genügt zu zeigen

$$\frac{1}{2} \leq \dim(G_q \cap E). \tag{1}$$

Die Idee ist nun die, geeignete Teilmengen von $G_q \cap E$ mit nicht zu schnellem (und zwar höchstens polynomialem) Wachstum der Teilnenner zu betrachten, für welche ein Schluß wie der oben erwähnte gerade noch zulässig ist, und damit die Dimension von unten zu approximieren. Es sei dazu $G_{q,p} (\subset G_q \cap E)$, $q, p \in \mathbf{N}$, die Menge der Zahlen $x = [a_1, a_2, \ldots]$, für die

$$q \leq a_1 < a_2 < \ldots \text{ sowie } a_n \leq (n+2)^p - 1 =: g(n) \quad (n \in \mathbf{N})$$

erfüllt ist. Offenbar ist $G_{q,\,p}$ abgeschlossen. Aus dem Satz von Heine-Borel folgt ohne Schwierigkeit, daß man sich bei der Bestimmung von $\Lambda_{s,\,\varepsilon}(G_{q,\,p})$ auf diejenigen Überdeckungssysteme $\mathfrak{U}$ beschränken kann, die aus endlich vielen abgeschlossenen Intervallen mit Endpunkten in $G_{q,\,p}$ bestehen. Nun zeigt man:

(2) *Es sei $\varepsilon>0$, $1>s>s-t>0$, q, $p\in\mathbf{N}$, und ein Überdeckungssystem $\mathfrak{U}$ von $G_{q,\,p}$ der beschriebenen Art mit Feinheit $<\varepsilon$ beliebig gegeben. Dann existiert ein endliches, nur aus fundamentalen Intervallen (eventuell variabler Ordnung) bestehendes Überdeckungssystem $\mathfrak{W}$ mit Feinheit $<l$ derart, daß $L_{s-t}(\mathfrak{U})>L_s(\mathfrak{W})$ gilt.*

Die Argumentation ist im wesentlichen dieselbe wie bei Good ([2], S. 207–209), bis auf die Beobachtung, daß die für die dortigen Zwecke hinreichende Forderung $g(n) = O(\log^p n)$ durch die schwächere $g(n) = O(n^p)$ ersetzt werden kann (Man beachte dort insbesondere die Beziehung (4.7)). Wir behaupten weiters:

(3) *Es sei $q\in\mathbf{N}$ beliebig gegeben und $p\in\mathbf{N}$, $2^p-1>q$, sowie $s=\frac{1}{2}-\frac{1}{p}$ gewählt. Dann gilt für jedes endliche nur aus fundamentalen Intervallen bestehende Überdeckungssystem $\mathfrak{W}$ von $G_{q,\,p}$ von beliebiger Feinheit $L_s(\mathfrak{W})>1$.*

Es ist klar, daß aus (2) und (3) zusammengenommen $\dim G_{q,\,p}\to\frac{1}{2}\ (p\to\infty)$ für jedes feste q folgt und daraus (1).

Ist $q_n = q(a_1, \ldots, a_n)$ der Nenner des reduzierten Bruches $p_n/q_n = [a_1, \ldots, a_n]$, dann ergibt sich für die Länge eines Intervalls $I(a_1, \ldots, a_n)$ mit Hilfe bekannter Regeln

$$|I^{(n)}| = |[a_1, \ldots, a_n]-[a_1, \ldots, a_{n-1}, a_n+1]| = 1/(q_n(q_n+q_{n-1})),$$

zusammen mit der trivialen Abschätzungen $q_{n-1}<q_n$ und $a_1\ldots a_n<q_n<(a_1+1)\ldots(a_n+1)$ also

$$(a_1\ldots a_n)^{-2s}>|I(a_1, \ldots, a_n)|^s>2^{-s}((a_1+1)\ldots(a_n+1))^{-2s}. \tag{4a, b}$$

Daß zur Überdeckung von $G_{q,\,p}$ überhaupt endliche aus fundamentalen Intervallen bestehende Überdeckungssysteme ausreichen, ist aus der Definition dieser Mengen unmittelbar klar. Zum Nachweis von (3) kann man ohne Beschränkung der Allgemeinheit annehmen, daß keine zwei Intervalle eines solchen Überdeckungssystems $\mathfrak{W}$ innere Punkte gemeinsam haben. Kommt ein Intervall $I(a_1, \ldots, a_n)$ vor, so entferne man alle eventuell auftretenden Teilintervalle $I(a_1, \ldots, a_n, k_1, k_2, \ldots)$. Dabei wird $L_s(\mathfrak{W})$ sicher nicht vergrößert. Enthält nun ein Überdeckungssystem der betrachteten Art ein Intervall $I(a_1, \ldots, a_{m-1}, a_m, \ldots, a_r)$, so enthält es zu jedem $k = a_{m-1}+1, \ldots, g(m)$ ein Intervall der Form $I(a_1, \ldots, a_{m-1}, k, b_1, \ldots, b_{\nu(k)})$ mit im allgemeinen von k abhängiger Ordnung $\nu(k)+m$. Berücksichtigt man alle

diese Tatsachen und die Beziehung (4b), so erhält man, wenn n die größte vorkommende Ordnung von Intervallen in $\mathfrak{W}$ bezeichnet,

$$L_s(\mathfrak{W}) = \sum_{I(a_1, \ldots, a_m) \in \mathfrak{W},\, m \in \mathbf{N}} |I^{(m)}|^s >$$

$$> 2^{-s} \sum_{a_1=q}^{g(1)} (a_1+1)^{-2s} \ldots \sum_{a_{n-1}=a_{n-2}+1}^{g(n-1)} (a_{n-1}+1)^{-2s} \sum_{a_n=a_{n-1}+1}^{g(n)} (a_n+1)^{-2s}.$$

Nun ist

$$\sum_{a_n=a_{n-1}+1}^{g(n)} (a_n+1)^{-2s} > \int_{x=a_{n-1}+1}^{g(n)} (x+1)^{-2s} dx = \frac{p}{2}\left((n+2)^2 - (a_{n-1}+2)^{2/p}\right) \geq$$

$$\geq \frac{p}{2}\left((n+2)^2 - ((n+1)^p+1)^{2/p}\right),$$

und wegen $p \geq 2$ ist dies ≥ 2 für $n \in \mathbf{N}$, wie man sich leicht überzeugt. So weiterschließend bestätigt man die Richtigkeit von (3).

Wir wenden uns den oberen Abschätzungen zu. Für $m = 0, 1, 2, \ldots$ bezeichne H_m die Menge der Zahlen mit einer Entwicklung der Form $[(1)_m, a_1, a_2, \ldots]$, wobei eine Sequenz von m Einsen am Anfang stehen und für die übrigen Teilnenner $2 \leq a_1, \leq a_2 \leq \ldots$ gelten soll. Offenbar ist

$$G_1 = \bigcup_{m=0}^{\infty} H_m \cup \{[(1)_\infty]\}.$$

Nach [2, Lemma 1] ist dim H_m = dim H_0 und daher (dim $G_q \leq$) dim G_1 = = dim H_0. Nach einer Bemerkung von oben genügt es also, folgende Aussagen über die fundamentalen Überdeckungssysteme $\mathfrak{J}^{(n)}$ von H_0 bzw. $F_q \cap E$) zu zeigen:

(5) *Es sei* $2 > \sigma = 2s = 1+t > 1$ *beliebig gegeben. Dann ist* $L_s(\mathfrak{J}^{(n)}(H_0)) < 1$ *für alle hinreichend großen* $n \in \mathbf{N}$.

(6) *Ist überdies* $q > (1/t)^{1/t}$, *dann gilt für alle* $n \in \mathbf{N}$ $L_s(\mathfrak{J}^{(n)}(F_{q+1} \cap E)) < 1$.

(Der Nachweis von (6) wäre natürlich entbehrlich, da die obere Abschätzung von dim $(F_q \cap E)$ wegen $F_q \cap E \subset F_q$ auch aus dem erwähnten Resultat von Good über dim F_q folgt. Wir geben hier einen davon unabhängigen einfachen Beweis). Wegen (4a) ist

$$L_s(\mathfrak{J}^{(n)}(H_0)) = \sum_{2 \leq a_1 \leq a_2 \leq \ldots \leq a_n} |I^{(n)}|^s < \sum_{a_1=2}^{\infty} a_1^{-\sigma} \sum_{a_2=a_1}^{\infty} a_2^{-\sigma} \ldots \sum_{a_n=a_{n-1}}^{\infty} a_n^{-\sigma}.$$

Nun ist

$$\sum_{a_n=a_{n-1}}^{\infty} a_n^{-\sigma} < a_{n-1}^{-\sigma} + \int_{a_{n-1}}^{\infty} x^{-\sigma} dx = \left(a_{n-1}^{-1} + \frac{1}{t}\right) a_{n-1}^{-t} \leq \left(\frac{1}{2} + \frac{1}{t}\right) a_{n-1}^{-t},$$

daher

$$\sum_{a_{n-1}=a_{n-2}}^{\infty} a_{n-1}^{-\sigma} \sum_{a_n=a_{n-1}}^{\infty} a_n^{-\sigma} < \left(\frac{1}{2}+\frac{1}{t}\right)\left(a_{n-2}^{-1}+\frac{1}{2t}\right)a_{n-2}^{-2t} \leq$$

$$\leq \left(\frac{1}{2}+\frac{1}{t}\right)\left(\frac{1}{2}+\frac{1}{2t}\right)a_{n-2}^{-2t}$$

und so fort bis

$$L_s(\mathfrak{F}^{(n)}(H_0)) < \left(2^{-n}\prod_{\nu=1}^{n}\left(1+\frac{1}{\nu}\frac{2}{t}\right)\right)2^{-nt}.$$

Der letzte Ausdruck strebt offenbar für jedes noch so kleine feste $t>0$ gegen 0 bei $n\to\infty$. Daraus folgt die Richtigkeit von (5). Wieder wegen (4a) ist

$$L_s(\mathfrak{F}^{(n)}(F_{q+1}\cap E)) < n! \sum_{q+1\leq a_1<a_2<\ldots<a_n} (a_1\cdot a_2\ldots a_n)^{-\sigma}.$$

Ausgehend von

$$\sum_{a_n=a_{n-1}+1}^{\infty} a_n^{-\sigma} < \int_{a_{n-1}}^{\infty} x^{-\sigma}dx = \frac{1}{t}a_{n-1}^{-t}$$

erhält man induktiv

$$n!\sum_{a_1=q+1}^{\infty} a_1^{-\sigma}\ldots\sum_{a_n=a_{n-1}+1}^{\infty} a_n^{-\sigma} < n!\,\frac{1}{n!t^n}\,q^{-nt},$$

und dies ist kleiner als 1 wegen der Annahme über q. Damit ist auch (6) gezeigt und der Beweis des Satzes abgeschlossen.

Literatur

[1] CUSICK, T. W.: Continuants with bounded digits, *Mathematika*, **24** (1977), 166–172; II., *Mathematika*, **25** (1978), 107–109. III., *Monatsh. Math.*, **99** (1985), 105–109.

[2] GOOD, I. J.: The fractional dimensional theory of continued fractions, *Proc. Cambridge Phil. Soc,*. **37** (1941), 199–228.

[3] JARNIK, V.: Zur metrischen Theorie der diophantischen Approximationen, *Prace* Mat. Fiz., **36** (1928), 91–106.

[4] KAUFMAN, R.: Continued fractions and Fourier transforms, *Mathematika*, **27** (1980), 262–267.

[5] LINDGREN, A.: One-sided minima of indefinite binary quadratic forms and one-sided diophantine approximation, *Ark. Math.*, **13** (1975), 287–302.

[6] RAMHARTER, G.: Über asymmetrische diophantische Approximationen, *Journ. Number Theory*, **14** (1982), 269–279.

[7] RAMHARTER, G.: Some metrical properties of continued fractions, *Mathematika*, **30** (1983), 117–132.

[8] ROGERS, C. A.: Some sets of continued fractions, *Proc. London Math. Soc.*, (3) **14** (1964), 29–44.

NOTES ON LACUNARY INTERPOLATION BY SPLINES, I.

(0,3) INTERPOLATION

By

THARWAT FAWZY

Suez Canal University, Ismailia, Egypt

(Received August 11, 1982)

1. Introduction. P. TURÁN and J. BALÁZS [1] in 1957 have initiated the study of "Lacunary Interpolation". Recently, A. MEIR and A. SHARMA [2], B. K. SWARTZ and R. S. VARGA [3], S. DEMKO [4], A. K. VARMA [5] and J. PRASAD and A. K. VARMA [6] considered special lacunary interpolation problems.

In this series of papers titled Notes on Lacunary Interpolation by Splines, we present new methods for the cases (0,3), (0,2), (0,4), (0, 1, 3), (0, 1, 4) and (0,2,4) interpolation and we get results of the same order as that of best approximation for the functions and their all possible derivatives. Moreover, the stability of interpolation in each case is proved.

In this paper we begin with the case (0,3) interpolation and it is convenient to state the results of J. PRASAD and A. K. VARMA [6] in the following two theorems.

THEOREM A. *Given arbitrary numbers* $f(x_i)$, $i = 0, 1, \ldots, n$; $f^{(p)}(z_i)$, $i = 0, 1, \ldots, n-1$, $p = 0, 3$; $2z_i = x_i + x_{i+1}$; $f'(x_0), f'(x_n)$; *there exists a unique* $S_n \in S_{n,5}^{(2)}$ *such that*

$$S_n(x_i) = f(x_i), \qquad i = 0, 1, \ldots, n,$$

$$S_n^{(p)}(z_i) = f^{(p)}(z_i), \qquad i = 0, 1, \ldots, n-1; \qquad p = 0, 3, \tag{1.1}$$

$$S_n'(x_n) = f'(x_n), \qquad S_n'(x_0) = f'(x_0).$$

THEOREM B. *Let* $f \in C^r[0, 1]$. *Then for the unique quintic spline* $S_n(x)$ *associated with* f *and satisfying* (1.1), *we have*

$$|S_n^{(j)}(x) - f^{(j)}(x)| \leqq \beta_{j,r}\delta^{r-j}\omega_r(\delta), \qquad j = 0, 1, 2 \text{ and } r = 3, 4, 5, \tag{1.2}$$

$$|S_n^{(j)}(x) - f^{(j)}(x)| \leqq \beta_{j,r}\delta^{6-j} \max_{0 \leqq x \leqq 1} |f^{(6)}(x)|, \quad j = 0, 1, 2 \text{ and } r = 6, \tag{1.3}$$

where $\omega_r(\cdot)$ *denotes the modulus of continuity of* $f^{(r)}$, $\delta = \max(x_{i+1}-x_i)$, $i = 0, 1, \ldots, n-1$ *and* $\beta_{j,r}$ *are constants independent of* f *and* δ.

Note that these constants $\beta_{j,r}$ have not been calculated while in our results their values are completely given. Also, the error estimation is not given in the above theorem for higher derivatives than the second derivative and in our method it is given for all possible derivatives. Moreover, the conditions $S_n'(x_0)$ and $S_n'(x_n)$ are released.

In this note, we study the (0,3) interpolation, $f \in C^r[0, 1]$, separately for $r = 3, 4$ and 5 in the three following cases:

2. Case. A. In this case $f \in C^3[0, 1]$ and we consider the partition

$$\Delta : 0 = x_0 < x_1 < \ldots < x_k < x_{k+1} < \ldots < x_n = 1,$$

where for $k = 0, 1, \ldots, n-1$,

$$h_k = x_{k+1} - x_k \quad \text{and} \quad h = \max h_k.$$

THEOREM 2.1. *Given arbitrary numbers* $f(x_k)$, $k = 0, 1, \ldots, n$, $f^{(p)}(z_k)$, $k = 0, 1, \ldots, n-1$, $p = 0, 3$; $2z_k = x_k + x_{k+1}$; *then there exists a unique spline* $S_n \in S_{n,3}^{(0)}$ such that

$$S_n(x_k) = f(x_k), \tag{2.1}$$

$$S_n(z_k) = f(z_k), \qquad k = 0, 1, \ldots, n-1, \tag{2.2}$$

$$S_n^{(3)}(z_k) = f^{(3)}(z_k), \qquad k = 0, 1, \ldots, n-1. \tag{2.3}$$

PROOF. For $x \in [x_k, x_{k+1}]$ and $k = 0, 1, \ldots, n-1$, set

$$S_n(x) = f(x_k) + a_k(x-x_k) + (1/2)b_k(x-x_k)^2 + (1/3!)c_k(x-x_k)^3. \tag{2.4}$$

Then the values

$$a_k = (1/h_k)[4f(z_k) - 3f(x_k) - f(x_{k+1}) + (h_k^3/12)f^{(3)}(z_k)], \tag{2.5}$$

$$b_k = [4f(x_{k+1}) + 4f(x_k) - 8f(z_k) - (h_k^3/2)f^{(3)}(z_k)]/h_k^2 \tag{2.6}$$

and

$$c_k = f^{(3)}(z_k) \tag{2.7}$$

prove the theorem.

THEOREM 2.2. *Let* $f \in C^3$ [0,1]. *Then for the unique cubic spline* S_n *associated with* f *and given in* Theorem 2.1. *we have for all* $x \in [0, 1]$

$$|S_n^{(i)}(x) - f^{(i)}(x)| \leq c_{3,i} h^{3-i} \omega_3(h) \quad i = 0, 1, 2, 3,$$

where $\omega_3(\cdot)$ *denotes the modulus of continuity of* $f^{(3)}$, $h = \max h_k$, $k = 0, 1, \ldots, n-1$ *and the values of* $c_{3,i}$ *are*

$$c_{3,0} = 2/3, \qquad c_{3,1} = 4/3, \qquad c_{3,2} = 5/3, \qquad c_{3,3} = 1.$$

PROOF. We have for $x \in [x_k, x_{k+1}]$, $x_k < \xi_k < x_{k+1}$ and $k = 0, 1, \ldots, n-1$,

$$|S_n(x) - f(x)| \leqq h|a_k - f'(x_k)| + (h^2/2)|b_k - f''(x_k)| + (h^3/3!)|c_k - f^{(3)}(\xi_k)|. \tag{2.8}$$

Using (2.5), (2.6) and (2.7), the following estimations could be easily obtained for all $k = 0, 1, \ldots, n-1$

$$|a_k - f'(x_k)| \leqq (1/6)h^2\omega_3(h), \tag{2.9}$$

$$|b_k - f''(x_k)| \leqq (2/3)\, h\omega_3(h) \tag{2.10}$$

and

$$|c_k - f^{(3)}(\xi_k)| \leqq \omega_3(h). \tag{2.10}$$

The above three estimations with (2.8) give

$$|S_n(x) - f(x)| \leqq (2/3)h^3\omega_3(h).$$

Similarly, it is easy to complete the proof for the derivatives.

3. Case B. In this case $f \in C^4[0, 1]$ and we consider the partition

$$\Delta : 0 = x_0 < x_1 < \ldots < x_k < x_{k+1} < \ldots < x_n = 1$$

where

$$h = x_{k+1} - x_k \qquad \text{and} \qquad k = 0, 1, \ldots, n-1.$$

THEOREM 3.1. *Given arbitrary numbers $f(x_k)$, $k = 0, 1, \ldots, n$, $f^{(p)}(z_k)$, $k = 0, 1, \ldots, n-1$; $p = 0, 3$; $2z_k = x_k + x_{k+1}$; then there exists a unique spline $S_\Delta \in C[0,1]$ such that*

$$S_\Delta(x) \in \pi_4 \text{ on each } [x_k, x_{k+1}],\ k = 0, 1, \ldots, n-1, \tag{3.1}$$

$$S_\Delta(x_k) = f(x_k),\ k = 1, \ldots, n: \tag{3.2}$$

$$S''_\Delta(x_k) = [f(x_{k+1}) - 2f(x_k) + f(x_{k-1}) - (h^2/12)(f^{(3)}(z_k) - f^{(3)}(z_{k-1}))]/h^2. \tag{3.3}$$

$$k = 1, 2, \ldots, n-1,$$

$$S_\Delta^{(p)}(z_k) = f^{(p)}(z_k), \qquad k = 1, 2, \ldots, n-1; \qquad p = 0, 3. \tag{3.4}$$

PROOF. Set for $x \in [x_k, x_{k+1}]$, $k = 0, 1, \ldots, n-1$

$$S_\Delta(x) = \begin{cases} S_0(x), & x_0 \leqq x \leqq x_1, \\ S_k(x), & x_k \leqq x \leqq x_{k+1}, \quad k = 1, 2, \ldots, n-1 \end{cases} \tag{3.5}$$

where,

$$S_k(x) = f(x_k) + a_k(x - x_k) + (1/2)b_k(x - x_k)^2 + (1/3!)c_k(x - x_k)^3 + \tag{3.6}$$
$$+ (1/4!)d_k(x - x_k)^4,$$

$$b_k = (1/h^2)[f(x_{k+1}) - 2f(x_k) + f(x_{k-1}) - (h^3/12)(f^{(3)}(z_k) - f^{(3)}(z_{k-1}))] \tag{3.7}$$

and

$$S_0(x) = f(x_1) + a_1(x-x_1) + (1/2)b_1(x-x_1)^2 + (1/3!)c_1(x-x_1)^3 + \tag{3.8}$$
$$+ (1/4!)d_1(x-x_1)^4.$$

Then, for $k = 1, 2, \ldots, n-1$, the values

$$d_k = -\left(\frac{8}{5h^4}\right) 4![f(x_{k+1}) - f(x_k) - (h^2/4)b_k - 2f(z_k) + 2f(x_k)] + (24/5h)f^{(3)}(z_k), \tag{3.9}$$

$$c_k = \left(\frac{4}{5h^3}\right) 4![f(x_{k+1}) - f(x_k) - (h^2/4)b_k - 2f(z_k) + 2f(x_k)] - (7/5)f^{(3)}(z_k) \tag{3.10}$$

and

$$a_k = (1/h)[f(x_{k+1}) - f(x_k) - (h^2/2)b_k - (h^3/3!)c_k - (h^4/4!)d_k] \tag{3.11}$$

prove the theorem.

THEOREM 3.2. *Let $f \in C^4[0,1]$. Then for the unique spline $S_\Delta(x)$ associated with f and given in* Theorem 3.1, *we have for all $x \in [x_k, x_{k+1}]$, $k = 1, 2, \ldots, n-1$*

$$|S_\Delta^{(i)}(x) - f^{(i)}(x)| \leqq c_{4,i} h^{4-i} \omega_4(h), \qquad i = 0, 1, 2, 3, 4 \tag{3.12}$$

and for $x \in [x_0, x_1]$,

$$|S_0^{(i)}(x) - f^{(i)}(x)| \leqq c^*_{4,i} h^{4-i} \omega_4(h), \qquad i = 0, 1, 2, 3, 4 \tag{3.13}$$

where $\omega_4(\cdot)$ denotes the modulus of continuity of $f^{(4)}$,

$$c_{4,0} = 23/30, \quad c_{4,1} = 49/30, \quad c_{4,2} = 179/60, \quad c_{4,3} = 23/5, \quad c_{4,4} = 17/5,$$
$$c^*_{4,0} = 97/120, \quad c^*_{4,1} = 18/10, \quad c^*_{4,2} = 209/60, \quad c^*_{4,3} = 28/5, \quad c^*_{4,4} = 22/5.$$

Before proving this theorem, we state and prove some lemmas which will help us in arriving at the proof.

LEMMA 3.1. For b_k given in (3.7), the estimation

$$|b_k - f''(x_k)| \leqq (h^2/12)\omega_4(h)$$

holds true for all $k = 1, 2, \ldots, n-1$.

PROOF. We have for all $k = 1, 2, \ldots, n-1$,

$$f(x_{k+1}) = \sum_{j=0}^{3} (h^j)/j!)f^{(j)}(x_k) + (h^4/4!)f^{(4)}(\xi_0^{(k)}), \quad x_k < \xi_0^{(k)} < x_{k+1}, \tag{3.14}$$

$$f(x_{k-1}) = \sum_{j=0}^{3} [(-h^j)/j!]f^{(j)}(x_k) + (h^4/4!)f^{(4)}(\xi_1^{(k-1)}), \quad x_{k-1} < \xi_1^{(k-1)} < x_k, \tag{3.15}$$

$$f^{(3)}(z_k) = f^{(3)}(x_k) + (h/2)f^{(4)}(\eta_3^{(k)}, \quad x_k < \eta_3^{(k)} < z_k \tag{3.16}$$

and

$$f^{(3)}(z_{k-1}) = f^{(3)}(x_k) - (h/2)f^{(4)}(\eta_4^{(k-1)}), \quad z_{k-1} < \eta_4^{(k-1)} < x_k. \tag{3.17}$$

Using (3.14)–(3.17) in (3.8), it is easy to prove this lemma.

LEMMA 3.2. For d_k given in (3.4), the inequality

$$|d_k - f^{(4)}(x)| \leqq (17/5)\omega_4(h) \tag{3.18}$$

holds true for all $x \in [x_k, x_{k+1}]$ and $k = 1, 2, \ldots, n-1$.

PROOF. We have for all $k = 1, 2, \ldots, n-1$,

$$f(z_k) = \sum_{j=0}^{3} (h^j/2^j j!) f^{(j)}(x_k) + (h^4/2^4 4!) f^{(4)}(\eta_0^{(k)}), \quad x_k < \eta_0^{(k)} < z_k. \tag{3.19}$$

Using (3.14), (3.19) and Lemma 3.1 in (3.9), we easily get (3.18).

LEMMA 3.3. For c_k given in (3.10), the estimation

$$|c_k - f^{(3)}(x_k)| \leqq (6/5)h\omega_4(h)$$

holds true for all $k = 1, 2, \ldots, n-1$.

PROOF. Using (3.14), (3.16), (3.19) and Lemma 3.1 in (3.10), we get the required results.

LEMMA 3.4. For a_k given in (3.11), the inequality

$$|a_k - f'(x_k)| \leqq (23/60)h^3\omega_4(h)$$

holds true for all $k = 1, 2, \ldots, n-1$.

PROOF. Using (3.14), Lemma 3.1 and Lemma 3.3 in (3.11), Lemma 3.4 easily follows.

PROOF of THEOREM 3.2. We have for all $x \in [x_k, x_{k+1}]$ and all $k = 1, 2, \ldots, n-1$

$$f^{(i)}(x) = \sum_{j=i}^{3} [f^{(j)}(x_k)/(j-i)!](x-x_k)^{(j-i)} + (1/(4-i)!) f^{(4)}(\xi_i^{(k)})(x-x_k)^{(4-i)} \tag{3.20}$$

where $x_k < \xi_i^{(k)} < x_{k+1}$ and $i = 0, 1, 2, 3$.

Using (3.20) and (3.7) with the help of Lemmas 3.1–3.4, it will be easy to prove the theorem for $k = 1, 2, \ldots, n-1$ and $i = 0, 1, 2, 3$.

If $i = 4$, we then get the situation of Lemma 3.2. Hence the proposition (3.12) for $k = 1, 2, \ldots, n-1$.

For $x \in [x_0, x_1]$, let

$$f^{(i)}(x) = \sum_{j=i}^{3} [f^{(j)}(x_1)/(j-i)!](x-x_1)^{(j-i)} + (1/(4-i)!) f^{(4)}(\zeta_i)(x-x_1)^{(4-i)} \tag{3.21}$$

where $x_0 < \zeta_i < x_1$ and $i = 0, 1, 2, 3$. From (3.8) and (3.21), with the help of Lemma 3.3 and Lemma 3.4, it is easy to get (3.13) for $i = 0, 1, 2, 3$.

If $i = 4$, then

$$|S_0^{(4)}(x) - f^{(4)}(x)| = |d_1 - f^{(4)}(x)| \leqq |d_1 - f^{(4)}(x_1)| + |f^{(4)}(x_1) - f^{(4)}(x)| \leqq$$
$$\leqq (17/5)\omega_4(h) + \omega_4(h) = (22/5)\omega_4(h)$$

and this completes the proof of Theorem 3.2.

4. Case C. In this case $f \in C^5[0,1]$ and we consider the partition

$$\Delta : 0 = x_0 < x_1 < \ldots < x_k < x_{k+1} < \ldots < x_n = 1$$

where,

$$x_{k+1} - x_k = h \quad \text{and} \quad k = 0, 1, \ldots, n-1.$$

THEOREM 4.1. *Given arbitrary numbers* $f(x_k)$, $k = 2, 3, \ldots, n-1$, $f^{(p)}(z_k)$, $k = 0, 1, \ldots, n-1$; $p = 0, 3$; $2z_k = x_k + x_{k+1}$; *then there exists a unique spline* S_Δ *such that*

(4.1) $\quad S_\Delta \in \pi_5$ *on each* $[x_k, x_{k+1}]$, $k = 0, 1, \ldots, n-1$

(4.2) $\quad S_\Delta(x_k) = f(x_k)$, $k = 2, 3, \ldots, n-1$,

(4.3) $\quad S_\Delta^{(p)}(z_k) = f^{(p)}(z_k)$, $k = 0, 1, \ldots, n-1$; $p = 0, 3$,

(4.4) $\quad S_\Delta \in C^{(0,3)}[0, 1]$,

that is, both of S_Δ *and its third derivative is continuos for all* $x \in [0, 1]$ *and*

(4.5) $$S_\Delta^{(5)}(z_k) = e_k,$$

where,

$$e_k = [f^{(3)}(z_{k+1}) - 2f^{(3)}(z_k) + f^{(3)}(z_{k-1})]/h^2 \quad \textit{and} \quad k = 1, 2, \ldots, n-2.$$
(4.6)

PROOF. For all $x \in [0, 1]$, set

(4.7) $$S_\Delta(x) = \begin{cases} S_0^*(x), & x_0 \leqq x \leqq z_0, \\ S_0(x), & z_0 \leqq x \leqq z_1, \\ S_k(x), & z_k \leqq x \leqq z_{k+1}, \quad k = 1, 2, \ldots, n-2, \\ S_{n-1}(x), & z_{n-1} \leqq x \leqq x_n \end{cases}$$

where,

$$S_0^*(x) = f(z_0) + S_0'(z_0)(x - z_0) + (1/2)S_0''(z_0)(x - z_0)^2 + (1/3!)f^{(3)}(z_0)(x - z_0)^3 +$$
(4.8) $$+ (1/4!)S_0^{(4)}(z_0)(x - z_0)^4 + (1/5!)S_0^{(5)})(z_0)(x - z_0)^5,$$

(4.9) $$S_0(x) = \sum_{j=0}^{5} [S_1^{(j)}(z_1)/j!](x - z_1)^j,$$

$$S_k(x) = f(z_k)a_k(x-z_k)+(1/2!)b_k(x-z_k)^2+(1/3!)c_k(x-z_k)^3+$$

$$+(1/4!)d_k(x-z_k)^4+(1/5!)e_k(x-z_k)^5, \tag{4.10}$$

$$S_{n-1}(x) = f(z_{n-1})+S'_{n-2}(z_{n-1})(x-z_{n-1})+(1/2)S''_{n-2}(z_{n-1})(x-z_{n-1})^2+$$

$$+(1/3!)f^{(3)}(z_{n-1})(x-z_{n-1})^3+(1/4!)S^{(4)}_{n-2}(z_{n-1})(x-z_{n-1})^4+$$

$$+(1/5!)S^{(5)}_{n-2}(z_{n-1})(x-z_{n-1})^5 \tag{4.11}$$

and e_k in (4.10) is given by (4.6).

Thus, for $k = 1, 2, \ldots, n-2$, the values

$$d_k = [f^{(3)}(z_{k+1})-f^{(3)}(z_k)-(h^2/2)e_k]/h, \tag{4.12}$$

$$b_k = (4/h^2)\left[f(z_{k+1})+f(z_k)-(3h^3/4!)f^{(3)}(z_k)-2f(x_{k+1})-\right.$$

$$\left.-\left(\frac{7h^4}{8\cdot 4!}\right)d_k-\left(\frac{15\,h^5}{16\cdot 5!}\right)e_k\right] \tag{4.13}$$

and

$$a_k = [f(z_{k+1})-f(z_k)-(h^3/3!)f^{(3)}(z_k)-(h^2/2)b_k-(h^4/4!)d_k-(h^5/5!)e_k]/h \tag{4.14}$$

complete the proof of Theorem 4.1.

THEOREM 4.2. *Let $f\in C^5[0, 1]$. Then for the unique spline S_Δ given in* Theorem 4.1, *we have for all $x\in[z_k, z_{k+1}]$ and $k = 1, 2, \ldots, n-2$*

$$|S_k^{(i)}(x)-f^{(i)}(x)\leqq\alpha_i h^{5-i}\omega_5(h), \tag{4.15}$$

for $z_0\leqq x\leqq z_1$,

$$|S_0^{(i)}(x)-f^{(i)}(x)|\leqq\beta_i h^{5-i}\omega_5(h), \tag{4.16}$$

for $x_0\leqq x\leqq z_0$,

$$|S_0^{*(i)}(x)-f^{(i)}(x)|\leqq\gamma_i h^{5-i}\omega_5(h) \tag{4.17}$$

and for $z_{n-1}\leqq x\leqq x_n$,

$$|S_{n-1}^{(i)}(x)-f^{(i)}(x)|\leqq\delta_i h^{5-i}\omega_5(h) \tag{4.18}$$

where in (4.15)–(4.18), $i = 0, 1, 2, 3, 4, 5$ *and*

$$\alpha_0 = 1/4,\quad \alpha_1 = 19/40,\quad \alpha_2 = 63/80,\quad \alpha_3 = 3/2,$$

$$\beta_0 = 31/120,\quad \beta_1 = 31/60,\quad \beta_2 = 229/240,\quad \beta_3 = 2,$$

$$\gamma_0 = 743/1920,\quad \gamma_1 = 4111/3840,\quad \gamma_2 = 43/30,\quad \gamma_3 = 33/16,$$

$$\delta = 527/1536,\quad \delta_1 = 1187/1280,\quad \delta_2 = 553/480,\quad \delta_3 = 25/16,$$

$$\alpha_4 = 5/2, \quad \alpha_5 = 3/2,$$
$$\beta_4 = 13/4, \quad \beta_5 = 5/2,$$
$$\gamma_4 = 5, \quad \gamma_5 = 7/2,$$
$$\delta_4 = 15/4, \quad \delta_5 = 5/2.$$

Before proving this theorem, we prove some lemmas which will help us in arriving at the proof of Theorem 4.2.

LEMMA 4.1. For e_k given in (4.6) we have

$$|e_k - f^{(5)}(x)| \leqq (3/2)\omega_5(h)$$

which holds true for all $x \in [z_k, z_{k+1}]$ and all $k = 1, 2, \ldots, n-2$.

PROOF. From (4.6) we have for all $x \in [z_k, z_{k+1}]$ and all $k = 1, 2, \ldots, n-2$

$$|e_k - f^{(5)}(x)| = [f^{(3)}(z_k) + hf^{(4)}(z_k) + (h^2/2)f^{(5)}(t_2) - 2f^{(3)}(z_k) + f^{(3)}(z_k) - \\ - hf^{(4)}(z_k) + (h^2/2)f^{(5)}(t_1)] - f^{(5)}(x)|$$

where $z_k < t_2 < z_{k+1}$ and $z_{k-1} < t_1 < z_1$.

Thus, we easily get

$$|e_k - f^{(5)}(x)| \leqq (1/2|f^{(5)}(t_2) - f^{(5)}(x)| + (1/2)|f^{(5)}(t_1) - f^{(5)}(x)| \leqq \\ \leqq (1/2)\omega_5(h) + (1/2)\omega_5(2h) \leqq (3/2)\omega_5(h).$$

LEMMA 4.2. For d_k given in (4.12), we have for all $k = 1, 2, \ldots, n-2$

$$|d_k - f^{(4)}(z_k)| \leqq (3/4)h\omega_5(h).$$

PROOF. From (4.12) we get

$$|d_k - f^{(4)}(z_k)| = |(1/h)[f^{(3)}(z_{k+1}) - f^{(3)}(z_k)] - (h/2)e_k - f^{(4)}(z_k)| \leqq \\ \leqq (h/2)|f^{(5)}(\xi_3^{(k)}) - e_k|, \quad z_k < \xi_3^{(k)} < z_{k+1}.$$

Using Lemma 4.1, we get the required result.

LEMMA 4.3. For b_k given in (4.13), we have for all $k = 1, 2, \ldots, n-2$

$$|b_k - f''(z_k)| \leqq (13/80)h^3\omega_5(h).$$

PROOF. From (4.13) we get

$$|b_k - f''(z_k)| = \Bigg| (4/h^2)\Bigg[(h^2/2)f''(z_k) + h^4\left(\frac{1}{4!} - \frac{1}{2\cdot 4!}\right)f^{(4)}(z_k) + \\ + h^4\left(\frac{1}{2\cdot 4!} - \frac{1}{4!}\right)d_k + (h^5/5!)\left(f^{(5)}(\xi_0^{(k)}) - e_k\right) + \\ + \left(\frac{1}{2^4\cdot 5!}\right)h^5\left(e_k - f^{(5)}(t_3)\right)\Bigg] - f''(z_k)\Bigg|$$

where $z_k < \xi_0^{(k)} < z_{k+1}$ and $z_k < t_3 < z_{k+1}$.

Using Lemma 4.2, it is easy to prove Lemma 4.3.

LEMMA 4.4. For a_k given in (4.14), we have for all $k = 1, 2, \ldots, n-2$

$$|a_k - f'(z_k)| \leqq (1/8)h^4\omega_5(h)$$

PROOF. We have from (4.14)

$$|a_k - f'(z_k)| = (h/2)|b_k - f''(z_k)| + (h^3/4!)|d_k - f^{(4)}(z_k)| + \\ + (h^4/5!)|e_k - f^{(5)}(\xi_0^{(k)})|$$

where $z_k < \xi_0^{(k)} < z_{k+1}$. Using Lemmas 4.3, 4.2 and 4.1, then Lemma 4.4 follows.

PROOF of THEOREM 4.2. For $x \in [z_k, z_{k+1}]$ and $k = 1, 2, \ldots, n-2$, using (4.10) we get

$$|S_k(x) - f(x)| \leqq h|a_k - f'(z_k)| + (h^2/2)|b_k - f''(z_k)| + (h^4/4!)|d_k - f^{(4)}(z_k)| + \\ + (h^5/5!)|e_k - f^{(5)}(\eta_0)|$$

where $z_k < \eta_0 < z_{k+1}$.

Using Lemmas 4.1 – 4.4, we get easily

$$|S_k(x) - f(x)| \leqq \alpha_0 h^5 \omega_5(h) \qquad \text{with } \alpha_0 = 1/4.$$

Similarly, using Lemmas 4.1 – 4.4 and (4.10), it is easy to prove (4.15) of the theorem for $i = 1, 2, 3, 4$ and 5.

Thus, using the results of (4.15) with (4.9) and (4.11), we easily get (4.16) and (4.18). Using (4.16) in (4.8), we get (4.17), and this completes the proof.

5. Stability. We conclude this note with stability results concerning the Cases **A, B** and **C.** We state the stability theorem for each case and prove it for the Case **A** only, while in the other two cases, the proofs are similar to that of Case **A.**

THEOREM 5.1. *Let $f \in C^3[0,1]$ and let $\bar{S}_n$ be the unique spline constructed in the same manner as that of* Theorem 2.1 *and satisfying the following data:*

$$\bar{S}_n(x_k) = \alpha_{k,0}, \qquad k = 0, 1, \ldots, n, \tag{5.1}$$

$$\bar{S}_n(z_k) = \beta_{k,0}, \qquad k = 0, 1, \ldots, n-1, \tag{5.2}$$

$$\bar{S}_n^{(3)}(z_k) = \beta_{k,3}, \qquad k = 0, 1, \ldots, n-1 \tag{5.3}$$

where we suppose that there exists a function $F(f, n)$ such that

$$\omega_3(h)\, h^3\, F(f, n) \geqq \max_k |f(x_k) - \alpha_{k,0}|, \tag{5.4}$$

$$\omega_3(h)\, h^3\, F(f, n) \geqq \max_k |f(z_k) - \beta_{k,0}| \tag{5.5}$$

and

$$\omega_3(h) F(f, n) \geqq \max_k |f^{(3)}(z_k) - \beta_{k,3}|. \tag{5.6}$$

Then there are constants K_i, independent of f, F and n, such that

$$F(S, n)K_i h^{3-i}\omega_3(h) \geqq ||D^i(f-\bar{S}_n)||_\infty, \qquad i = 0, 1, 2, 3$$

where $\|\cdot\|_\infty \equiv \|\cdot\|_{L_\infty[0,1]}$ *and*

$$K_0 = 109/6, \qquad K_1 = 317/12, \qquad K_2 = 115/6 \qquad K_3 = 2.$$

PROOF. Analogous to (2.4), then for $x \in [x_k, x_{k+1}]$ and $k = 0, 1, \ldots, n-1$

$$\bar{S}_n(x) = \alpha_{k,0} + \bar{a}_k(x-x_k) + (1/2)\bar{b}_k(x-x_k)^2 + (1/3!)\bar{c}_k(x-x_k)^3 \tag{5.7}$$

where

$$\bar{a}_k = (1/h_k)[4\beta_{k,0} - 3\alpha_{k,0} - \alpha_{k+1,0} + (h^3/12)\beta_{k,3}], \tag{5.8}$$

$$\bar{b}_k = (1/h_k^2)[4\alpha_{k+1,0} + 4\alpha_{k,0} - 8\beta_{k,0} - (h_k^3/2)\beta_{k,3}] \tag{5.9}$$

and

$$\bar{c}_k = \beta_{k,3}. \tag{5.10}$$

Using (2.5) with (5.8), (2.6) with (5.9) and (2.7) with (5.10), with the help of (5.4), (5.5) and (5.6), the following estimations could be easily obtained for all $k = 0, 1, \ldots, n-k$

$$|\bar{a}_k - a_k| \leqq (79/12)h^2\omega_3(h)\, F(f, n), \tag{5.11}$$

$$|\bar{b}_k - b_k| \leqq (33/2)h\omega_3(h)\, F(f, n) \tag{5.12}$$

and

$$|\bar{c}_k - c_k| \leqq \omega_3(h)F(f, n). \tag{5.13}$$

Thus, the above three inequalitites (5.11)–(5.13) with (5.7) and (2.4) give for all $x \in [x_k, x_{k+1}]$ and $k = 0, 1, \ldots, n-1$,

$$|\bar{S}_n(x) - S_n(x)| \leqq (35/2)h^3\omega_3(h)\, F(f, n), \tag{5.14}$$

$$|\bar{S}_n'(x) - S_n'(x)| \leqq (301/12)h\omega_3(h)F(f, n), \tag{5.15}$$

$$|\bar{S}_n''(x) - S_n''(x)| \leqq (35/2)h\omega_3(h)\, F(f, n) \tag{5.16}$$

and

$$|\bar{S}_n^{(3)}(x) - S_n^{(3)}(x)| \leqq \omega_3(h)Ff, n). \tag{5.17}$$

Using (5.14)–(5.17) with the help of Theorem 2.2, we easily get

$$|\bar{S}_n(x) - f(x)| \leqq |\bar{S}_n(x) - S_n(x)| + |S_n(x) - f(x)| \leqq (109/6)h^3\omega_3(h)F(f, n), \tag{5.18}$$

$$|\bar{S}_n'(x) - f'(x)| \leqq |\bar{S}_n'(x) - S_n'(x)| + |S_n'(x) - f'(x)| \leqq (317/12)h^2\omega_3(h)F(f, n), \tag{5.19}$$

$$|\bar{S}_n''(x)-f''(x)| \leqq |\bar{S}_n''(x)-S_n''(x)|+|S_n''(x)-f''(x)| \leqq (115/6)h\omega_3(h)F(f,n) \tag{5.20}$$

and

$$|\bar{S}_n^{(3)}(x)-f^{(3)}(x)| \leqq |\bar{S}_n^{(3)}(x)-S_n^{(3)}(x)|+|S_n^{(3)}(x)-f^{(3)}(x)| \leqq 2\omega_3(h)F(f,n) \tag{5.21}$$

and the proof of Theorem 5.1 is now complete.

THEOREM 5.2. *Let $f \in C^4[0, 1]$ and let $\bar{S}_\Delta^*$ be the unique spline constructed in the same manner as that of* Theorem 3.1 *and satisfying the following data :*

$$\bar{S}_\Delta^*(x_k) = \alpha_{k,0}^*, \qquad k = 1, 2, \ldots, n, \tag{5.22}$$

$$\bar{S}_\Delta^*(z_k) = \beta_{k,0}^*, \qquad k = 1, 2, \ldots, n-1, \tag{5.23}$$

$$\bar{S}_\Delta^{*(3)}(z_k) = \beta_{k,3}^*, \quad k = 1, 2, \ldots, n-1, \tag{5.24}$$

$$\bar{S}_\Delta^{*''}(x_k) = (1/h^2)[\alpha_{k+1,0}^*-2\alpha_{k,0}^*+ \tag{5.25}$$
$$+\alpha_{k-1,0}^*-(h^3/12)(\beta_{k,0}^*-\beta_{k-1,0}^*)], \quad k = 1, 2, \ldots, n-1,$$

where we suppose that there exists a function $F^*(f, n)$ such that

$$\omega_4(h)h^4F(f,n) \geqq \max_k |f(x_k)-\alpha_{k,0}^*|, \tag{5.26}$$

$$\omega_4(h)h^4F(f,n) \geqq \max_k |f(z_k)-\beta_{k,0}^*| \tag{5.27}$$

and

$$\omega_4(h)hF(f,n) \geqq \max_k |f^{(3)}(z_k)-\beta_{k,3}^*|. \tag{5.28}$$

Then there are constants K_i^*, independent of f, F^* and n, such that

$$K_i^*h^{4-i}\omega_4(h)F(f,n) \geqq ||D^i(f-\bar{S}^*)||_\infty, \qquad i = 0, 1, 2, 3, 4$$

where $\|\cdot\|_\infty \equiv \|\cdot\|_{L_\infty[0,1]}$.

THEOREM 5.3. *Let $f \in C^5[0, 1]$ and let $\bar{S}_\Delta^{**}$ be the unique spline constructed in the same manner as that of* Theorem 4.1 *and satisfying the following data:*

$$\bar{S}_\Delta^{**}(x_k) = \alpha_{k,0}^{**}, \qquad k = 2, 3, \ldots, n-1, \tag{5.29}$$

$$\bar{S}_\Delta^{**}(z_k) = \beta_{k,0}^{**}, \qquad k = 0, 1, \ldots, n-1, \tag{5.30}$$

$$\bar{S}_\Delta^{**(3)}(z_k) = \beta_{k,3}^{**}, \qquad k = 0, 1, \ldots, n-1, \tag{5.31}$$

$$\bar{S}_\Delta^{**(5)}(z_k) = (1/h^2)[\beta_{k+1,3}^{**}-2\beta_{k,3}^{**}+\beta_{k-1,3}^{**}], \quad k = 1, 2, \ldots, n-2 \tag{5.32}$$

where we suppose that there exists a function $F^{**}(f, n)$ such that

$$\omega_5(h)h^5F^{**}(f, n) \geqq \max_k |f(x_k) - \alpha_{k,0}^{**}|, \tag{5.33}$$

$$\omega_5(h)h^5F^{**}(f, n) \geqq \max_k |f(z_k) - \beta_{k,0}^{**}|, \tag{5.34}$$

$$\omega_5(h)h^2F^{**}(f, n) \geqq \max_k |f^{(3)}(z_k) - \beta_{k,3}^{**}|. \tag{5.35}$$

Then there exist constants K_i^{**}, independent of f, F^{**} and n, such that

$$K_i^{**}h^{5-i}\omega_5(h) \geqq \|D^i(f - S_\Delta^{**})\|_\infty, \qquad i = 0, 1, 3, 4, 5$$

where $\|\cdot\|_\infty \equiv \|\cdot\|_{L_\infty[0,1]}$.

As we have mentioned before, the proofs of the last two theorems are similar to that of Theorem 5.1.

Finally, to illustrate our method, a numerical example is given. The method described in Case **B** is applied to the function $f(x) = 1 + xe^x$ and the following results are obtained for $x = 0{,}86$ and $h = 0{,}1$.

	Numerical value	Exact value	The error
$f(0.86)$	3.0323231	3.0323176	$5.5 \cdot (10)^{-6}$
$f'(0.86)$	4.3949245	4.3954776	$5.531 \cdot (10)^{-4}$
$f''(0.86)$	6.754668	6.7586376	$3.9696 \cdot (10)^{-3}$
$f^{(3)}(0.86)$	9.1491812	9.127976	$2.73836 \cdot (10)^{-2}$
$f^{(4)}(0.86)$	14.15292	11.484957	2.6679624

Note that for $h = 0.1$, $\omega_4(h) = 1.53936$.

References

[1] BALÁZS, J. and TURÁN, P.: Notes on interpolation II, III, IV, *Acta Math. Acad. Sci. Hungar.*, **8** (1957), 201–215, **9** (1958), 195–214, **9** (1958), 243–252.

[2] MEIR, A. and SHARMA, A.: Lacunary interpolation by splines, *SIAM. J. Num. Anal.*, **10** (1973), 433–442.

[3] SWARTZ, B, K. and VARGA, R. S.: A note on lacunary interpolation by splines, *SIAM. J. Num. Anal.*, **10** (1973), 443–447.

[4] DEMKO, S.: Lacaunary polynomial spline interpolation, *SIAM. J. Num. Anal.*, **13** (1976), 369–381.

[5] VARMA, A. K.: Lacunary interpolation by splines I. II, *Acta Math. Sci. Hungar.*, **31** (1978), 185–192, 193–203.

[6] PRASAD, J. and VARMA, A. K.: Lacunary interpolation by quintic splines, *SIAM. J. Num. Anal.*, **16** (1979), 1075–1079.

GENERALISED DIRECT SUMMANDS OF ABELIAN GROUPS

By

M. J. SCHOEMAN

University of Pretoria

(Received December 8, 1982)

Dedicated to F. Loonstra, a great teacher of mathematics

1. Introduction

The investigation into the existence of an automorphism α of a group A such that α acts as the identity on the torsion subgroup T of A and as -1 on the factor group A/T, has attracted attention in various papers. From [4] we know conditions under which the existence of such an α implies that A is splitting. It is still an open question whether A splits if there exists to each pair (β, γ) of automorphisms of T and A/T respectively, an automorphism inducing β and γ.

MADER [4] considered the following more general situation. Let B be an arbitrary subgroup of A and α an automorphism of A which acts as the identity on B and as -1 on A/B. We then have the following ([4] Proposition 2.2):

(*) Let such an α exist and let $A/B\,[2] = 0$. If either $2B = B$ or $2(A/B) = A/B$, then B is a direct summand of A.

In [3] it was noted that if α is an automorphism of A which acts as the identity on T and as -1 on A/T, then $2A \leq T \oplus C$ for some subgroup C of A, and hence A is quasi-splitting (in the sense of WALKER [5]). It is the purpose of this paper to elaborate on results in [3] by considering *generalised direct summands* of a group A; defined as subgroups B of A such that $rA \leq B \oplus C$ for some subgroup C of A and some non-zero integer r.

In section 2 we shall consider endomorphisms of A which act as multiplication by (different) integers n and m on the subgroup B and the factor group A/B respectively. It will be seen (Proposition 2.4) that if either $B[n-m] = 0$ or $(n-m)\,a \in B$, $a \in A$, implies $a \in B$; then such an endomorphism exists if and only if $(n-m)A \leq B \oplus C$ for some subgroup C of A. Conditions under which generalised direct summands are direct summands (in the ordinary sense) are also given, showing, inter alia, how MADER's result mentioned above, fits into a much more general theory.

In section 3 we go a step further by considering subroups B of A for which there exists a subgroup C of A such that $\Phi A \leq B \oplus C$, for some endomorphism Φ of A. For reasons which will become clear, it will be necessary to restrict

ourselves to those endomorphisms Φ which map B into itself. Given endomorphisms α and β of A (each of which maps B into itself), we shall inquire into the existence of an endomorphism Φ of A which agrees with α on B, while the endomorphisms induced on A/B by Φ and β are equal. Proposition 3.2 places the results of section 2 in full perspective.

The notation followed is that of FUCHS [1] and [2]. In particular, $E(A)$ denotes the endomorphism ring of the group A. If $r \neq 0$ is an integer, the endomorphism $a \to ra$, $a \in A$, is denoted by $\bar{r}_A$, and if no danger of confusion exists we simply write $\bar{r}$. For a subgroup B of A and an integer $r \neq 0$, let

$$r^{-1}B = \{a \in A \mid ra \in B\} \text{ and } B[r] = \{x \in B \mid rx = 0\}.$$

We say that B satisfies *conditions* $S(r)$ if either $r^{-1}B = B$ or $B[r] = 0$. The torsion subgroup of A satisfies condition $S(r)$ for all integers $r \neq 0$, and so does any subgroup of a torsionfree group. Also, if A is a mixed group and B a pure subgroup of A containing the torsion subgroup of A, then B satisfies condition $S(r)$ for all integers $r \neq 0$.

2. Generalised direct summands

A subgroup B of A is called a *generalised direct summand* if there exists an integer $r \neq 0$ such that $rA \leq B \oplus C$ for some subgroup C of A. If B is the torsion subgroup of A then this notion corresponds with that of quasi-splitting groups introduced by WALKER [5].

We commence by giving some necessary and sufficient conditions for B to be a generalised direct summand of A.

(2.1) If $rA \leq B \oplus C \leq A$, $0 \neq r \in \mathbf{Z}$, then there exists a commutative triangle

$$\begin{tikzcd} 0 \arrow[r] & B \arrow[r] \arrow[d, "\bar{r}"'] & A \arrow[dl, dashed, "\alpha"] \\ & B & \end{tikzcd}$$

Fig. 1.

with exact row.

Conversely, if such a triangle exists and if $B[r] = 0$ then $rA \leq B \oplus \operatorname{Ker} \alpha$.

PROOF. If $rA \leq B \oplus C \leq A$ and $\pi : B \oplus C \to B$ is the projection, then $\alpha = \pi \bar{r}_A$ makes the triangle commutative.

Conversely, if such an α exists, then $f = \bar{r}_A - \alpha \in E(A)$ maps B onto 0, and hence $f\alpha = 0$. For every $a \in A$ we therefore have $ra = \alpha a + fa$ where $\alpha a \in B$ and $fa \in \operatorname{Ker} \alpha$. Moreover, since $B[r] = 0$ we have $B \cap \operatorname{Ker} \alpha = 0$ and so $rA \leq B \oplus \operatorname{Ker} \alpha \leq A$. ∎

(2.2) If $rA \leq B \oplus C \leq A$, $0 \neq r \in \mathbf{Z}$, then there exists a commutative triangle

$$\begin{tikzcd} A \arrow[r] & A/B \arrow[r] & 0 \\ & A/B \arrow[ul, dashed, "\beta"] \arrow[u, "\bar{r}"'] & \end{tikzcd}$$

Fig. 2.

with exact row.

Conversely, if such a triangle exists and if $r^{-1}B = B$ then $rA \leq B \oplus \operatorname{Im} \beta$.

PROOF. If $a \in A$ and $ra = b + c$, $b \in B$, $c \in C$, then β defined by $\beta(a+B) = c$ is a homomorphism making the triangle commutative.

Conversely, if such a β exists then $\beta(a+B)+B = ra+B$ for all $a \in A$, and hence $rA \leq B + \operatorname{Im} \beta$. Furthermore, if for some $a \in A$ we have $\beta(a+B) \in B$ then $\beta(a+B)+B = ra+B = B$ and so $r^{-1}B = B$ implies that $a \in B$, that is $\beta(a+B) = 0$. Hence $B \cap \operatorname{Im} \beta = 0$ and thus $rA \leq B \oplus \operatorname{Im} \beta \leq A$. ■

It is well-known that there is a one-to-one correspondence between direct summands of A and endomorphisms π of A satisfying $\pi^2 = \pi$. Using the same techniques as in (2.1) and (2.2) it is not difficult to show the following analogy for generalised direct summands.

(2.3) If $rA \leq B \oplus C \leq A$, $0 \neq r \in \mathbf{Z}$, then there exists an endomorphism Φ of A such that $\Phi^2 = r\Phi$.

Conversely, if B satisfies condition $S(r)$ and if $\Phi : A \to B$ is an epimorphism with $\Phi^2 = r\Phi$, then $rA \leq B \oplus (\bar{r} - \varphi)\, A$. ■

Let B be a subgroup of A and suppose there exists an endomorphism α of A such that $\alpha x = nx$, $x \in B$, and $\alpha a - ma \in B$, $a \in A$, where n and m are integers. Note that if $n = m$ then always such an endomorphism exists, namely $\alpha = \bar{n}_A$. This trivial case will henceforth be excluded. If B is fully invariant in A then we can describe this situation by saying that $\alpha \in E(A)$ induces the pair $(\bar{n}, \bar{m}) \in E(B) \times E(A/B)$.

The first proposition gives a correspondence between generalised direct summands of A and endomorphisms of A acting in the way mentioned above.

PROPOSITION 2.4. *Let B be a subgroup of A and α an endomorphism of A such that $\alpha x = nx$, $x \in B$ and $\alpha a - ma \in B$, $a \in A$, where m and n are integers, $m \neq n$. If B satisfies condition $S(n-m)$ then $(n-m)A \leq B \oplus C$ for some $C \leq A$.*

Conversely, if $rA \leq B \oplus C \leq A$ for some integer $r \neq 0$, then for any pair n, m of integers with $n - m = r$, there exists an $\alpha \in E(A)$ such that $\alpha x = nx$, $x \in B$, and $\alpha a - ma \in B$, $a \in A$.

PROOF. If $a \in A$ then the hypothesis on α implies that $\alpha a - ma \in B$. Define $\beta : A \to B$ by $\beta a = \alpha a - ma$, and $\gamma : A/B \to A$ by $\gamma(a+B) = na - \alpha a$, $a \in A$. We have commutative triangles

with exact rows and so the second parts of (2.1) and (2.2) complete the first part of the theorem.

$$\begin{tikzcd} 0 \arrow[r] & B \arrow[r] \arrow[d, "\overline{(n-m)}"'] & A \arrow[dl, dashed] \\ & B & \end{tikzcd} \qquad \text{and} \qquad \begin{tikzcd} A \arrow[r] & A/B \arrow[r] & 0 \\ & A/B \arrow[ul, dashed, "\gamma"] \arrow[u, "\overline{(n-m)}"'] & \end{tikzcd}$$

Fig. 3.

Conversely, if $rA \leq B \oplus C \leq A$, $0 \neq r \in \mathbf{Z}$, then (2.1) and (2.2) imply that we have commutative triangles

$$\begin{tikzcd} A \arrow[r] & A/B \arrow[r] & 0 \\ & A/B \arrow[ul, dashed, "\psi"] \arrow[u, "\bar{r}"'] & \end{tikzcd} \qquad \text{and} \qquad \begin{tikzcd} 0 \arrow[r] & B \arrow[r] \arrow[d, "\bar{r}"'] & A \arrow[dl, dashed, "\phi"] \\ & B & \end{tikzcd}$$

Fig. 4.

with exact rows. Let $r = n-m$ and define $\alpha : A \to A$ by $\alpha a = \Phi a + \Phi\psi(a+B) + ma$, $a \in A$. If $x \in B$ then $\alpha x = \Phi x + mx = nx - mx + mx = nx$. Also, if $a \in A$ then $\alpha a - ma = \Phi a + \Phi\psi(a+B) \in B$. ■

We now give two conditions under which generalised direct summands are direct summands.

LEMMA 2.5 [3]. Let B be a direct summand of rA for some $0 \neq r \in \mathbf{Z}$. If $rB = B$ then B is a direct summand of A. ■

COROLLARY 2.6. If $rA \leq B \oplus C \leq A$, $0 \neq r \in \mathbf{Z}$, and $rB = B$ then B is a direct summand of A.

PROOF. If $rA \leq B \oplus C$ then $rA + B = B \oplus (C \cap (rA+B))$. Since $rB = B$ we have $rA = B \oplus (C \cap (rA+B))$, and so Lemma 2.5 completes the proof. ■

LEMMA 2.7. If $rA \leq B \oplus C \leq A$, $0 \neq r \in \mathbf{Z}$, and $r(A/B) = A/B$, then B is a direct summand of A.

PROOF. Since any $a \in A$ is of the form $a = rx + b$, $x \in A$, $b \in B$, the result follows immediately. ■

We are now ready to show how MADER's result (mentioned in the introduction) fits into the more general theory described above. It also shows how the requirement $A/B[2] = 0$ can be replaced by the weaker condition that either $A/B[2] = 0$ or $B[2] = 0$. (The reader should bear in mind that an endomorphism of A which acts as the identity on B and as -1 on A/B, is in fact an automorphism of A).

PROPOSITION 2.8. *Let B be a subgroup of A satisfying condition $S(r)$ and $n, m \in \mathbf{Z}$ with $n-m = r$. Suppose further that either $rB = B$ or $r(A/B) = A/B$. Then B is a direct summand of A if and only if there exists an $\alpha \in E(A)$ such that $\alpha x = nx$, $x \in B$ and $\alpha a - ma \in B$, $a \in A$.*

PROOF. If B is a direct summand of A, say $A = B \oplus C$, and $a \in A$ with $a = b+c$, $b \in B$, $c \in C$; then α defined by $\alpha a = nb+mc$ is an endomorphism of A having the required properties.

The converse follows from Proposition 2.4, Corollary 2.6 and Lemma 2.7. ■

The following places Proposition 2.4 of [4] in perspective.

COROLLARY 2.9. Let A be a torsion group and B a subgroup of A such that either $A[p] = 0$ of $A/B\,[p] = 0$, p a prime number. Then B is a direct summand of A if and only if there exists an $\alpha \in E(A)$ such that $\alpha x = nx$, $x \in B$, and $\alpha a - ma \in B$, $a \in A$, where m and n are integers with $n-m = p$.

PROOF. A torsion group which does not contain elements of prime order p, is divisible by p. ■

3. A more general situation

In this section we go a step further by considering a subgroup B of A and a pair $(\alpha', \bar{\beta}) \in E(B) \times E(A/B)$ where α' can be extended to an endomorphism α of B and $\bar{\beta}$ can be lifted to an endomorphism β of A, $\alpha \neq \beta$. We wish to find conditions under which $(\alpha-\beta)A \leq B \oplus C$ for some subgroup C of A, thus extending the results of the previous section to, what we believe, the limit.

Let $\operatorname{Ann} B = \{f \in E(A) | fB = 0\}$ and $E_B(A) = \{\Phi \in E(A) | \Phi B \leq B\}$. Every $\Phi \in E_B(A)$ thus induces (by restriction) an endomorphism Φ' of B and an endomorphism $\bar{\Phi}$ of $E(A/B)$, the latter being defined by $\bar{\Phi}(a+B) = \Phi a + B$, $a \in A$, such that

$$\begin{array}{ccccccccc} E: 0 & \to & B & \to & A & \to & A/B & \to & 0 \\ & & \downarrow{\scriptstyle \Phi'} & & \downarrow{\scriptstyle \Phi} & & \downarrow{\scriptstyle \bar{\Phi}} & & \\ E: 0 & \to & B & \to & A & \to & A/B & \to & 0 \end{array}$$

is a commutative diagram. It is well-known that Φ' and $\bar{\Phi}$ induce endomorphisms Φ'_* and $\bar{\Phi}^*$ of Ext $(A/B, B)$. These maps are given by

$$\Phi'_* : E \to \Phi' E \quad \text{and} \quad \bar{\Phi}^* : E \to E\bar{\Phi}$$

where $\Phi' E$ and $E\bar{\Phi}$ are defined in [1] (Section 50). Since the above diagram commutes we have $\Phi' E \equiv E\bar{\Phi}$ and thus Ker $\Phi'_* =$ Ker $\bar{\Phi}^*$. This fact will play an important role in the rest of this section.

For $\Phi \in E_B(A)$, let $\Phi^{-1}B = \{a \in A | \Phi a \in B\}$. Then $\bar{\Phi}$ is injective if and only if $\Phi^{-1}B = B$, while Φ' is injective if and only if $B \cap \operatorname{Ker} \Phi = 0$.

LEMMA 3.1. Let E denote the extension $0 \to B \to A^j \to A/B \to 0$ and let $\Phi \in E_B(A)$. If $\Phi A \leq B \oplus C \leq A$ then $E \in \operatorname{Ker} \Phi'_* = \operatorname{Ker} \overline{\Phi}^*$.

Conversely, let $E \in \operatorname{Ker} \Phi'_* = \operatorname{Ker} \overline{\Phi}^*$. If either

(i) Φ' is injective and $f\Phi = \Phi f$ for all $f \in \operatorname{Ann} B$ or

(ii) $\overline{\Phi}$ is injective

then $\Phi A \leq B \oplus C \leq A$.

PROOF. If $\Phi A \leq B \oplus C \leq A$ and $\pi : B \oplus C \to B$ the corresponding projection, then $\pi\Phi$ is an endomorphism of A such that $\pi\Phi x = \Phi' x, x \in B$. Hence $\Phi' E$ splits ([1] p 218) and so $E \in \operatorname{Ker} \Phi'_* = \operatorname{Ker} \overline{\Phi}^*$.

Conversely, assume that $E \in \operatorname{Ker} \Phi'_* \equiv \operatorname{Ker} \overline{\Phi}^*$. Thus ([1] p 218) we know that homomorphisms η and ψ exist such that the following triangles commute

$$\begin{tikzcd} 0 \arrow[r] & B \arrow[r] \arrow[d, "\Phi'"'] & A \arrow[dl, dashed, "\eta"] \\ & B & \end{tikzcd} \qquad \begin{tikzcd} A \arrow[r, "j"] & A/B \arrow[r] & 0 \\ & A/B \arrow[ul, dashed, "\psi"] \arrow[u, "\overline{\Phi}"'] & \end{tikzcd}$$

Fig. 5.

Suppose that (i) holds. From the first diagram we have $\Phi - \eta = f \in \operatorname{Ann} B$ and hence that $f\Phi = \Phi f$. From $f\Phi - f\eta = f^2 = \Phi f - \eta f$ and the fact that $f\eta = 0$ it follows that $\eta f = 0$. It is now easy to show that $\Phi A \leq B \oplus \operatorname{Ker} \eta \leq A$.

Next, suppose that (ii) holds. If $x \in B \cap \operatorname{Im} \psi$ then $x = \psi(a+B)$ for some $a \in A$. Hence $B = jx = j\psi(a+B) = \overline{\Phi}(a+B) = \Phi a + B$. Thus $\Phi a \in B$ and since $\overline{\Phi}$ is injective we have $a \in B$, proving that $B \cap \operatorname{Im} \psi = 0$. For any $a \in A$ we have $j\psi(a+B) = \Phi a + B$ and so $\Phi a \in B + \operatorname{Im} \psi$. Thus $\Phi A \leq B \oplus \operatorname{Im} \psi \leq A$. ∎

It is clear that if α' is an endomorphism of B which can be extended to an endomorphism α of A, then $\alpha \in E_B(A)$. Also, if $\bar{\beta}$ is an endomorphism of A/B which can be lifted to an endomorphism β of A, then $\beta \in E_B(A)$. Consequently, $\alpha - \beta \in E_B(A)$. For obvious reasons the case $\alpha = \beta$ will be excluded. We are now ready to discuss the situation $(\alpha - \beta)A \leq B \oplus C$ for some subgroup C of A.

As before, $(\alpha - \beta)'$ denotes the restriction of $\alpha - \bar{\beta}$ to B and $\overline{(\alpha - \beta)}$ denotes the endomorphism of A/B with

$$\overline{(\alpha - \beta)}(a+B) = (\alpha - \beta)a + B, \ a \in A.$$

PROPOSITION 3.2. *Let B be a subgroup of A and $(\alpha', \bar{\beta}) \in E(B) \times E(A/B)$ where α' can be extended to an endomorphism α of A and $\bar{\beta}$ be lifted to an en-*

domorphism β of A, $\alpha \neq \beta$. Suppose further that a $\Phi \in E_B(A)$ exists such that $\Phi x = \alpha x$, $x \in B$ and $\Phi a - \beta a \in B$, $a \in A$. If either

(i) *$(\alpha - \beta')$ is injective and $f(\alpha - \beta) = (\alpha - \beta) f$ for all $f \in \operatorname{Ann} B$ or*

(ii) *$(\alpha - \beta)$ is injective*

then $(\alpha - \beta) A \leq B \oplus C \leq A$.

Converseley, if α, $\beta \in E_B(A)$ and $(\alpha - \beta) A \leq B \oplus C \leq A$, then a $\Phi \in E_B(A)$ exists such that $\Phi x = \alpha x$, $x \in B$ and $\alpha a - \beta a \in B$, $a \in A$.

PROOF. To prove the first part of the theorem, we notice that the hypothesis on α' and $\bar{\beta}$ implies the commutativity of the diagrams

$$\begin{tikzcd} E: 0 \arrow[r] & B \arrow[r] \arrow[d, "\alpha'"] & A \arrow[r] \arrow[d, "\alpha"] & A/B \arrow[r] \arrow[d, "\bar{\alpha}"] & 0 \\ E: 0 \arrow[r] & B \arrow[r] & A \arrow[r] & A/B \arrow[r] & 0 \end{tikzcd} \quad \text{and} \quad \begin{tikzcd} E: 0 \arrow[r] & B \arrow[r] \arrow[d, "\beta'"] & A \arrow[r] \arrow[d, "\beta"] & A/B \arrow[r] \arrow[d, "\bar{\beta}"] & 0 \\ E: 0 \arrow[r] & B \arrow[r] & A \arrow[r] & A/B \arrow[r] & 0 \end{tikzcd}$$

We therefore have $\alpha' E \equiv E\bar{\alpha}$ and $\beta' E \equiv E\bar{\beta}$. On the other hand, we also have a commutative diagram

$$\begin{tikzcd} E: 0 \arrow[r] & B \arrow[r] \arrow[d, "\alpha'"] & A \arrow[r] \arrow[d, "\Phi"] & A/B \arrow[r] \arrow[d, "\bar{\beta}"] & 0 \\ E: 0 \arrow[r] & B \arrow[r] & A \arrow[r] & A/B \arrow[r] & 0 \end{tikzcd}$$

which implies that $\alpha' E \equiv E\bar{\beta}$. Thus $\alpha' E \equiv \beta' E$ and so $E \in \operatorname{Ker} (\alpha' - \beta')_*$. Since $\alpha - \beta \in E_B(A)$ we may apply Lemma 3.1 to complete the first part of the theorem.

Conversely, if α, $\beta \in E_B(A)$ are such that $(\alpha - \beta) A \leq B \oplus C \leq A$, then the first part of Lemma 3.1 implies that $\alpha' E \equiv \beta' E$ and since $\beta' E \equiv E\bar{\beta}$ we have $\alpha' E \equiv E\bar{\beta}$. Hence there exists a commutative diagram (cf[1])

$$\begin{tikzcd} & E: 0 \arrow[r] & B \arrow[r] \arrow[d, "\alpha'"] & A \arrow[r] \arrow[d, "\varepsilon"] & A/B \arrow[r] \arrow[d, equal] & 0 \\ & \alpha' E: 0 \arrow[r] & B \arrow[r] \arrow[d, equal] & A_1 \arrow[r] \arrow[d, "\gamma"] & A/B \arrow[r] \arrow[d, equal] & 0 \\ & E\bar{\beta}: 0 \arrow[r] & B \arrow[r] \arrow[d, equal] & A_2 \arrow[r] \arrow[d, "\delta"] & A/B \arrow[r] \arrow[d, "\bar{\beta}"] & 0 \\ & E: 0 \arrow[r] & B \arrow[r] & A \arrow[r] & A/B \arrow[r] & 0 \end{tikzcd}$$

It is clear that $\Phi = \delta\gamma\varepsilon$ is an endomorphism of A having the desired properties. ■

LEMMA 3.3. Let $\Phi A \leq B \oplus C \leq A$ for some $\Phi \in E_B(A)$. If either Φ' is an automorphism of B or $\bar{\Phi}$ is surjective, then B is a direct summand of A.

PROOF. First assume that Φ' is an automorphism of B. We have $\Phi A+ +B = B\oplus(C\cap(\Phi A+B))$ and hence $\Phi A = \Phi B\oplus(C\cap(\Phi A+B))$. It is not difficult to see that $A = B\oplus C'$ where $C' = \{x\in A|\Phi x\in C\}$. On the other hand, if $\bar{\Phi}(A/B) = A/B$ then for all $x\in A$ we have $x = \Phi a+b$, $a\in A$, $b\in B$, and hence $x\in B\oplus C$. ■

COROLLARY 3.4. Let B be a subgroup of A and $(\alpha', \bar{\beta})\in E(B)\times E(A/B)$ where α' can be extended to an $\alpha\in E(A)$ and $\bar{\beta}$ be lifted to a $\beta\in E(A)$, $\alpha\neq\beta$. Suppose further that either

(i) $(\alpha-\beta)'$ is an automorphism of B and $f(\alpha-\beta) = (\alpha-\beta)f$ for all $f\in \operatorname{Ann} B$ or

(ii) $\overline{(\alpha-\beta)}$ is an automorphism of A/B.

Then B is a direct summand of A if and only if there exists a $\Phi\in E_B(A)$ such that $\Phi x = \alpha x$, $x\in B$ and $\Phi a-\beta a\in B$, $a\in A$.

References

[1] L. FUCHS: *Infinite abelian groups*, Volume 1. Academic Press (1970).
[2] L. FUCHS: *Infinite abelian groups*, Volume 2. Academic Press (1973).
[3] F. LOONSTRA and M. J. SCHOEMAN: On a paper by Mader (Submitted).
[4] A. MADER: On the automorphism group and endomorphism ring of an abelian group *Annales Univ. Sci. Budapest*, Sectio Mathematica, 8 (1965), 3–12.
[5] C. P. WALKER: Properties of Ext and quasi-splitting of abelian groups. *Acta Math. Acad. Sci. Hungar.*, 15 (1964), 157–160.

RELATIVES OF 3-PERMUTABILITY AND PRINCIPAL TOLERANCE TRIVIAL VARIETIES

By

IVAN CHAJDA

Přerov, Czechoslovakia

(Received September 30, 1982)

J. T. BALDWIN and J. BERMAN [1] described a close connection between definability of principal congruences and relatives of congruence permutability and 3-permutability. Some of these results and problems can be translated into the terminology of tolerance relations, [3], [5] and can be useful for characterizing so called principal tolerance trivial varieties. These varieties form a very large class of varieties with "nice" properties which can be used in applications.

1. Permutability and its relatives

By a *tolerance* on an algebra $\mathfrak{A} = (A, F)$ is meant a reflexive and symmetric binary relation on $\mathfrak{A}$ having the Substitution Property with respect to F, i.e. it is a symmetric and diagonal subalgebra of $\mathfrak{A} \times \mathfrak{A}$. It is easy to show that the set $LT(\mathfrak{A})$ of all tolerances on $\mathfrak{A}$ forms an algebraic lattice with respect to set inclusion, [2], [3], [5]. If a, b are elements of $\mathfrak{A}$, denote by $T(a, b)$ or $T_A(a, b)$ the least tolerance on $\mathfrak{A}$ containing the pair $\langle a, b\rangle$. By Con $(\mathfrak{A})$ we denote the congruence lattice of $\mathfrak{A}$ and by $\Theta(a, b)$ or $\Theta_A(a, b)$ the *principal congruence* on $\mathfrak{A}$ containing the pair $\langle a, b\rangle$.

An algebra $\mathfrak{A}$ is *tolerance trivial* if every tolerance on $\mathfrak{A}$ is a congruence; $\mathfrak{A}$ is *principal tolerance* trivial if $T(a, b) = \Theta(a, b)$ for each two elements a, b of $\mathfrak{A}$. $T(a, b)$ is called a *principal tolerance*. A variety $\mathcal{V}$ is *(principal) tolerance trivial* if each $\mathfrak{A} \in \mathcal{V}$ has this property.

The following statement is well-known (see e.g. [3], [4]):

PROPOSITION 1. For a variety $\mathcal{V}$, the following conditions are equivalent:

(1) $\mathcal{V}$ is tolerance trivial;
(2) $\mathcal{V}$ is permutable.

This result motivated our effort to characterize principal tolerance trivial varieties by relatives of permutability and use such characterizations for creating polynomial conditions.

PROPOSITION 2. Let x, y, a, b be element of an algebra $\mathfrak{A}$. Then

$$\langle x, y\rangle \in T(a, b)$$

if and only if there exist a $(2+n)$-ary polynomial p and elements $c_1, \ldots, c_n$ of $\mathfrak{A}$ such that

$$x = p(a, b, c_1, \ldots, c_n)$$
$$y = p(b, a, c_1, \ldots, c_n).$$

For the proof, see e.g. Lemma 1 in [2].

Thus, in the terminology of [1], principal tolerance trivial algebras are exactly algebras with $1-$*step principal congruences*. We can adopt another concept on [1]: $\mathfrak{A}$ has $3-$*permutable principal congruences* if

$$\Theta(a, b)\cdot\Theta(c, d)\cdot\Theta(a, b) = \Theta(c, d)\cdot\Theta(a, b)\cdot\Theta(c, d)$$

for each elements a, b, c, d of $\mathfrak{A}$, or equivalently, if

$$\Theta(a, b)\vee\Theta(c, d) = \Theta(a, b)\cdot\Theta(c, d)\cdot\Theta(a, b)$$

in Con $(\mathfrak{A})$. Now we can translate Theorem 3.7 in [1] for $n = 1$ in our terminology:

PROPOSITION 3. Let $\mathcal{V}$ be a principal tolerance trivial variety. Then for each $\mathfrak{A}\in\mathcal{V}$, each $\Theta\in$ Con $(\mathfrak{A})$ and every elements a, b of $\mathfrak{A}$,

$$\Theta\vee\Theta(a, b) = \Theta\cdot\Theta(a, b)\cdot\Theta$$

in Con $(\mathfrak{A})$, i.e. principal tolerance trivial varieties have 3-permutable principal congruences.

It is worthy to say that principal tolerance trivial varieties constitute a very large class of varieties containing among others these "nice" varieties:

(i) all permutable varieties (i.e. all varieties of groups, quasigroups, rings, modules, etc.) as follows by Proposition1;
(ii) a variety of distributive lattices, varieties of distributive p-algebras, a variety of Heyting algebras etc., see e.g. [6].

It follows that some theorems on permutable varieties remain true also for e.g. the variety of distributive lattices provided those proofs use only principal congruences. One such case will be given in the third part.

It is evident that the identity from Proposition 3 is equivalent to the inclusion

$$(*) \qquad \Theta\cdot\Theta(a, b)\cdot\Theta\supseteq\Theta(a, b)\cdot\Theta\cdot\Theta(a, b).$$

However, it is not characterizable by a Mal'cev condition. Such characterization is possible for the converse inclusion:

THEOREM 1. *For a variety $\mathcal{V}$, the following conditions are equivalent:*
(1) *For each $\mathfrak{A} \in \mathcal{V}$, each $\Theta \in \mathrm{Con}\,(\mathfrak{A})$ and each element a, b of $\mathfrak{A}$,*

$$\Theta \cdot \Theta(a, b) \cdot \Theta \subseteq \Theta(a, b) \cdot \Theta \cdot \Theta(a, b);$$

(2) *there exist 6-ary polynomials $p_1, \ldots, p_n$ such that*

$$x = p_1(x, y, z, z, x, y, z, z)$$

$$y = p_n(z, z, x, y, x, y, z, z)$$

$$p_i(z, v, x, y, x, y, z, v) = p_{i+1}(x, y, z, v, x, y, z, v) \text{ for } i = 1, \ldots, n-1.$$

PROOF. (1)⇒(2): Let $\mathfrak{A} = F_4(x, y, z, v)$ be a free algebra of $\mathcal{V}$ with four free generators x, y, z, v. Put $\Theta = \Theta(x, z) \vee \Theta(v, y)$.

Then $\langle x, y\rangle \in \Theta \cdot \Theta(z, v) \cdot \Theta$, thus, by (1),

$$\langle x, y\rangle \in \Theta(z, v) \cdot \Theta \cdot \Theta(z, v).$$

i.e. there exist elements c, d of $\mathfrak{A}$ such that

$$\langle x, c\rangle \in \Theta(z, v),\ \langle c, d\rangle \in \Theta,\ \langle d, y\rangle \in \Theta(z, v).$$

Hence, there exist 6-ary polynomials $p_1, \ldots, p_n$ such that

$$c = p_1(x, y, z, v, x, y, z, v)$$

$$p_i(z, v, x, y, x, y, z, v) = p_{i+1}(x, y, z, v, x, y, z, v) \quad \text{for } i = 1, \ldots, n-1$$

$$d = p_n(z, v, x, y, x, y, z, v)$$

as follows from $\langle c, d\rangle \in \Theta(x, z) \vee \Theta(v, y)$ in $F_4(x, y, z, v)$. However $\langle x, c\rangle \in \Theta(z, v)$ and $\langle d, y\rangle \in \Theta(z, v)$ give immediately

$$x = p_1(x, y, z, z, x, y, z, z)$$

$$y = p_n(z, z, x, y, x, y, z, z).$$

(2)⇒(1): Let $\mathfrak{A} \in \mathcal{V}$, $\Theta \in \mathrm{Con}\,(\mathfrak{A})$ and a, b be elements of $\mathfrak{A}$ with

$$\langle x, y\rangle \in \Theta \cdot \Theta(a, b) \cdot \Theta.$$

Then there exist c, d of $\mathfrak{A}$ such that

$$\langle x, c\rangle \in \Theta,\ \langle c, d\rangle \in \Theta(a, b),\ \langle d, y\rangle \in \Theta.$$

Put $r = p_1(x, y, c, d, x, y, c, d)$, $s = p_n(c, d, x, y, x, y, c, d)$.
Thus $\langle r, s\rangle \in \Theta(x, c) \vee \Theta(d, y) \subseteq \Theta$ and, by the identities of (2),

$$\langle x, r\rangle = \langle p_1(x, y, c, c, x, y, c, c), p_1(x, y, c, d, x, y, c, d)\rangle \in \Theta(a, b)$$

$$\langle s, y\rangle = \langle p_n(c, d, x, y, x, y, c, d), p_n(c, c, x, y, x, y, c, c)\rangle \in \Theta(a, b),$$

i.e. $\langle x, y\rangle \in \Theta(a, b) \cdot \Theta \cdot \Theta(a, b)$ proving (1). ∎

REMARK 1. The identity (1) of Theorem 1 is probably the best approximation of 3-permutability of principal congruences which can be characterized by a Mal'cev condition, since there exist varieties whose free algebras have 3-permutable principal congruences but the whole $\mathcal{V}$ has not this property, see e.g. [1].

REMARK 2. If we replace (1) of Theorem 1 by an analogous identity for tolerances, namely:

(1·) For each $\mathfrak{A} \in \mathcal{V}$, each $T \in LT(\mathfrak{A})$ and each a, b of $\mathfrak{A}$,

$$T \cdot T(a, b) \cdot T \subseteq T(a, b) \cdot T \cdot T(a, b),$$

then it can be proven in a way similar as in the proof of Theorem 1 (only Proposition 2 is used instead of the description of $T(a, b)$ or $T(x, z) \vee T(v, y)$ in $LT(\mathfrak{A})$, see e.g. [2], [3], [5]) that (1·) is equivalent with:

(2·) there exists a 6-ary polynomial p over $\mathcal{V}$ with

$$x = p(x, y, z, z, x, y, z, z)$$

$$y = p(z, z, x, y, x, y, z, z).$$

It is easy to show that (2·) implies the permutability:

if $\mathfrak{A} \in \mathcal{V}$, $R, S \in \mathrm{Con}\,(\mathfrak{A})$ and $\langle x, y\rangle \in R \cdot S$, then $\langle x, z\rangle \in R$ and $\langle z, y\rangle \in S$ for some $z \in \mathfrak{A}$ and (2·) implies $\langle x, p(x, z, z, y, x, y, z, z)\rangle \in S$, $\langle p(x, z, z, y, x, y, z, z), y\rangle \in R$, thus $\langle x, y\rangle \in S \cdot R$.

By Proposition 1, (1·) as well as (2·) are equivalent ot the permutability of congruences.

2. Characterizations of principal tolerance trivial varieties

Firstly, we give a characterization of principal tolerance trivial varieties in the terms of 3-permutable tolerances:

An algebra $\mathfrak{A}$ has 3-*permutable principal tolerances* if for each a, b, c, of $\mathfrak{A}$,

$$T(a, b) \cdot T(c, d) \cdot T(a, b) = T(c, d) \cdot T(a, b) \cdot T(c, d).$$

A variety $\mathcal{V}$ has 3-*permutable principal tolerances* if each $\mathfrak{A} \in \mathcal{V}$ has this property.

THEOREM 2. *For a variety $\mathcal{V}$, the following two conditions are equivalent:*

(1) *$\mathcal{V}$ is principal tolerance trivial;*

(2) *$\mathcal{V}$ has 3-permutable principal tolerances.*

PROOF. (1)⇒(2): Let $\mathfrak{A} \in \mathcal{V}$, a, b, c, d be elements of $\mathfrak{A}$ and $\langle x, y\rangle \in T(a, b) \cdot T(c, d) \cdot T(a, b)$. Then, by (1),

$$\langle x, y\rangle \in \Theta(a, b) \cdot \Theta(c, d) \cdot \Theta(a, b).$$

By Proposition 2, $\mathfrak{A}$ has 1-step principal congruences and, by the remark after Theorem 3.5 in [1],

$$h(\Theta(a, b)) = \Theta(h(a), h(b))$$

for any homomorphism h of $\mathfrak{A}$. Let h be a canonical homomorphism of $\mathfrak{A}$ onto $\mathfrak{A}/\Theta(c, d)$, thus

$$\langle h(x), h(y)\rangle \in \Theta(h(a), h(b)) \cdot \Theta(h(a), h(b)) = \Theta(h(a), h(b)),$$

i.e.

$$\langle x, y\rangle \in \Theta(c, d) \cdot \Theta(a, b) \cdot \Theta(c, d) = T(c, d) \cdot T(a, b) \cdot T(c, d),$$

proving

$$T(a, b) \cdot T(c, d) \cdot T(a, b) \subseteq T(c, d) \cdot T(a, b) \cdot T(c, d),$$

whence (2) is evident.

(2)$\Rightarrow$(1): Let $\mathfrak{A} \in \mathcal{V}$ and a, b be elements of $\mathfrak{A}$. Then $T(a, b) = \omega \cdot T(a, b) \cdot \omega = T(a, b) \cdot \omega \cdot T(a, b) = T(a, b) \cdot T(a, b)$, where $\omega = T(a, a)$, i.e. it is the identity relation on $\mathfrak{A}$. Hence $T(a, b)$ is transitive, i.e. $T(a, b) = \Theta(a, b)$. ■

The polynomial characterizations formulated in $\forall\,\exists$-conditions are used for characterizing of tolerance modularity, distributivity, decomposability of tolerances on direct products etc., see [2], [3]. Such type of condition is used also for principal tolerance triviality. For the sake of brevity, denote the sequence $z_1, \ldots, z_n$ only by $\mathbf{z}$.

THEOREM 3. *For a variety $\mathcal{V}$, the following two conditions are equivalent:*

(1) *$\mathcal{V}$ is principal tolerance trivial;*

(2) *for every $(2+n)$-ary polynomials f, g there exist $(4+n)$-ary polynomials p, q, r such that*

$$f(x, y, \mathbf{z}) = q(f(y, x, \mathbf{z}), g(x, y, \mathbf{z}), x, y, \mathbf{z})$$

$$p(x, y, x, y, \mathbf{z}) = q(g(x, y, \mathbf{z}), f(y, x, \mathbf{z}), x, y, \mathbf{z})$$

$$p(y, x, x, y, \mathbf{z}) = r(f(y, x, \mathbf{z}), g(x, y, \mathbf{z}), x, y, \mathbf{z})$$

$$g(y, x, \mathbf{z}) = r(g(x, y, \mathbf{z}), f(y, x, \mathbf{z}), x, y, \mathbf{z}).$$

PROOF. (1)$\Rightarrow$(2): Let $\mathfrak{A} = F_{2+n}(x, y, z_1, \ldots, z_n)$ be a free algbera over $\mathcal{V}$ with $2+n$ free generators $x, y, z_1, \ldots, z_n$. Let f, g be $(2+n)$-ary polynomials over $\mathcal{V}$. By Proposition 2, we have

$$\langle f(x, y, \mathbf{z}), f(y, x, \mathbf{z})\rangle \in T(x, y),$$

$$\langle g(x, y, \mathbf{z}), g(y, x, \mathbf{z})\rangle \in T(x, y),$$

thus clearly

$$\langle f(x, y, \mathbf{z}), g(y, x, \mathbf{z})\rangle \in T(x, y) \cdot T(f(y, x, \mathbf{z}), g(x, y, \mathbf{z})) \cdot T(x, y).$$

By Theorem 2, it implies

$$\langle f(x, y, \mathbf{z}), g(y, x, \mathbf{z})\rangle \in T(f(y, x, \mathbf{z}), g(x, y, \mathbf{z})) \cdot T(x, y) \cdot T(f(y, x, \mathbf{z}), g(x, y, \mathbf{z})).$$

Hence there exist c, d of $\mathfrak{A}$ such that

(i) $\langle c, d\rangle \in T(x, y)$
(ii) $\langle f(x, y, \mathbf{z}), c\rangle \in T(f(y, x, \mathbf{z}), g(x, y, \mathbf{z}))$
(iii) $\langle d, g(y, x, \mathbf{z})\rangle \in T(f(y, x, \mathbf{z}), g(x, y, \mathbf{z}))$.

By Proposition 2, (ii) implies the existence of a $(4+n)$-ary polynomial q such that

$$f(x, y, \mathbf{z}) = q(f(y, x, \mathbf{z}), g(x, y, \mathbf{z}), x, y, \mathbf{z})$$
$$c = q(g(x, y, \mathbf{z}), f(y, x, \mathbf{z}), x, y, \mathbf{z}).$$

Analogously, (iii) implies the existence of a $(4+n)$-ary r with

$$d = r(f(y, x, \mathbf{z}), g(x, y, \mathbf{z}), x, y, \mathbf{z})$$
$$g(y, x, \mathbf{z}) = r(g(x, y, \mathbf{z}), f(y, x, \mathbf{z}), x, y, \mathbf{z}).$$

Finally, (i) implies the existence of a $(4+n)$-ary polynomial p with

$$c = p(x, y, x, y, \mathbf{z})$$
$$d = p(y, x, x, y, \mathbf{z})$$

and (2) is proved.

(2)$\Rightarrow$(1): Let $\mathfrak{A} \in \mathcal{V}$, a, b, c, d, x, y be elements of $\mathfrak{A}$ and

$$\langle x, y\rangle \in T(a, b) \cdot T(c, d) \cdot T(a, b).$$

Then $\langle x, z\rangle \in T(a, b)$, $\langle z, v\rangle \in T(c, d)$, $\langle v, y\rangle \in T(a, b)$ for some z, v of $\mathfrak{A}$. By Proposition 2, there exist $(2+n)$-ary polynomials f, g and elements $e_1, \ldots, e_n$ of $\mathfrak{A}$ such that

$$x = f(a, b, \mathbf{e}), \qquad z = f(b, a, \mathbf{e})$$

and

$$v = g(a, b, \mathbf{e}), \qquad y = g(b, a, \mathbf{e}).$$

By (2), there exist $(4+n)$-ary p, q, r such that

$$x = f(a, b, \mathbf{e}) = q(f(b, a, \mathbf{e}), g(a, b, \mathbf{e}), a, b, \mathbf{e}) = q(z, v, a, b, \mathbf{e})$$
$$p(a, b, a, b, \mathbf{e}) = q(g(a, b, \mathbf{e}), f(b, a, \mathbf{e}), a, b, \mathbf{e}) = q(v, z, a, b, \mathbf{e}),$$

hence, by Proposition 2, $\langle z, v\rangle \in T(c, d)$ implies

$$\langle x, p(a, b, a, b, \mathbf{e})\rangle \in T(c, d).$$

Analogously we obtain

$$\langle p(b, a, a, b, \mathbf{e}), y\rangle \in T(c,d).$$

Clearly

$$\langle p(a, b, a, b, \mathbf{e}), p(b, a, a, b, \mathbf{e})\rangle \in T(a, b), \quad \text{thus}$$
$$\langle x, y\rangle \in T(c, d) \cdot T(a, b) \cdot T(c, d)$$

whence the 3-permutability of principal tolerances is clear. By Theorem 2, (1) is evident. ∎

3. Some applications

A variety $\mathcal{V}$ has *directly decomposable congrounce* if for all $\mathfrak{A}$, $\mathfrak{B}$ of $\mathcal{V}$ and each $\Theta \in \mathrm{Con}\,(\mathfrak{A} \times \mathfrak{B})$ there exist $\Theta_1 \in \mathrm{Con}\,(\mathfrak{A})$ and $\Theta_2 \in \mathrm{Con}\,(\mathfrak{B})$ such that $\Theta = \Theta_1 \times \Theta_2$. G. A. FRASER and A. HORN [8] gave a Mal'cev conditions characteziting such varieties. This condition is, however, rather long and complicated. Here we show how it can be simplified if $\mathcal{V}$ is assumed to be principal tolerance trivial:

THEOREM 4. *Let $\mathcal{V}$ be a principal tolerance trivial variety, then the following conditions are equivalent:*

(1) *$\mathcal{V}$ has directly decomposable congruences;*

(2) *there exist a $(2+n)$-ary polynomial p, binary polynomials $q_1, \ldots, q_n$ and ternary polynomials $r_1, \ldots, r_n$ such that*

$$x = p(x, y, q_1(x, y), \ldots, p_n(x, y))$$
$$y = p(y, x, q_1(x, y), \ldots, q_n(x, y))$$
$$z = p(x, y, r,(x, y, z), \ldots, r_n(x, y, z)) = p(y, x, r_1(x, y, z), \ldots, r_n(x, y, z)).$$

PROOF. (1)$\Rightarrow$(2) : Let $\mathfrak{A} = F_2(x, y)$, $F = F_3(x, y, z)$ be free algebras over $\mathcal{V}$ and let $\mathcal{V}$ have directly decomposable congruences. By Theorem 4 in [8], we have

$$\langle [x, z], [y, z] \rangle \in \Theta([x, x], [y, y]).$$

By (1) it means

$$\langle [x, z], [y, z] \rangle \in T([x, x], [y, y])$$

and, by Proposition 2, there exist a $(2+n)$-ary polynomial p and elements $c_1, \ldots, c_n$ of $\mathfrak{A} \times \mathfrak{B}$ such that

$$[x, z] = p([x, x], [y, y], c_1, \ldots, c_n)$$
$$[y, z] = p([y, y], [x, x], c_1, \ldots, c_n).$$

Since $c_i \in \mathfrak{A} \times \mathfrak{B}$, we have $c_i = [q_i(x, y), r_i(x, y, z)]$ for some binary or ternary polynomials q_i or r_i, respectively. If we write it componentwise, we obtain (2).

(2)$\Rightarrow$(1) : Let $\mathfrak{A}, \mathfrak{B} \in \mathcal{V}$ and a_1, a_2 be elements of $\mathfrak{A}$ and b_1, b_2, b be elements of $\mathfrak{B}$. Put $c_i = [q_i(a_1, a_2,), r_i(b_1, b_2, b)]$.
By (2),

$$\langle [a_1, b], [a_2, b] \rangle = \langle [p(a_1, a_2, \mathbf{q}(a_1, a_2)), p(b_1, b_2, \mathbf{r}(b_1, b_2, b))],$$
$$[p(a_2, a_1, \mathbf{q}(a_1, a_2)), p(b_2, b_1, \mathbf{r}(b_1, b_2, b))] \rangle =$$
$$= \langle p([a_1, b_1], [a_2, b_2], c_1, \ldots, c_n), p([a_2, b_2], [a_1, b_1], c_1, \ldots, c_n) \rangle \in$$
$$\in T([a_1, b_1], [a_2, b_2]).$$

Thus $\langle [a_1, b], [a_2, b] \rangle \in \Theta([a_1, b_1], [a_2, b_2])$ and by Theorem 4 in [8], (1) is proved. ∎

The easy way of using the condition (2) of Theorem 4 can be illustrated by these examples:

EXAMPLE 1. Let $\mathcal{V}$ be a variety of rings with unit element. Since $\mathcal{V}$ is permutable, it is principal tolerance trivial and we can take $n = 2$, $p(x_0, x_1, x_2, x_3) = x_0 \cdot x_2 + x_3$, $q(x, y) = 1$, $q_2(x, y) = 0 = r_1(x, y, z)$ and $r_2(x, y, z) = z$. Clearly

$$p(x, y, q_1, q_2) = x \cdot 1 + 0 = x$$

$$p(y, x, q_1, q_2) = y \cdot 1 + 0 = y$$

$$p(x, y, r_1, r_2) = x \cdot 0 + z = z = y \cdot 0 + z = p(y, x, r_1, r_2)$$

proving the direct decomposability of congruences.

EXAMPLE 2. Let $\mathcal{D}$ be the variety of distributive lattices. By [6], $\mathcal{D}$ is principal tolerance trivial and we can take $n = 2$,

$$p(x_0, x_1, x_2, x_3) = (x_0 \wedge x_2) \vee x_3,\ q_1(x, y) = x \vee y,\ q_2(x, y) = x \wedge y,$$

$$r_1 = x \wedge y \wedge z, \quad r_2(x, y, z) = z.$$

Then

$$p(x, y, q_1, q_2) = [x \wedge (x \vee y)] \vee [x \wedge y] = x$$

$$p(y, x, q_1, q_2) = [y \wedge (x \vee y)] \vee [x \wedge y] = y$$

$$p(x, y, r_1, r_2) = [x \wedge (x \wedge y \wedge z)] \vee z = z =$$

$$= [y \wedge (x \wedge y \wedge z)] \vee z = p(y, x, r_1, r_2).$$

The next application is in the case of Congruence Extension Property. A variety $\mathcal{V}$ satisfies the *Congruence Extension Property* (briefly CEP) if for each $\mathfrak{A} \in \mathcal{V}$ and every subalgebra $\mathfrak{B}$ of $\mathfrak{A}$ and each $\Theta \in \mathrm{Con}\,(\mathfrak{B})$ there exists $\Theta^* \in \mathrm{Con}\,(\mathfrak{A})$ such that $\Theta = \Theta^* \cap (\mathfrak{B} \times \mathfrak{B})$. Such varieties are not characterizable by Mal'cev type conditions, see [7]. In the case of principal tolerance trivial varieties, we obtain a short polynomial $\forall\ \exists$-condition:

THEOREM 5. *Let $\mathcal{V}$ be a principal tolerance trivial variety. The following conditions are equivalent:*

(1) *$\mathcal{V}$ satisfies* CEP;

(2) *for every $(2+n)$-ary polynomial p there exists a 6-ary polynomial q such that*

$$p(x, y, \mathbf{z}) = q(x, y, x, y, p(x, y, \mathbf{z}), p(y, x, \mathbf{z}))$$

$$p(y, x, \mathbf{z}) = q(y, x, x, y, p(x, y, \mathbf{z}), p(y, x, \mathbf{z})).$$

PROOF. By [7], $\mathcal{V}$ satisfies CEP if and only if for each $\mathfrak{A} \in \mathcal{V}$, every subalgebra $\mathfrak{B}$ of $\mathfrak{A}$ and each element, a, b of $\mathfrak{B}$,

$$\Theta_B(a, b) = \Theta_A(a, b) \cap (\mathfrak{B} \times \mathfrak{B}).$$

The inclusion $\subseteq$ is evidently true in any case, hence CEP is equivalent to the converse inclusion only. Since $\mathcal{V}$ is principal tolerance trivial, CEP is equivalent to

$$(**) \qquad T_B(a, b) \supseteq T_A(a, b) \supseteq (\mathfrak{B} \times \mathfrak{B}).$$

(1)$\Rightarrow$(2): Let $\mathfrak{A} = F_{2+n}(x, y, z_1, \ldots, z_n)$ be a free algebra of $\mathcal{V}$ with free generators $x, y, z_1, \ldots, z_n$ and let p be a $(2+n)$-ary polynomial over $\mathcal{V}$. Denote by

$$c = p(x, y, \mathbf{z}), \qquad d = p(y, x, \mathbf{z}).$$

Let $\mathfrak{B}$ be an algebra of $\mathcal{V}$ generated by the generators x, y, c, d. Then clearly $\mathfrak{B}$ is a subalgebra of $\mathfrak{A}$ and, by Proposition 2,

$$\langle c, d\rangle \in T_A(x, y) \cap (\mathfrak{B} \times \mathfrak{B}).$$

By $(**)$ it implies $\langle c, d\rangle \in T_B(x, y)$, thus, by Proposition 2, there exists a 6-ary polynomial q such that

$$c = q(x, y, x, y, c, d)$$
$$d = q(y, x, x, y, c, d)$$

proving (2).

(2)$\Rightarrow$(1): Suppose $\mathfrak{A}, \mathfrak{B} \in \mathcal{V}$, $\mathfrak{B}$ is a subalgebra of $\mathfrak{A}$, c, d, x, y are elements of $\mathfrak{B}$ and $\langle c, d\rangle \in T_A(x, y) \cap (\mathfrak{B} \times \mathfrak{B})$. By Proposition 2, there exist a $(2+n)$-ary polynomial p and elements $a_1, \ldots, a_n$ of $\mathfrak{A}$ such that

$$c = p(x, y, a_1, \ldots, a_n)$$
$$d = p(y, x, a_1, \ldots, a_n).$$

By (2), there exists a 6-ary polynomial q with

$$c = q(x, y, x, y, c, d)$$
$$d = q(y, x, x, y, c, d).$$

Since x, y, c, d are elements of $\mathfrak{B}$, it yields $\langle c, d\rangle \in T_B(x, y)$. ■

4. Connections with modularity

Now, we can study connections of varieties investigated in the first part and congruence modularity. A variety $\mathcal{V}$ is *modular* if Con $(\mathfrak{A})$ is modular for each $\mathfrak{A} \in \mathcal{V}$. It is well-known that the 3-permutability of congruences implies the modularity. The first theorem shows what can be said in this sense on principal tolerance trivial varieties and the second one is a strengthening of the result on 3-permutability.

THEOREM 6. *Let $\mathcal{V}$ be a principal tolerance trivial variety. Then for each $\mathfrak{A} \in \mathcal{V}$, $\Theta \in$ Con $(\mathfrak{A})$ and elements a, b, c, d of $\mathfrak{A}$, $\Theta(c, d) \subseteq \Theta$ implies $\Theta \wedge (\Theta(c, d) \vee \vee \Theta(a, b)) = \Theta(c, d) \vee (\Theta \wedge \Theta(a, b))$.*

PROOF. Suppose $\mathfrak{A}\in\mathcal{V}$, $\mathcal{V}$ is principal tolerance trivial and $\Theta(c, d)\subseteq\Theta$. Let

$$\langle x, y\rangle\in\Theta\wedge\big(\Theta(c, d)\vee\Theta(a, b)\big).$$

Then $\langle x, y\rangle\in\Theta$ and, by Proposition 3,

$$\langle x, y\rangle\in\Theta(c, d)\cdot\Theta(a, b)\cdot\Theta(c, d),$$

i.e. there exist elements p, q of $\mathfrak{A}$ with

(i) $$\langle x, p\rangle\in\Theta(c, d),\quad \langle p, q\rangle\in\Theta(a, b),\ \langle q, y\rangle\in\Theta(c, d).$$

By the assumption, (i) implies

$$\langle p, x\rangle\in\Theta,\qquad \langle y, q\rangle\in\Theta$$

which with $\langle x, y\rangle\in\Theta$ gives $\langle p, q\rangle\in\Theta$. By (i), we have $\langle p, q\rangle\in\Theta\wedge\Theta(a, b)$, thus (i) also implies

$$\langle x, y\rangle\in\Theta(c, d)\cdot\big(\Theta\wedge\Theta(a, b)\big)\cdot\Theta(c, d).$$

By Proposition 3, this yields

$$\langle x, y\rangle\in\Theta(c, d)\vee\big(\Theta\wedge\Theta(a, b)\big).$$

The converse inclusion is trivial. ∎

THEOREM 7. *Let $\mathcal{V}$ be a variety of algebras such that for each $\mathfrak{A}\in\mathcal{V}$, every $\Theta\in$ Con $(\mathfrak{A})$ and each a, b of $\mathfrak{A}$,*

$(***)$ $$\Theta\cdot\Theta(a, b)\cdot\Theta\subseteq\Theta(a, b)\cdot\Theta\cdot\Theta(a, b).$$

Then $\mathcal{V}$ is congruence modular.

PROOF. It is evident that $(***)$ of Theorem 7 implies

$$\Theta\vee\Theta(a, b) = \Theta\cdot\Theta(a, b)\cdot\Theta$$

in Con $(\mathfrak{A})$.

(A) Firstly suppose R, $T\in$ Con $(\mathfrak{A})$ and $\Theta(c, d)\subseteq R$. In a routine way, analogous to that of the proof of Theorem 6, we can easily obtain

$$R\wedge\big(\Theta(c, d)\vee T\big) = \Theta(c, d)\vee(R\wedge T).$$

(B) Now suppose the general case R, S, $T\in$ Con $(\mathfrak{A})$, $S\subseteq R$ and proceed to prove the modular identity by an induction, using (A) as an induction hypothesis. Let

$$\langle x, y\rangle\in R\wedge(S\vee T).$$

Clearly $S = \vee\{\Theta(c_\alpha, d_\alpha);\ \alpha\in I\}$ in Con $(\mathfrak{A})$, thus the previous formula implies

$$\langle x, y\rangle\in\vee\{\Theta(c_\alpha, d_\alpha);\ \alpha\in I\}\vee T.$$

By the Mal'cev lemma, there exists a finite subset, say $\{1, \ldots, n\}$ of I such that

$$\langle x, y\rangle \in \Big(\bigvee_{i=1}^{n} \Theta(c_i, d_i)\Big) \vee T.$$

Clearly $\Theta(c_i, d_i) \subseteq S$ for $i = 1, \ldots, n$ and

$$\langle x, y\rangle \in R \wedge \Big[\Big(\bigvee_{i=1}^{n} \Theta(c_i, d_i)\Big) \vee T\Big],$$

i.e.

$$\langle x, y\rangle \in R \wedge \Big[\Theta(c_1, d_1) \vee \Big(\bigvee_{i=2}^{n} \Theta(c_i, d_i) \vee T\Big)\Big].$$

Applying (A), we have

$$\langle x, y\rangle \in \Theta(c_1, d_1) \vee \Big(R \wedge \Big[\bigvee_{i=2}^{n} \Theta(c_i, d_i) \vee T\Big]\Big)$$

and after n steps of this procedure we conclude

$$\langle x, y\rangle \in \bigvee_{i=1}^{n} \Theta(c_i, d_i) \vee (R \wedge T) \subseteq S \vee (R \wedge T)$$

proving the modularity of $\mathcal{O}$. ■

References

[1] Baldwin J. T., Berman J.: Definable principal congruence relations: kith and kin, *Acta Sci. Math. (Szeged)*, **44** (1982), 255–270.

[2] Chajda I.: Distributivity and modularity of lattice of tolerance relations, *Algebra Univ.*, **12** (1981), 247–255.

[3] Chajda I.: Recent results and trends in tolerances on algebras and varieties, *Colloq. Math. Soc. J. Bolyai* 28., Finite algebra and multiple-valued logic, Szeged, 1979, North-Holland, 1981, 69–95.

[4] Chajda I.: Tolerance trivial algebras and varieties, *Acta Sci. Math, (Szeged)*, **46** (1983), 35–40.

[5] Chajda I., Zelinka B.: Lattices of tolerances, *Časop. pěst. matem.*, **102** (1977), 10–24.

[6] Chajda I., Zelinka B.: Minimal compatible tolerances on lattices, *Czech. Math, J.* **27** (1977), 452–459.

[7] Day A.: A note on the Congruence Extension Property, *Algebra Univ.*, **1** (1971), 234–235.

[8] Fraser G. A., Horn A.: Congruence relations in direct products. *Proc. Amer. Math, Soc.*, **26** (1970), 390–394.

ON SOME FIXED POINT THEOREMS AND THEIR COMPARISONS

By

M. G. DESHPANDE and G. G. HAMEDANI

Department of Mathematics, Statistics, and Computer Science
Marquette University Milwaukee

(Received July 28, 1983)

0. Introduction

While considering conditions under which a mapping T of a complete metric space X into itself has a unique fixed point, KANNAN [3] and subsequently FISHER [1], [2] have considered the first three of the six conditions listed below as $(T_1)-(T_6)$.

For any $x, y \in X$:

(T_1) $d(Tx, Ty) \leq c\{d(x, Tx)+d(y, Ty)\}$, where $0 \leq c < \frac{1}{2}$.

(T_2) $d(Tx, Ty) \leq c\{d(x, Ty)+d(y, Tx)\}$, where $0 \leq c < \frac{1}{2}$.

(T_3) $\{d(Tx, Ty)\}^2 \leq c\{d(x, Tx)d(x, Ty)+d(y, Ty)d(y, Tx)\}$, where $0 \leq c < \frac{1}{2}$.

(T_4) $\{d(Tx, Ty)\}^2 \leq c\{\{d(x, Tx)\}^2+\{d(y, Ty)\}^2\}$, where $0 \leq c < \frac{1}{2}$.

(T_5) $\{d(Tx, Ty)\}^2 \leq c\{d(x, Tx)d(y, Tx)+d(x, Ty)d(y, Ty)\}$, where $0 \leq c < \frac{1}{2}$.

(T_6) $\{d(Tx, Ty)\}^2 \leq c\{d(x, Tx)d(y, Ty)+d(x, Ty)d(y, Tx)\}$, where $0 \leq c < \frac{1}{2}$.

KANNAN has proved that (T_1) implies a unique fixed point for T which is the limit of the sequence $\{T^n x\}$ for arbitrary $x \in X$. Fisher later proved that (T_2) and (T_3) likewise imply a unique fixed point for T. The object of this note is to show (in part 1 below) that each of the conditions (T_4), (T_5), (T_6) imply the same result; and then to study (in part 2) the relation between conditions $(T_1)-(T_6)$. Specifically, we show that $(T_1) \Rightarrow (T_4)$, which implies

that our Theorem 1 is stronger than that of KANNAN; that the theorem is strictly stronger is shown by means of an example. Various examples are then given to show that (T_1), (T_2) and (T_3) are all uncomparable and also that (T_6) does not imply any of the other five conditions. In the special case of a linear mapping $Tx = ax$ on $(\mathbf{R}, d)$, where $\mathbf{R}$ is the set of all real numbers and d is the usual metric, exact intervals have been obtained in which the coefficient a must be in order that each of the conditions $(T_1)-(T_6)$ may hold. All of our counterexamples such as (T_1) no$\Rightarrow(T_2)$, (T_1) no$\Rightarrow(T_2)$, ... etc., come from this linear mapping, in some cases restricted to a smaller domain. While there are some implications for which we neither have proofs nor counterexamples, it has been demostrated that the linear mappings on $\mathbf{R}$ cannot be employed for counterxamples in these cases, and some mappings on other complete metric spaces might well be useful.

1. Fixed Point Theorems

THEOREM 1. *If T is a mapping of the complete metric space X into itself, satisfying* (T_4), *then T has a unique fixed point.*

PROOF. Let x be an arbitrary point in X. Then

$$\{d(T^n x, T^{n+1}x)\}^2 \leq c\{\{d(T^{n-1}x, T^n x)\}^2 + \{d(T^n x, T^{n+1}x)\}^2\},$$

which implies that

$$d(T^n x, T^{n+1}x) \leq \left(\frac{c}{1-c}\right)^{1/2} d(T^{n-1}x, T^n x)$$

for $n = 1, 2, \ldots$. Since $0 \leq c < \frac{1}{2}$, $\{T^n x\}$ is a Cauchy sequence in X and hence has a limit z in X.

We now have

$$\{d(T^n x, Tz)\}^2 \leq c\{\{d(T^{n-1}x, T^n x)\}^2 + \{d(z, Tz)\}^2\},$$

which implies that

$$\{d(z, Tz)\}^2 \leq c\{d(z, Tz)\}^2.$$

Since $0 \leq c < \frac{1}{2}$, it follows that $d(z, Tz) = 0$, and hence $Tz = z$, so that z is a a fixed point of T. If z' is a second fixed point of T, then

$$\{d(z, z')\}^2 = \{d(Tz, Tz')\} \leq c\{\{d(z, Tz)\}^2 + \{d(z', Tz')\}^2\} = 0, \text{ so } z = z'.$$

THEOREM 2. *If T is a mapping of the complete metric space X into itself satisfying* (T_5), *then T has a unique fixed point.*

PROOF. Let x be an arbitrary point in X. If for some positive integer n, $d(T^n x, T^{n+1}x) = 0$, then we have a fixed point; if not, then from (T_5) we obtain

$$d(T^n x, T^{n+1}x) \leq \left(\frac{c}{1-c}\right) d(T^{n-1}x, T^n x)$$

for $n = 1, 2, \ldots$. Now the rest of the proof is similar to that of Theorem 1.

THEOREM 3. *If T is a mapping of the complete metric space X into itself, satisfying* (T_6), *then T has a unique fixed point.*

PROOF. Let x be an arbitrary point in X. If for some positive integer n, $d(T^n x, T^{n+1}x) = 0$, then we have a fixed point; if not, then from (T_6) we obtain

$$d(T^n x, T^{n+1}x) \leq c\, d(T^{n-1}x, T^n x)$$

for $n = 1, 2, \ldots$. Since $0 \leq c < 1$, $\{T^n x\}$ is a Cauchy sequence in X and hence has a limit z in X, which in fact is a fixed point of T. It is easy to see that the fixed point is unique.

REMARKS. (a) If T is a continuous mapping of the compact metric space X into itself, satisfying the inequality

$$\{d(Tx, Ty)\}^2 < \{d(x, Tx)d(y, Ty) + d(x, Ty)d(y, Tx)\}$$

for all distinct x, y in X, then T has a unique fixed point.

(b) The result in (a) is similar to Theorem 4 of [2].

(c) A remark similar to (a) also holds for each of the Theorems 2 and 3.

2. Comparison of $(T_1)-(T_6)$

THEOREM 4. *Let T be the linear mapping $Tx = ax$ for $|a| < 1$ defined on* $\mathbf{R}$ *and let $d(x, y) = |x-y|$. Then T satisfies*

(i) *the condition* (T_1) *if and only if* $-1 < a < \frac{1}{3}$;

(ii) *the condition* (T_2) *if and only if* $-\frac{1}{3} < a < 1$;

(iii) *the condition* (T_3) *if and only if* $-\frac{1}{\sqrt{5}} < a < \frac{1}{\sqrt{5}}$,

however, by restricting T to a suitable subspace of $\mathbf{R}$, (T_3) *can be satisfied also for* $\frac{1}{\sqrt{5}} \leq a < \frac{1}{2}$;

(iv) *the condition* (T_4) *if and only if* $-1 < a < \frac{1}{3}$,

however, by restricting to a suitable subspace of $\mathbf{R}$, (T_4) *can be satisfied also for* $\frac{1}{3} \leq a < \frac{1}{1+\sqrt{2}}$;

(v) *the condition* (T_5) *if and only if* $-\frac{1}{3}<a<\frac{1}{3}$;

(vi) *the condition* (T_6) *if and only if* $-1<a<1$.

PROOF. We first consider (i), (ii) and (vi). The "if" parts of (i), (ii) and (vi) follow from the respective inequalities: for all $x, y \in \mathbf{R}$,

$$|a(x-y)| \leq \left|\frac{a}{1-a}\right| \{|x-ax|+|y-ay|\},$$

$$|a(x-y)| \leq \left|\frac{a}{1+a}\right| \{|x-ay|+|y-ax|\},$$

and

$$a^2(x-y)^2 \leq |a| \ \{|(x-ax)\,(y-ay)|+|(x-ay)\,(y-ax)|\}.$$

The "only if" part follows from the observation that each of the inequalities becomes equality along the lines which make the value of either of the summands on the right hand side equal to zero.

To prove (iv), we note that $(u-v)^2 \leq u^2+v^2$ if $uv \geq 0$ and $(u-v)^2 \leq \leq 2(u^2+v^2)$ if $uv<0$ with equality along $u = 0$ or $v = 0$ and along $u = -v$ respectively. We can now express condition (T_4) in the following form

$$a^2(x-y)^2 \leq \frac{2a^2}{(1-a)^2}\{(x-ax)^2+(y-ay)^2\},$$

which determines the range $-1<a<\frac{1}{3}$ for (T_4). If however, we select a subspace of $\mathbf{R}$ as $X = [0, \infty)$ and consider restriction of T to X (with $a>0$), then $(x-ax)\,(y-ay) \geq 0$ and hence condition (T_4) can be expressed as

$$a^2(x-y)^2 \leq \frac{a^2}{(1-a)^2}\{(x-ax)^2+(y-ay)^2\}.$$

Thus (T_4) also holds for $\frac{1}{3} \leq a < \frac{1}{1+\sqrt{2}}$, where T is taken on some subpaces of $\mathbf{R}$.

(iii) and (v) are essentially similar, so we will only prove (iii).

Case **1** ($0<a<1$). We observe that lines $x = 0$, $y = 0$, $x = ay$, $y = ax$ and $y = x$ partition the plane $\mathbf{R}\times\mathbf{R}$ into 10 regions as shown in figure 1.

Since condition (T_3) is symmetric in x and y, we consider the maximum value attained by the function

$$f(x, y) = \frac{\{d(Tx, Ty)\}^2}{d(x, Tx)d(x, Ty)+d(y, Tx)d(y, Ty)}$$

in each of the five regions below the line $y = x$. Of course on the line $y = x$ condition (T_3) (and all the other conditions) are trivially satisfied for any value of a.

If we let $u = (x-ax)(x-ay)$ and $v = (y-ax)(y-ay)$, then u and v are both positive in regions I, III and V; $u \geq 0$ and $v \leq 0$ in region II; $u \leq 0$ and $v \geq 0$ in region IV. Upon substituting for the various distances in $f(x, y)$ we obtain

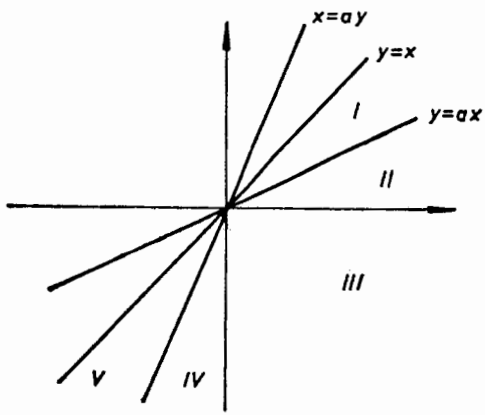

Fig. 1.

$$f(x, y) = \begin{cases} \dfrac{a^2(x-y)}{(1-a)(x+y)}, & \text{for } (x, y) \text{ in region II} \\ \dfrac{a^2(y-x)}{(1-a)(y+x)}, & \text{for } (x, y) \text{ in region IV} \\ \dfrac{a^2(x-y)^2}{(1-a)\{(x-y)^2+2(1-a)xy\}}, & \text{for } (x, y) \text{ in regions I, III and V.} \end{cases}$$

Clearly, f has its maximum along $y = 0$ in region II and along $x = 0$ in region IV. Also, by considering values of f along the lines $y = \lambda x$ in regions I and V and along circles $x^2+y^2 = \lambda$ in region III, it can be seen that f attains its maximum along $y = ax$ in region I, along $y = -x$ in region III, and along $x = ay$ in region V. Thus:

$$f(x, y) \leq \begin{cases} \dfrac{a^2}{1-a}, & \text{for } (x, y) \text{ in regions II and IV,} \\ \dfrac{a^2}{1+a}, & \text{for } (x, y) \text{ in regions I and V,} \\ \dfrac{2a^2}{1-a^2}, & \text{for } (x, y) \text{ in region III.} \end{cases}$$

Therefore, f attains its absolute maximum in region III and condition (T_3) will be satisfied if and only if $\frac{2a^2}{1-a^2} < \frac{1}{2}$ $\left(\text{i.e. } 0 < a < \frac{1}{\sqrt{5}}\right)$. However, since in the first quadrant, the absolute maximum of f is only $\frac{a^2}{1-a}$, by choosing $X = [0, \infty)$ we can require

$$\frac{a^2}{1-a} < \frac{1}{2} \left(\text{i.e. } 0 < a < \frac{1}{2}\right).$$

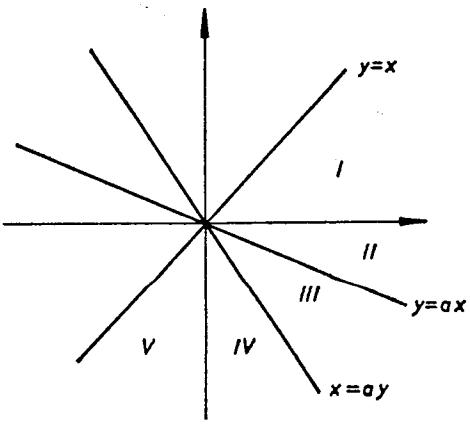

Fig. 2.

Case 2 ($-1 < a < < 0$). Referring to figure 2, we note that $u \geq 0$ and $v \geq 0$ in regions I, III and V; $u \geq 0$ and $v \leq 0$ in region II; $u \leq 0$ and $v \geq 0$ in region IV. Using the same expressions for $f(x, y)$ as before, and by similar reasoning we have

$$f(x, y) \leq \begin{cases} \dfrac{a^2}{1+a}, & \text{for } (x, y) \text{ in regions II and IV} \\ \dfrac{a^2}{1-a}, & \text{for } (x, y) \text{ in regions I and V} \\ \dfrac{2a^2}{1-a^2}, & \text{for } (x, y) \text{ in region III.} \end{cases}$$

From this, it follows that condition (T_3) is satisfied if and only if $a > -\frac{1}{\sqrt{5}}$ which completes the proof of the theorem.

Squaring both sides of the inequality in (T_1) and using the fact that $(u+v)^2 \leq 2(u^2+v^2)$ we obain (T_4). This shows that $(T_1) \Rightarrow (T_4)$. If (T_5) holds and z is a fixed point of T, then $d(z, Tx) \leq c\ d(x, Tx)$ for all x in X and consequently $d(Tx, Ty) \leq c\{d(x, Tx)+d(y, Ty)\}$ for all x and y in X. Thus $(T_5) \Rightarrow (T_1)$ and (T_4). By choosing various values of a as indicated by the above theorem, we can construct examples of linear mappings which satisfy one or more of the conditions $(T_1)-(T_6)$ and fail the others. At least for the linear mappings discussed above, it seems that $(T_5) \Rightarrow (T_i)$ $i = 2, 3, 6$ and that $(T_i) \Rightarrow (T_6)$ for $i = 1, 2, 3, 4$. We do not know whether these implications are true in general; we believe $(T_5) \Rightarrow (T_i)$, $i = 2, 3, 6$ are, but $(T_i) \Rightarrow (T_6)$, $i = 1, 2, 3, 4$ are not. Apart from these unsettled cases and $(T_1) \Rightarrow (T_4)$, $(T_5) \Rightarrow (T_1)$ (and (T_4)), every possible implication $(T_i) \Rightarrow (T_j)$ does not hold (for $i \neq j$) as can be seen by choosing different values of a. Thus, for example, $Tx = \frac{4}{5}x$ satisfies (T_6) and (T_2) but fails to satisfy (T_1), (T_3), (T_4) and (T_5). $Tx = -\frac{4}{5}x$ satisfies (T_6) and (T_1) but fails to satisfy (T_2), (T_3) and (T_5). $Tx = \frac{10}{21}x$ on $X = [0, \infty)$ satisfies (T_3) but fails (T_1) and (T_5), while $Tx = \frac{2}{5}x$ satisfies (T_4) but not (T_1).

References

[1] FISHER, B. A fixed point theorem, *Math. Mag.*, **48** (1975), 223–225.
[2] FISHER, B.: Some theorems on fixed points, *Studia Scien. Math. Hung.*, **12** (1977), 159–160.
[3] KANNAN, R.: Some results on fixed points, *Bull. Calcutta Math. Soc.*, **60** (1968), 71–76.

ON INTERPOLATION POLYNOMIALS USING THE ROOTS OF ULTRASPHERICAL POLYNOMIALS

By

S. A. N. ENEDUANYA

University of Technology, Minna, Nigeria

(Received October 31, 1983)

1. In this paper a method for approximating functions and their derivatives is introduced on the basis of Hermite–Fejér interpolation. Only the function values are needed for the use of the method and there is no need to know the values of the derivative.

The approximation problems based on the Hermite–Fejér interpolation have been discussed by several authors. First, L. FEJÉR [2] and later G. SZEGŐ [7], E. EGERVÁRY and P. TURÁN [1], G. GRÜNWALD [4], P. SZÁSZ [6], G. FREUD [3] and others have investigated this problem, but the simultaneous approximation of the function and its derivative and the rate of convergence have not been examined before.

In this paper all of these three problems will be discussed.

The interpolation points are selected as the roots of ultraspherical polynomials. The ultraspherical polynomials are defined as

$$\omega_n(x) = P_n^{(\alpha)}(x) = \frac{(-1)^n}{2^n \cdot n!}\left\{\frac{d^n}{dx^n}[(1-x^2)^{n+\alpha}]\right\}(1-x^2)^{-\alpha}, \quad n = 0, 1, 2, \ldots \tag{1.1}$$

where $\alpha > -1$.

All roots of polynomials (1.1) are single and are between -1 and $+1$, that is,

$$\Delta: -1 < x_n < x_{n-1} < \ldots < x_\nu < x_{\nu-1} < \ldots < x_1 < 1 \quad (n = 1, 2, \ldots). \tag{1.2}$$

On the basis of the nodes (1.2) and function $f(x) \in C^{(1)}([-1,1])$ define the following spline function:

If $x \in [x_\nu, x_{\nu-1}]$, $(2 \leqq \nu \leqq n+1,\ x_{n+1} = -1)$ let

$$F_\Delta(x; f) \equiv y_\nu + a_\nu(x-x_\nu) + b_\nu(x-x_\nu)^2 + c_\nu(x-x_\nu)^3 \equiv f_\nu(x) \tag{1.3}$$

and in the interval $x \in [x_1, 1 = x_0]$, let

$$F_\Delta(x; f) \equiv y_1 + \frac{y_0 - y_1}{h_1}(x - x_1) \equiv f_1(x). \tag{1.4}$$

In (1.3)

$$a_\nu = \frac{y_{\nu-1} - y_\nu}{h_\nu}; \quad b_\nu = -c_\nu h_\nu, \tag{1.5}$$

$$c_\nu = \frac{1}{h_\nu}\left\{\frac{y_{\nu-2} - y_{\nu-1}}{h_{\nu-1}} - \frac{y_{\nu-1} - y_\nu}{h_\nu}\right\}. \tag{1.6}$$

In relations (1.4), (1.5) and (1.6), $h_\nu = x_{\nu-1} - x_\nu$, $(1 \leqq \nu \leqq n+1)$ and $y_\nu = f(x_\nu)$, $(0 \leqq \nu \leqq n+1)$.

It is easy to verify that

$$F_\Delta(x_\nu; f) = y_\nu, \quad (0 \leqq \nu \leqq n+1 \tag{1.7}$$

and

$$f_\nu^{(s)}(x_{\nu-1}) = f_{\nu-1}^{(s)}(x_{\nu-1}), \quad (2 \leqq \nu \leqq n+1, s = 0, 1) \tag{1.8}$$

that is, $F_\Delta(x; f) \in C^{(1)}([-1,1])$.

Let $H_{2n-1}(x; f)$ denote the Hermite – Fejér polynomial of degree at most $(2n-1)$ which satisfies equalities

$$H_{2n-1}(x_\nu; f) = y_\nu = f(x_\nu), \quad (1 \leqq \nu \leqq n) \tag{1.9}$$

and

$$H'_{2n-1}(x_\nu; f) = y'_\nu = F'_\Delta(x_\nu; f), \quad (1 \leqq \nu \leqq n) \tag{1.10}$$

at the nodes (1.2).

It is well known that polynomials $\{H_{2n-1}(x; f)\}_{n=1}^{\infty}$ can be uniquely determined.

The following theorems will be proved in this paper.

THEOREM 1. *Let* $f(x) \in C^{(1)}([-1,1])$.

If either $x \in [-1,1]$ *and* $-\frac{1}{2} \geqq \alpha > -1$ *then*

$$|f(x) - H_{2n-1}(x; f)| = O(1)\omega\left(\frac{1}{n}; f'\right)\frac{\log n}{n}, \quad (n = 1, 2, \ldots), \tag{1.11}$$

where $\omega(\cdot\,; f')$ *denotes the modulus of continuity of* $f'(x)$.

THEOREM 2. *Let* $f(x) \in C^{(1)}([-1,1])$. *If* $x \in [-1+\varepsilon, 1-\varepsilon]$, *then*

$$|f'(x) - H'_{2n-1}(x; f)| = O(1)\omega\left(\frac{1}{n}; f'\right)\log n, \quad (n = 1, 2, \ldots) \tag{1.12}$$

where $0 < \varepsilon < \frac{1}{2}$ *and* $\omega(\cdot\,; f')$ *denotes the modulus of continuity of* $f'(x)$.

Expression (1.12) tends to zero for $\omega\left(\frac{1}{n}; f'\right)\log n = o(1)$. This condition is obviously satisfied of for example $f'(x) \in \mathrm{Lip}\ \mu$, $(0<\mu\leqq 1)$.

2. Before proving the above theorems, the following well known facts should be mentioned. Polynomials $\omega_n(x) = P_n^{(\alpha)}(x)$ satisfy the relations (Szegő [7])

$$P_n^{(\alpha)}(x) = \begin{cases} O(1-x^2)^{-\frac{1}{4}-\frac{\alpha}{2}} \cdot n^{-\frac{1}{2}}, & x\in(1,1),\ \alpha \geqq -\frac{1}{2} \\ O(1)n^{\alpha}, & x\in[-1,1],\ -1<\alpha\leqq -\frac{1}{2} \end{cases} \tag{2.1}$$

and for the roots (1.2), $x_\nu = \cos\vartheta_\nu$, $(x_\nu = x_{n-\nu+1})$,

$$\vartheta_\nu = \frac{1}{n}[\nu\pi + O(1)], \quad (1\leqq \nu \leqq n,\ n = 1, 2, \ldots) \tag{2.2}$$

and the derivatives satisfy inequalities

$$|P_n^{(\alpha)\prime}(x_\nu)| > \begin{cases} c\nu^{-\frac{3}{2}-\alpha} \cdot n^{2+\alpha}, & (0\leqq x_\nu < 1) \\ c(n-\nu+1)^{-\frac{3}{2}-\alpha} \cdot n^{2+\alpha}, & (-1<x_\nu\leqq 0), \end{cases} \tag{2.3}$$

where $O(1)$ does not depend on ν and n, and constant c depends on only α.

Polynomials $H_{2n-1}(x; f)$ satisfying (1.9) and (1.10) can be written in the form

$$H_{2n-1}(x; f) = \sum_{\nu=1}^{n} y_\nu h_\nu(x) + \sum_{\nu=1}^{n} y'_\nu (x-x_\nu) l_\nu^2(x), \tag{2.4}$$

where

$$h_\nu(x) = \left[1 - \frac{2(\alpha+1)x_\nu}{1-x_\nu^2}(x-x_\nu)\right] l_\nu^2(x) \tag{2.5}$$

and

$$l_\nu(x) = \frac{P_n^{(\alpha)}(x)}{P_n^{(\alpha)\prime}(x_\nu)(x-x_\nu)}. \tag{2.6}$$

If $r_m(x)$ is a polynomial of degree m and $m\leqq 2n-1$, then it is well known that

$$r_m(x) \equiv \sum_{\nu=1}^{n} r_m(x_\nu) h_\nu(x) + \sum_{\nu=1}^{n} r'_m(x_\nu)(x-x_\nu) l_\nu^2(x), \tag{2.7}$$

which implies that for $r_m(x) \equiv 1$,

$$\sum_{\nu=1}^{n} h_\nu(x) \equiv 1. \tag{2.8}$$

FEJÉR [2] has proved that if $-1<\alpha<0$, then for $x\in[-1,1]$

$$h_\nu(x)\geqq 0, \quad (1\leqq \nu\leqq n). \tag{2.9}$$

The "Lebesgue constants" of interpolation satisfy inequalities (SZEGŐ [7])

$$\lambda_n = \max_{x\in[-1,1]}\sum_{\nu=1}^{n}|l_\nu(x)| = O(1)\log n, \text{ if } -\frac{1}{2}\geqq\alpha>-1 \tag{2.10}$$

and for $\alpha>-\frac{1}{2}$

$$\lambda_n = \max_{x\in[-1+\varepsilon,\,1-\varepsilon]}\sum_{\nu=1}^{n}|l_\nu(x)| = O(1)\log n, \quad \left(0<\varepsilon<\frac{1}{2}\right). \tag{2.11}$$

If $f(x)\in C^{(1)}([-1,1])$ and function $F_\Delta(x;f)$ is defined by identities (1.3) and (1.4), then (ENEDUANYA [8])

$$|f^{(s)}(x)-F_\Delta^{(s)}(x;f)| = O(1)\omega\left(\frac{1}{n};f'\right)\frac{1}{n^{1-s}}, \quad (s=0,1) \tag{2.12}$$

for any $x\in[-1,1]$, where $\omega\left(\frac{1}{n};f'\right)$ denotes the modulus of continuity of functions $f'(x)$. Furthermore,

$$\omega\left(\frac{1}{n};F_\Delta'\right) = O(1)\omega\left(\frac{1}{n};f'\right). \tag{2.13}$$

In the proof the following relation will be used (JACKSON [5], SZEGŐ [7]): If $\varphi(x)\in C^{(1)}([-1,1])$ then there exists a polynomial $\varrho_m(x)$ of degree m such that for $x\in[-1,1]$,

$$|\varphi^{(s)}(x)-\varrho_m^{(s)}(x)| = O(1)\omega\left(\frac{1}{m};\varphi'\right)\frac{1}{m^{1-s}}, \quad (s=0,1), \tag{2.14}$$

where $\omega(\cdot\,;\varphi')$ denotes the modulus of continuity of function $\varphi'(x)$.

3. In this part of the paper the proofs of Theorems 1 and 2 are presented. Let $f(x)\in C^{(1)}([-1,1])$ and let $\varrho_{2n-1}(x;F_\Delta)$ be the polynomial defined for function $F_\Delta(x;f)$. The relations (2.12), (2.13), (2.14), (2.7), (2.4) and the facts $y_\nu=F_\Delta(x_\nu;f)$, $y_\nu'=F_\Delta'(x_\nu;f)$ (which are simple consequences of relations (1.7) and (1.9)) imply that

$$|f(x)-H_{2n-1}(x;f)|\leqq|f(x)-F_\Delta(x;f)|+|F_\Delta(x;f)-\varrho_{2n-1}(x;F_\Delta)|+$$

$$+\left|\sum_{\nu=1}^{n}\{\varrho_{2n-1}(x_\nu)-F_\Delta(x_\nu)\}h_\nu(x)+\sum_{\nu=1}^{n}\{\varrho_{2n-1}'(x_\nu)-F_\Delta'(x_\nu)\}(x-x_\nu)l_\nu^2(x)\right|=$$

$$= O(1)\omega\left(\frac{1}{n};f'\right)\frac{1}{n}+|V_{2n-1}(x)+W_{2n-1}(x)|. \tag{3.1}$$

If $-1<\alpha\leqq-\frac{1}{2}$, then from (2.14), (2.13), (2.9), (2.8), (2.1), (2.3), (2.6) and (2.10) imply the assertion of Theorem 1 since

$$|V_{2n-1}(x)+W_{2n-1}(x)|\leqq\max_{\nu}|\varrho_{2n-1}(x_\nu)-F_\Delta(x_\nu)|\sum_{\nu=1}^{n}h_\nu(x)+$$

$$+\max_{\nu}|\varrho'_{2n-1}(x_\nu)-F'_\Delta(x_\nu)|\sum_{\nu=1}^{n}|x-x_\nu|l_\nu^2(x)=O(1)\omega\left(\frac{1}{n};f'\right)\frac{\log n}{n}. \tag{3.2}$$

If $\alpha>-\frac{1}{2}$, then relations (2.12), (2.6), (2.1), (2.3) and (2.11) imply that for $x\in[-1+\varepsilon,\ 1-\varepsilon]$, $\left(0<\varepsilon<\frac{1}{2}\right)$,

$$|W_{2n-1}(x)|=O(1)\omega\left(\frac{1}{n};f'\right)|P_n^{(\alpha)}(x)|\sum_{\nu=1}^{n}\frac{1}{|P_n^{(\alpha)\prime}(x_\nu)|}|l_\nu(x)|=$$

$$=O(1)\omega\left(\frac{1}{n};f'\right)\frac{\log n}{n}. \tag{3.3}$$

The mean value theorem of Lagrange implies that

$$\varrho_{2n-1}(x_\nu)-F_\Delta(x_\nu)=[\varrho_{2n-1}(x)-F_\Delta(x)]+$$
$$+[\varrho'_{2n-1}(\eta_\nu)-F'_\Delta(\eta_\nu)](x_\nu-x),\quad(\eta_\nu\in(x,x_\nu))$$

then from (2.8), (2.5), (2.6) we may conclude that

$$|V_{2n-1}(x)|=\left|\varrho_{2n-1}(x)-F_\Delta(x)+\sum_{\nu=1}^{n}\{\varrho'_{2n-1}(\eta_\nu)-F'_\Delta(\eta_\nu)\}\times\right.$$

$$\left.\times\left\{-1+\frac{2(\alpha+1)x_\nu}{1-x_\nu^2}(x-x_\nu)\right\}\frac{P_n^{(\alpha)}(x)}{P_n^{(\alpha)\prime}(x_\nu)}\right| \tag{3.4}$$

and by using relations (2.14), (2.1), (2.2), (2.3) (2.11) simple calculations show that for $x\in[-1+\varepsilon,\ 1-\varepsilon]$, $0<\varepsilon<\frac{1}{2}$

$$|V_{2n-1}(x)|=O(1)\omega\left(\frac{1}{n};f'\right)\frac{\log n}{n}. \tag{3.5}$$

Relations (3.1), (3.3) and (3.5) imply the assertion of Theorem 1 even for $\alpha>-\frac{1}{2}$.

Thus Theorem 1 is proved.

The proof of Theorem 2 is based on the relation

$$|f'(x)-H'_{2n-1}(x;f)| \leqq |f'(x)-F'_d(x;f)| + |F'_d(x;f)-\varrho'_{2n-1}(x;F_d)| + \tag{3.6}$$
$$+ |V'_{2n-1}(x)+W'_{2n-1})|,$$

which are similar to relations (3.1) applied for the derivatives.

Since function $V_{2n-1}(x)+W_{2n-1}(x)$ is a polynomial, the inequality of Bernstein, and relations (3.3) and (3.5) imply that

$$|V'_{2n-1}(x)+W'_{2n-1}(x)| = O(1)\omega\left(\frac{1}{n};f'\right)\frac{\log n}{\sqrt{1-x^2}}. \tag{3.7}$$

From relations (3.6), (2.12) and (3.7) one may conclude that in interval $-1+\varepsilon \leqq x \leqq 1+\varepsilon$, $\left(0<\varepsilon<\frac{1}{2}\right)$

$$|f'(x)-H'_{2n-1}(x;f)| = O(1)\omega\left(\frac{1}{n};f'\right)\log n.$$

Thus Theorem 2 is proved.

References

[1] E. Egerváry – P. Turán, Notes on interpolation V., *Acta Math. Acad. Sci. Hung.*, **9** (1958), 259–267.

[2] L. Fejér, Lagrangesche Interpolation und die zugehörigen konjugierten Punkte, *Math. Ann.* **106** (1932), 1–55.

[3] G. Freud, On Hermite-Fejér interpolation sequences, *Acta Math. Acad. Sci. Hung.*, **23** (1972), 175–178.

[4] G. Günwald, On the theory of interpolation, *Acta Math.*, 75 (1943), 219–245.

[5] D. Jackson, The theory of approximation, *Amer. Math. Soc. Coll. Publ.*, **11** (1930).

[6] P. Szász, On quasi-Hermite-Fejér interpolation, *Acta Math. Acad. Sci. Hung.*, **10** (1959), 413–439.

[7] G. Szegő, *Orthogonal polynomials.* Amer. Math. Soc. Coll. Publ., New York, (1939).

[8] S. A. N. Eneduanya, On Hermite-Fejér interpolation polynomials using Tchebyshev abscissa, *Annales Univ. Sci. Budapest*, Sectio Mathematica, 28 (1985), 69–76.

ON THE DERIVATIVE OF INTERPOLATION POLYNOMIALS

By

SYLVANUS A. N. ENEDUANYA

University of Technology, Minna, Nigeria

(Received October 31, 1983)

In this paper we shall investigate a system of interpolation polynomials, that converges to the derivative of function.

It is a well-known fact (see FEJÉR [3]), that if f is a continuous function on $[-1,1]$ and if the Hermite–Fejér interpolation polynomial $H_{2n-1}(x; f)$ satisfies the following conditions

$$H_{2n-1}(x_\nu; f) = f(x_\nu), \quad H'_{2n-1}(x_\nu; f) = 0, \tag{1}$$

where

$$-1 = x_n < x_{n-1} < \ldots < x_\nu < x_{\nu-1} < \ldots < x_2 < x_1 = 1 \tag{2}$$

are the roots of the integrated Legendre polynomial

$$\omega_n(x) \equiv -n(n-1)\int_{-1}^{x} P_{n-1}(t)dt \equiv (1-x^2)P'_{n-1}(x) \tag{3}$$

with the normalization $P_{n-1}(1) = 1$, then converges uniformly to $f(x)$ on $[-1,1]$.

It is, however, trivially true that $H'_{2n-1}(x; f)$ does not converge to $f'(x)$, if $f'(x) \in C([-1,1])$. We shall give with the help of the polynomials $H_{2n-1}(x; f)$ a system of interpolation polynomials that converge to $f'(x)$. For the polynomial $H_{2n-1}(x; f)$ the following explicit form can be given (see [3])

$$H_{2n-1}(x; f) = \sum_{\nu=1}^{n} f(x_\nu) A_\nu(x), \tag{4}$$

where

$$A_1(x) = \left[1 - \frac{n(n-1)}{2}(x-1)\right] l_1^2(x),$$

$$A_n(x) = \left[1 + \frac{n(n-1)}{2}(x+1)\right] l_n^2(x),$$

(5) $$A_\nu(x) = l_\nu^2(x), \ (2 \leqq \nu \leqq n-1;\ n = 1, 2, \ldots)$$

and

(6) $$l_\nu(x) = \frac{\omega_n(x)}{\omega_n'(x_\nu)(x - x_\nu)}.$$

We define the numbers y_j in the following way:

(7) $$y_j \overset{\text{def}}{=} H_{2n-1}(\xi_j)\left[2 - \frac{(1-\xi_j^2)P_{n-1}'(\xi_j)^2}{2}\right]^{-1}$$

$$1 \leqq j \leqq n-1,\ n = 2, 3, \ldots),$$

where

$$1 < \xi_{n-1} < \xi_{n-2} < \ldots < \xi_j < \xi_{j-1} < \ldots < \xi_1 < 1$$

are the roots of the polynomial

(8) $$\omega_n'(x) \equiv -n(n-1)P_{n-1}(x).$$

We shall now determine the following interpolatorial polynomials $L_{n-2}(x; f')$ of degree $\leqq n-2$, such that

(9) $$L_{n-2}(\xi_j;\ f') = y_j,\ (1 \leqq j \leqq n-1,\ n = 2, 3, \ldots).$$

The polynomials that satisfy the condition (9) are given explicitly by

(10) $$L_{n-2}(x;\ f') = \sum_{j=1}^{n=1} y_j l_j^*(x),$$

where

(11) $$l_j^*(x) = \frac{\omega_n'(x)}{\omega_n''(\xi_j)(x - \xi_j)} = \frac{P_{n-1}(x)}{P_{n-1}(\xi_j)(x - \xi_j)}.$$

We prove the following

THEOREM. *If $f(x) \in C^{(1)}([-1,1]$, then we have*

$$|f'(x) - L_{n-2}(x;\ f')| =$$

(12) $$= \begin{cases} O(1)\omega\left(\dfrac{\log n}{n};\ f'\right)\log n, \text{ for } -1+\varepsilon \leqq x \leqq 1+\varepsilon \\ O(1)\omega\left(\dfrac{\log n}{n};\ f'\right)\sqrt{n}, \text{ for } -1 \leqq x \leqq 1, \end{cases}$$

where $0 < \varepsilon < \frac{1}{2}$ and $\omega(\cdot\,; f')$ denotes the modulus of continuity of $f'(x)$.

REMARK. The polynomials $L_{n-2}(x; f')$ converge uniformly to $f'(x)$ in every interval $[-1+\varepsilon, 1+\varepsilon]$ where $0<\varepsilon<\frac{1}{2}$ provided that

$$\omega\left(\frac{\log n}{n}; f'\right)\log n = o(1).$$

In the interval $[-1,1]$ we have uniform convergence if

$$\omega\left(\frac{\log n}{n}; f'\right)\sqrt{n} = o(1).$$

Now differentiating (4) we get

$$H'_{2n-1}(x; f) = \sum_{\nu=1}^{n} f(x_\nu)A'_\nu(x). \tag{13}$$

From the Lagrange mean-value theorem we obtain

$$f(x_\nu) = f(\xi_j)+f'(\xi_j)(x_\nu-\xi_j)+[f'(\eta_\nu)-f'(\xi_j)](x_\nu-\xi_j)$$

$$(\eta_\nu\in(x_\nu, \xi_j),\ 1\leqq\nu\leqq n,\ 1\leqq j\leqq n-1) \tag{14}$$

where x_ν and ξ_j are the roots of $\omega_n(x)$ and $\omega_n^*(x) = P_{n-1}(x)$ respectively.

Hence from (13) and (14) we get, if $x = \xi_j$

$$H'_{2n-1}(\xi_j; f) = f(\xi_j)\sum_{\nu=1}^{n} A'_\nu(\xi_j)+f'(\xi_j)\sum_{\nu=1}^{n}(x_\nu-\xi_j)A'_\nu(\xi_j)+$$

$$+\sum_{\nu=1}^{n}[f'(\eta_\nu)-f'(\xi_j)](x_\nu-\xi_j)\quad A'_\nu(\xi_j),\ (1\leqq j\leqq n-1). \tag{15}$$

It is a well-known fact, that

$$\sum_{\nu=1}^{n} A_\nu(x)\equiv 1;\quad \sum_{\nu=1}^{n} A'_\nu(x)\equiv 0 \tag{16}$$

and from (16), (5) and (3),

$$\sum_{\nu=1}^{n} l_\nu^2(x)\equiv 1-\frac{(1-x^2)P'_{n-1}(x)^2}{n(n-1)}. \tag{17}$$

Finally we have from (15), (5), (16) and (17) for $x = \xi_j$

$$H'_{2n-1}(\xi_j; f) = f'(\xi_j)\left[2-\frac{(1-\xi_j^2)P'_{n-1}(\xi_j)^2}{n(n-1)}\right]+2\sum_{\nu=1}^{n}[f'(\xi_\nu)-f'(\xi_j)]l_\nu^2)\xi_j)+$$

$$+\frac{n(n-1)}{2}[f'(\eta_1)-f'(\xi_j)](\xi_j-1)l_1^2(\xi_j)-$$

$$-\frac{n(n-1)}{2}[f'(\eta_n)-f'(\xi_j)](\xi_j+1)l_n^2(\xi_j) = I_1+I_2+I_3+I_4. \tag{18}$$

We shall now proceed to prove the

LEMMA. We have for $1 \leqq j \leqq n-1$ the inequality

$$|y_j - f'(\xi_j)| \leqq O(1)\omega\left(\frac{\log n}{n};\ f'\right). \tag{19}$$

PROOF. Let $\omega(\delta) = \omega(\delta; f')$ be the modulus of continuity of $f'(x)$ then for $\eta_\nu \in (x_\nu, \xi_j)$ the well-known relation

$$|f'(\eta_\nu) - f'(\xi_j)| \leqq \omega(|x_\nu - \xi_j|) \leqq \omega\left(\frac{\log n}{n}\right)\left(\frac{n}{\log n}|x_\nu - \xi_j| + 1\right) \tag{20}$$

holds.

For the polynomial $\omega_n(x)$, we have S. BERNSTEIN's inequality [2] for $n \geqq 4$ and for $-1 \leqq x \leqq 1$

$$|\omega_n(x)| \leqq \sqrt{\frac{2}{\pi}n},\ \sqrt{1-x^2}\,|P'_{n-1}(x)| = O(1)n. \tag{21}$$

For the numbers $|P_{n-1}(x_\nu)|$ we have the estimation [1]

$$|P_{n-1}(x_\nu)| = O(1)\frac{1}{\sqrt{\nu}},\ (1 \leqq \nu \leqq n). \tag{22}$$

We have for the "Lebesgue function" for $-1 \leqq x \leqq 1$ the estimate [3]

$$\lambda_n(x) = \sum_{\nu=1}^{n} |l_\nu(x)| = O(1)\log n \tag{23}$$

and from (17)

$$\sum_{\nu=1}^{n} l_\nu^2(x) \leqq 1. \tag{24}$$

Hence from (18), (17), (20), (21), (22), (23), (3) and (17) we have the inequality

$$|I_2|\left[2 - \frac{(1-\xi_j^2)P'_{n-1}(\xi_j)^2}{n(n-1)}\right]^{-1} = |I_2|\left[1 + \sum_{\nu=1}^{n} l_\nu^2(\xi_j)\right]^{-1} \leqq |I_2| =$$

$$= O(1)\omega\left(\frac{\log n}{n}\right)\left[\frac{n}{\log n}\sum_{\nu=1}^{n}\left|\frac{\omega_n(\xi_j)}{\omega'_n(x_\nu)}\right| |l_\nu(\xi_j)| + \sum_{\nu=1}^{n} l_\nu^2(\xi_j)\right] =$$

$$= O(1)\omega\left(\frac{\log n}{n}\right)\left[\sum_{\nu=1}^{n}\frac{\sqrt{\nu}}{\sqrt{n}}|l_\nu(\xi_j)| \cdot \frac{1}{\log n} + 1\right] = O(1)\omega\left(\frac{\log n}{n}\right) \tag{25}$$

and

$$|I_3|\left[1 + \sum_{\nu=1}^{n} l_\nu^2(\xi_j)\right]^{-1} \leqq |I_3| \leqq$$

$$\leqq O(1)\omega\left(\frac{\log n}{n}\right)\left[\frac{n^3}{\log n}\frac{\omega_n^2(\xi_j)}{\omega'_n(1)^2} + n^2\frac{(1-\xi_j^2)^D{}_{n-1}(\xi_j)}{\omega_n(1)^2}\right] \leqq O(1)\omega\left(\frac{\log n}{n}\right). \tag{26}$$

The same holds for

$$|I_4|\left[1+\sum_{\nu=1}^{n} l_\nu^2(\xi_j)\right]^{-1}.$$

From (18), (25), (26) and (7) we have

$$|y_j'-f'(\xi_j)| = O(1)\omega\left(\frac{\log n}{n};\ f'\right). \blacksquare$$

Now we can turn to the proof of our theorem.

Since $f'(x)$ is continuous on $[-1,1]$, a polynomial $\varrho(x)$ of degree $n-2$ exists such that [4]

$$|f'(x)-\varrho(x)| = O(1)\omega\left(\frac{1}{n};\ f'\right). \tag{27}$$

It is known that

$$\varrho(x)\equiv\sum_{j=1}^{n-1}\varrho(\xi_j)l_j^*(x)\equiv L_{n-2}(x;\ \varrho). \tag{28}$$

From (9), (27) and (19)

$$|f'(x)-L_{n-2}(x;\ f)| \leqq |f'(x)-\varrho(x)|+\sum_{j=1}^{n-1}\{|\varrho(\xi_j)-f'(\xi_j)|+$$

$$+|f'(\xi_j)-y_j|\}|l_j^*(x)| = O(1)\omega\left(\frac{1}{n};\ f'\right)+O(1)\omega\left(\frac{1}{n};\ f'\right)\sum_{j=1}^{n-1}|l_j^*(x)|+$$

$$+O(1)\omega\left(\frac{\log n}{n};\ f'\right)\sum_{j=1}^{n-1}|l_j^*(x)| \tag{29}$$

holds.

It is a well-known fact that [4]

$$\sum_{j=1}^{n-1}|l_j^*(x)| = \begin{cases} O(1)\log n, \text{ for } -1+\varepsilon\leqq x\leqq 1-\varepsilon,\ 0<\varepsilon<\dfrac{1}{2}, \\ O(1)\sqrt{n}, \text{ for } -1\leqq x\leqq 1. \end{cases} \tag{30}$$

(29) and (30) complete the proof of Theorem.

References

[1] J. Balázs – P. Turán, Notes on interpolation III, *Acta Math. Acad. Sci. Hungar.*, **9** (1958), 195–214.

[2] S. Bernstein, Sur les polynómes orthogonaux relatifs a un segment fini II., *Journ. Math. Pures et Appl.*, **10** (1931), 219–286.

[3] L. Fejér, Lagrangesche Interpolation und die zugehörigen konjugierten Punkte, *Math. Annalen*, **106** (1932), 1–55.

[4] G. Szegő. *Orthogonal polynomials*, Amer Math. Soc. Coll. Publ., 1939.

ON HERMITE–FEJÉR INTERPOLATION POLYNOMIALS USING TCHEBYSHEV ABSCISSA

By

S. A. N. ENEDUANYA

University of Technology, Minna, Nigeria

(Received October 31, 1983)

1. In this paper a method based on the Hermite–Fejér interpolation will be introduced for the simultaneous approximation of functions and their derivatives. It will be assumed that the function to be approximated is continuously differentiable. Only the functional values should be known at the interpolating points and there is no need for the knowledge of the values of the derivatives. The method to be introduced in this paper can be applied in several fields of applied and numerical mathematics.

Assume that the interpolating points are either the roots

$$\Delta: \left\{x_\nu = \cos\frac{2\nu-1}{2n}\pi\right\}_{\nu=1}^{n}, \quad (n = 1, 2, 3, \ldots) \tag{1.1}$$

of the first kind Tchebyshev polynomials $T_n(x) = \cos(n \arccos x)$ or the points

$$\left\{\xi_j = \frac{j\pi}{n}\right\}_{j=1}^{n-1}, \quad (n = 2, 3, 4, \ldots) \tag{1.2}$$

which are the roots of the second kind Tchebyshev polynomials $T_n'(x)$ of degree $n-1$, furthermore the points $\xi_n = -1$ and $\xi_0 = 1$.

Let $f(x)$ be a continuously differentiable function on the interval $[-1,1]$, that is, $f \in C^{(1)}([-1,1])$.

Let $F_\Delta(x; f)$ denote the spline functions such that for $x_\nu \leqq x \leqq x_{\nu-1}$ and $2 \leqq \nu \leqq n+1$

$$F_\Delta(x; f) \equiv y_\nu + \frac{y_{\nu-1} - y_\nu}{h_\nu}(x - x_\nu) + a_\nu(x - x_\nu)^2 + b_\nu(x - x_\nu)^3 \equiv f_\nu(x) \tag{1.3}$$

and for $x_\nu \leqq x \leqq x_0 = 1$

$$F_\Delta(x; f) \equiv y_1 + \frac{y_0 - y_1}{h_1}(x - x_1) \equiv f_1(x) \tag{1.4}$$

where $y_\nu = f(x_\nu)$, $(0 \leqq \nu \leqq n+1)$ and

$$b_\nu = \frac{1}{h_\nu^2}\left\{\frac{y_{\nu-2}-y_{\nu-1}}{h_{\nu-1}} - \frac{y_{\nu-1}-y_\nu}{h_\nu}\right\},$$

(1.5) $$a_\nu = -b_\nu h_\nu,\ 2 \leqq \nu \leqq n+1;\ h_\nu = x_{\nu-1}-x_\nu,\ 1 \leqq \nu \leqq n+1,$$

furthermore

$$x_0 = 1,\ x_{n+1} = -1 \text{ and } x_\nu = \cos\frac{2\nu-1}{2n}\pi.$$

It is easy to verify that

(1.6) $$F_\Delta(x_\nu;\ f) = y_\nu,\ (0 \leqq \nu \leqq n+1)$$

and

(1.7) $$f_\nu^{(s)}(x_{\nu-1}) = f_{\nu-1}^{(s)}(x_{\nu-1}),\ (2 \leqq \nu \leqq n+1,\ s = 0, 1)$$

that is $F_\Delta(x; f) \in C^{(1)}([-1,1])$.

Let $H_{2n-1}(x; f)$ denote the Hermite–Fejér interpolating polynomial of degree not greater than $(2n-1)$ such that for interpolating points (1.1).

$$H_{2n-1}(x_\nu, f) = y_\nu = f(x_\nu),\ (1 \leqq \nu \leqq n),$$

(1.8) $$H'_{2-1}(x_\nu; f) = F'_\Delta(x_\nu; f) = y'_\nu,\ (1 \leqq \nu \leqq n).$$

It is well known that polynomials $\{H_{2n-1}(x; f)\}_{n=1}^{\infty}$ exist and are unique. The following theorem will be first proved in this paper.

THEOREM 1. *Let $f(x) \in C^{(1)}([-1,1])$, then for all $n \geqq 9$,*

(1.9*a*) $$|f(x) - H_{n-1}(x; f)| = O(1)\omega\left(\frac{1}{n}, f'\right)\frac{\log n}{n},\ x \in [-1, 1].$$

(1.9*b*) $$|f'(x) - H'_{2n-1}(x, f)| = O(1)\omega\left(\frac{1}{n}; f'\right)\log n,\ x \in [-1+\varepsilon, 1-\varepsilon],$$

where $0 < \varepsilon < \frac{1}{2}$ and $\omega(\cdot\,; f')$ denotes the modulus of continuity of functions $f'(x)$.

Relation (1.9*b*) implies convergence only if $\omega\left(\frac{1}{n}; f'\right)\log n = o(1)$. This condition is obviously satisfied if $f'(x) \in \mathrm{Lip}\,\alpha$ $(0 < \alpha \leqq 1)$.

Let $L_n(x; f)$ be the Lagrange interpolation polynomial of degree at most n, such that for the interpolating points $\xi_0 = 1$, $\xi_n = -1$ and the points defined by (1.2),

(1.10) $$L_n(\xi_j; f) = \frac{y_j - y_{j+1}}{x_j - x_{j+1}} = y_j,\ (0 \leqq j \leqq n),$$

where $y_j = f(x_j)$, $x_0 = 1$, $x_{n+1} = -1$ and points $\{x_j\}_{j=1}^n$ are defined by (1.1), and $f(x) \in C^{(1)}([-1,1])$, The following theorem will also be proved:

THEOREM 2. *Let $f(x) \in C^{(1)}([-1,1])$ then for all $n \geqq 9$ and $x \in [-1,1]$,*

$$|f'(x) - L_n(x; f)| = O(1)\omega\left(\frac{1}{n}; f'\right) \log n, \tag{1.11}$$

where $\omega(\cdot\,; f')$ denotes the modulus of continuity of $f'(x)$.

Relation (1.11) implies convergence only if $\omega\left(\frac{1}{n}; f'\right) \log n = o(1)$.

2. Preliminaries and some Lemmas. It is well known (FEJÉR [1]) that polynomial $H_{2n-1}(x; f)$ satisfying equalities (1.8) can be written as

$$H_{2n-1}(x; f) = \sum_{\nu=1}^{n} y_\nu h_\nu(x) + \sum_{\nu=1}^{n} y'_\nu (x - x_\nu) l_\nu^2(x) \tag{2.1}$$

where

$$h_\nu(x) = \left[1 - \frac{x_\nu}{1 - x_\nu^2}(x - x_\nu)\right] l_\nu^2(x) = v_\nu(x) l_\nu^2(x) \tag{2.2}$$

and

$$l_\nu(x) = \frac{T_n(x)}{T'_n(x_\nu(x - x_\nu)}. \tag{2.3}$$

For all $x \in [-1,1]$, the following relations are satisfied (SZEGŐ [3]):

$$\sum_{\nu=1}^{n} |l_\nu(x)| = O(1) \log n, \tag{2.4}$$

$$\sum_{\nu=1}^{n} |x - x_\nu| l_\nu^2(x) = O(1) \frac{\log n}{n}, \tag{2.5}$$

$$v_\nu(x) \geqq \frac{1}{2} \quad \text{and} \quad h_\nu(x) \geqq 0. \tag{2.6}$$

If $g_m(x)$ is a polynomial of degree m and $m \leqq 2n - 1$, then it is known that

$$g_m(x) \equiv \sum_{\nu=1}^{n} g_m(x_\nu) h_\nu(x) + \sum_{\nu=1}^{n} g'_m(x_\nu)(x - x_\nu) l_\nu^2(x). \tag{2.7}$$

If $g_m(x) \equiv 1$ then relation (2.7) implies that

$$\sum_{\nu=1}^{n} h_\nu(x) \equiv 1. \tag{2.8}$$

Before proving the above theorems, the following lemmas will be verified.

LEMMA 1. If $f(x)\in C^{(1)}([-1,1])$, then for all $x\in[-1,1]$

$$|f^{(s)}_{(x)}-F_{\Delta}^{(s)}(x;f)| = O(1)\omega\left(\frac{1}{n};f'\right)\frac{1}{n^{1-s}},\quad (s=0,1) \tag{2.9}$$

where $F_{\Delta}(x;f)$ denotes function (1.3), (1.4) and $\omega(\cdot\,;f')$ is the modulus of continuity of function $f'(x)$.

PROOF. For $x_\nu\leqq x\leqq x_{\nu-1}$ and $2\leqq\nu\leqq n-1$, the mean value theorem of Lagrange and relations (1.3), (1.5) imply that

$$\begin{aligned}|f'(x)-F'_{\Delta}(x;f)| &= \left|f'(x)-\frac{y_{\nu-1}-y_\nu}{h_\nu}+b_\nu[2h_\nu(x-x_\nu)-3(x-x_\nu)^2\right| \leqq \\ &\leqq |f'(x)-f'(\eta_\nu)|+\frac{1}{h_\nu^2}|f'(\eta_{\nu-1})-f'(\eta_\nu)|\frac{h_\nu^2}{3}\leqq\omega(h_\nu;f')+\frac{1}{3}\omega(h_\nu+h_{\nu-1};f')\end{aligned} \tag{2.10}$$

since $\eta_\nu\in(x_\nu,x_{\nu-1})$ and $\eta_{\nu-1}\in(x_{\nu-1},x_{\nu-2})$. Since $f(x_\nu)=F_{\Delta}(x_\nu;f)$, relation (2.10) implies that

$$\begin{aligned}|f(x)-F_{\Delta}(x;f)| &\leqq \int_{x_\nu}^{x}|f'(t)-F'_{\Delta}(t;f)|dt\leqq \\ &\leqq\left[\omega(h_\nu;f')+\frac{1}{3}\omega(h_\nu+h_{\nu-1};f')\right]h_\nu.\end{aligned} \tag{2.11}$$

The well known inequality $\omega(\delta\lambda)\leqq\omega(\delta)(\lambda+1)$ and inequalities (2.10), (2.11) and $h_\nu=O\left(\frac{1}{n}\right)$ imply the assertion for all $x\in[-1,1]$. The case when $x\in[x_1,1]$ can be verified by a similar way. Thus Lemma 1 is proved.

Polynomials $L_n(x;f)$ satisfying equations (1.10) can be written as

$$L_n(x;f)=y_0'\frac{1+x}{2}\cdot\frac{T_n'(x)}{T_n'(1)}+\frac{1-x}{2}\frac{T_n'(x)}{T_n'(-1)}+\sum_{\nu=1}^{n}y_\nu'\frac{1-x^2}{1-x_\nu^2}l_\nu^*(x) \tag{2.12}$$

where

$$l_j^*(x)=\frac{T_n'(x)}{T_n''(\xi_j)(x-\xi_j)}. \tag{2.13}$$

LEMMA 2. If $x\in[-1,1]$, then

$$\begin{aligned}\lambda_n(x) &= \frac{1}{2}(1+x)\left|\frac{T_n'(x)}{T_n'(1)}\right|+\frac{1}{2}(1-x)\left|\frac{T_n'(x)}{T_n'(-1)}\right|+ \\ &+\sum_{j=1}^{n}\frac{1-x^2}{1-\xi_j^2}|l_j^*(x)| = O(1)\log n.\end{aligned} \tag{2.14}$$

PROOF. It is well known that polynomials $T_n(x)$ satisfy the differential equations

$$(1-x^2)T_n''(x)-xT_n'(x)+n^2T_n(x)=0. \quad (n=0,1,2,\ldots). \tag{2.15}$$

Since $|T_n(1)|=|T_n(-1)=|T_n(\xi_j)|=1$ and for $x\in[-1,1]$, $|T_n(x)|\leqq 1$, relation (2.15) implies that

$$(1-\xi_j^2)|T_n''(\xi_j)|=n^2;\ |T_n'(1)|=|T_n'(-1)|=n^2. \tag{2.16}$$

furthermore

$$|T_n'(x)|=\frac{n}{\sqrt{1-x^2}},\ x\in(-1,1). \tag{2.17}$$

The inequality of Markov implies that

$$|T_n'(x)|\leqq n^2,\ x\in[-1,1]. \tag{2.18}$$

From (2.16) and (2.18) are may conclude that

$$\frac{1+x}{2}\left|\frac{T_n'(x)}{T_n'(1)}\right|\leqq 1;\ \frac{1-x}{2}\left|\frac{T_n'(x)}{T_n'(-1)}\right|\leqq 1,\ x\in[-1,1]. \tag{2.19}$$

Since (2.13) and (2.14) hold,

$$\lambda_n(1)=\lambda_n(-1)=\lambda_n(\xi_j)=1,\ (1\leqq j\leqq n-1). \tag{2.20}$$

If $x\in(\xi_{k+1},\xi_k)$, $(1\leqq k\leqq n-2)$, then the mean value theorem of Lagrange and relations (1.2), (2.13), (2.15), (2.16) imply that

$$\frac{1-x^2}{1-\xi_k^2}|l_k^*(x)|=\frac{1-x^2}{(1-\xi_k^2)|T_n''(\xi_k)|}\left|\frac{T_n'(x)-T_n'(\xi_k)}{x-\xi_k}\right|=$$

$$=\frac{1-x^2}{n^2}|T_n''(\eta_k)|=\frac{1-x^2}{n^2(1-\eta_k^2)}(1-\eta_k^2)|T_n''(\eta_k)|=O(1), \tag{2.21}$$

and similarly for $x\in(\xi_{k+1},\xi_k)$, $(0\leqq k\leqq n-2)$

$$\frac{1-x^2}{1-\xi_{k+1}^2}|l_{k+1}^*(x)|=O(1). \tag{2.22}$$

If $x\in(\xi_{k+1},\xi_k)$, $(1\leqq k\leqq n-2)$, then relations (2.13), (2.16), (2.17) and (1.2) imply that

$$\sum_{j=1}^{k-1}\frac{1-x^2}{1-\xi_j^2}|l_j^*(x)|=O(1)\sum_{j=1}^{k-1}\frac{k}{n^2(\xi_j-\xi_k)}=O(1)\sum_{j=1}^{k-1}\frac{1}{k-j}=O(1)\log n \tag{2.23}$$

and similarly

$$\sum_{j=k+2}^{n}\frac{1-x^2}{1-\xi_j^2}|l_j^*(x)|=O(1)\log n. \tag{2.24}$$

If $x\in(\xi_1, 1)$, then the mean value theorem of Lagrange, Markov's inequality and relation (2.13), (2.16), (1.2) and (2.18) imply that

$$\frac{1-x^2}{1-\xi_1^2}\,|l_1^*(x)| \leqq \frac{1-x_1^2}{n^2}\,|T_n''(\eta_1)| = O(1),\ \eta_1\in(\xi_1, 1). \tag{2.25}$$

From (3.13), (2.16), (2.17) and (1.2) one may conclude that

$$\sum_{j=2}^{n-1}\frac{1-x^2}{1-\xi_j^2}\,|l_j^*(x)| \leqq \sum_{j=2}^{n-1}\frac{\sqrt{1-x_1^2}}{n(x_1-x_j)} = O(1). \tag{2.26}$$

Similarly to the relations (2.25) and (2.26) it can be proved that for $x\in(-1, \xi_{n-1})$

$$\frac{1-x^2}{1-\xi_{n-1}^2}\,|l_{n-1}^*(x)| = O(1);\quad \sum_{j=1}^{n-2}\frac{1-x^2}{1-\xi_j^2}\,|l_j^*(x)| = O(1). \tag{2.27}$$

And finally, relations (2.20)–(2.27) imply the assertion.

The proofs of the theorems will be based on Gopengaus' inequality [2], which can be stated as follows. If $\varphi(x)\in C^{(1)}([-1,1])$, then there exists a polynomial $G_m(x;\varphi)$ of degree m, $m\geqq 9$, such that for $x\in[-1,1]$,

$$|\varphi^{(s)}(x)-G_m^{(s)}(x;\varphi)| = O(1)\omega\left(\frac{\sqrt{1-x^2}}{m};\varphi'\right)\left(\frac{\sqrt{1-x^2}}{m}\right)^{1-s},\ (s=0,\ 1), \tag{2.28}$$

where $\omega(\cdot\,;\varphi')$ denotes the modulus of continuity of function $\varphi'(x)$.

3. The proofs of the theorems. Assume that $f(x)\in C^{(1)}([-1,1])$. Then relations (2.9), (2.28), (2.7), (1.6), (1.8) imply that for $x\in[-1,1]$,

$$\begin{aligned}|f(x)-H_{2n-1}(x;f)| &\leqq |f(x)-F_\Delta(x;f)|+|F_\Delta(x;f)-G_{2n-1}(x;F_\Delta)|+\\ &+\left|\sum_{\nu=1}^{n}\{G_{2n-1}(x_\nu;F_\Delta)-F_\Delta(x_\nu;f)\}h_\nu(x)+\right.\\ &\left.+\sum_{\nu=1}^{n}\{G_{2n-1}'(x_\nu;F_\Delta)-F_\Delta'(x_\nu;f)\}(x-x_\nu)l_\nu^2(x)\right| = O(1)\omega\left(\frac{1}{n};f'\right)+\\ &+O(1)\omega\left(\frac{\sqrt{1-x^2}}{n};F_\Delta'\right)\frac{\sqrt{1-x^2}}{n}+|V_{2n-1}(x)|.\end{aligned} \tag{3.1}$$

Since $\omega(\delta_1)\leqq\omega(\delta_2)$ for $\delta_1\leqq\delta_2$, and relations (2.28), (2.6), (2.8) and (2.5) are true, one may conclude that for all $x\in[-1,1]$,

$$|V_{2n-1}(x)| = O(1)\omega\left(\frac{1}{n};F_\Delta'\right)\frac{1}{n}+O(1)\left(\frac{1}{n};F_\Delta'\right)\frac{\log n}{n}, \tag{3.2}$$

where $\omega(\cdot\,;F_\Delta')$ denotes the modulus of continuity of function $F_\Delta'(x;f)$.

If $\omega(\cdot\,;F_\Delta')$ and $\omega(\cdot\,;f')$ denote the moduli of continuity of functions $F_\Delta'(x;f)$ and $f'(x)$ respectively, then

$$\omega\left(\frac{1}{n};\ F_\Delta'\right) = O(1)\omega\left(\frac{1}{n};\ f'\right). \tag{3.3}$$

In order to prove this inequality observe that $x', x''\in[-1,1]$ and $|x'-x''|\leqq \leqq\frac{1}{n}$, then relation (2.9) and the fact that

$$\max_{|x'-x''|\leqq\frac{1}{n}} |f'(x')-f'(x'')| = \omega\left(\frac{1}{n};\ f'\right)$$

imply that

$$|F_\Delta'(x';\ f)-F_\Delta'(x'';\ f)| \leqq |F_\Delta'(x';\ f)-f'(x')| + |f'(x')-f'(x'')| + \\ + |f'(x'')-F_\Delta'(x'',\ f) = O(1)\omega\left(\frac{1}{n};\ f'\right)$$

and since

$$\max_{|x'-x''|\leqq\frac{1}{n}} |F_\Delta'(x';f)-F_\Delta'(x'';\ f)| = \omega\left(\frac{1}{n};\ F_\Delta'\right)$$

inequality (3.3) is necessarily true.

Equality (1.9*a*) is a consequence of relations (3.1), (3.2) and (3.3).

By the use of Bernstein's inequality for polynomial $V_{2n-1}(x)$ and applying relations (3.2) and (3.3) we get the following inequality:

$$|V_{2n-1}'(x)| = O(1)\omega\left(\frac{1}{n};\ f'\right)\frac{\log n}{\sqrt{1-x^2}},\quad x\in(-1,\ 1). \tag{3.4}$$

By using relations (2.9), (2.28), (3.3) and (3.4) inequalities similar to (3.1) can be verified for the derivatives, that is, for $x\in[-1+\varepsilon,\ 1-\varepsilon]$, $\left(0<\varepsilon<\frac{1}{2}\right)$

$$|f'(x)-H_{2n-1}'(x;\ f)| = O(1)\omega\left(\frac{1}{n}:\ f'\right)\log n \tag{3.5}$$

which is equivalent to inequality (1.9*b*). Thus, Theorem 1 is completely proven.

Next, Theorem 2 will be verified.

If $f(x)\in C^{(1)}([-1,1])$ and $x\in[-1,1]$, then it is easy to show that

$$|f'(x)-L_n^*(x;\ f')| = O(1)\omega\left(\frac{1}{n};\ f'\right)\log n, \tag{3.6}$$

where function

$$L_n^*(x; f') = f'(1)\frac{(1+x)T_n'(x)}{2T_n'(1)} + f'(-1)\frac{(1-x)T_n'(x)}{2T_n'(-1)} +$$

$$+\sum_{j=1}^{n-1} f'(\xi_j)\frac{1-x^2}{1-\xi_j^2} l_j^*(x) \tag{3.7}$$

is a Lagrange interpolation polynomial with nodes defined by (1.2). Let $G_n(x; f')$ denote the polynomial satisfying inequality (2.28), then relations (2.28) and (2.14) imply that for $x \in [-1,1]$,

$$|f'(x) - L_n^*(x; f')| \leq |f'(x) - G_n(x; f')| + |L_n^*(x; G_n - f')| =$$

$$= O(1)\omega\left(\frac{1}{n}; f'\right)\log n,$$

which is equivalent to inequality (3.6). The mean value theorem of Lagrange and relations (3.6), (1,10) and (2.14) imply that

$$|f'(x) - L_n(x; f)| \leq |f'(x) - L_n^*(x; f')| +$$

$$\max_{0 \leq j \leq n}\left|f'(\xi_j) - \frac{y_j - y_{j+1}}{x_j - x_{j+1}}\right| \lambda_n(x) = O(1)\omega\left(\frac{1}{n}; f'\right)\log n +$$

$$+ \max_{0 \leq j \leq n} |f(\xi_j) - f'(\eta_j)|\lambda_n(x) = O(1)\omega\left(\frac{1}{n}; f'\right)\log n \tag{3.8}$$

since $\xi_j, \eta_j \in (x_{j+1}, x_j)$, consequently relation (1.1) implies that

$$|f'(\xi_j) - f'(\eta_j)| = O(1)\omega\left(\frac{1}{n}; f'\right).$$

Thus, Theorem 2 is proved.

References

[1] L. Fejér, Lagrangesche Interpolation und die zugehörigen konjugierte Punkte, *Math, Ann.*, **106** (1932), 1–55.

[2] И. Э. ГОПЕНГАУЗ, К теореме А. Ф. Тимана о приближении функции многочленами на конечном отрезке, *Мат. Заметки*, **1** (1967), 163–172.

[3] G. Szegö, *Orthogonal polynomials*, Amer. Math. Soc. Coll. Publ., New York, 1939.

ON THE CONVERGENCE OF SPECIAL HERMITE–FEJÉR INTERPOLATION POLYNOMIALS

By

SYLVANUS A. N. ENEDUANYA

University of Technology, Minna, Nigeria

(Received October 31, 1983)

In this paper we define a Hermite–Fejér type interpolation process which approximates in "Jackson order". This interpolation process is interesting for the numerical analysis, too.

Let us denote by

$$\Delta:\{x_\nu^{(n)}\}_{\nu=1}^n = \left\{\cos\frac{(2\nu-1)}{2n}\pi\right\}_{\nu=1}^n \tag{1}$$

the roots of the Tchebyshev polynomials

$$T_n(x) = \cos(n \arccos x), \quad (n = 0, 1, 2, \ldots). \tag{2}$$

To every real function $f(x)$ defined on the closed interval $[-1,1]$ we can order the system of the generalized Hermite–Fejér interpolation polynomials $\{H_{2n-1}(x; f)\}_{n=1}^\infty$ where for every index n

$$H_{2n-1}(x; f) = H_{2n-1}(x)$$

is a polynomial of degree $\leqq (2n-1)$ satisfying the conditions

$$\begin{aligned} H_{2n-1}(x_\nu) &= f(x_\nu) = y_\nu, \\ H'_{2n-1}(x_\nu) &= y'_{\nu n} = y'_\nu \end{aligned} \quad (1 \leqq \nu \leqq n;\ n = 1, 2, \ldots) \tag{3}$$

$$\{y_\nu\}_{\nu=1}^n \text{ and } \{y'_\nu\}_{\nu=1}^n$$

may be any preassigned system of real numbers.

L. FEJÉR [1] has proved that if f(x) is continuous on $[-1,1]$ and for every pair of indices (ν, n) $y'_{\nu n} = 0$ then $H_{2n-1}(x; f)$ tends uniformly to $f(x)$ on $[-1,1]$ as $n \to \infty$.

Later FEJÉR [2] improved on this result by proving that if $f(x)$ is continuous on $[-1,1]$, and the preassigned values of the derivatives satisfy the condition

$$y'_\nu = y'_{\nu n} = H'_{2n-1}(x_\nu) = o\left(\frac{n}{\sqrt{1-x_\nu^2}\,(\log n)}\right) \tag{4}$$

then the relation $\lim\limits_{n\to\infty} H_{2n-1}(x;f) = f(x)$ holds also uniformly on $[-1,1]$.

The aim of the present paper is to investigate the *special Hermite–Fejér interpolation sequence*

$$\{\overline{H}_{2n-1}(x;f)\}_{n=1}^{\infty} \tag{5}$$

for abscissas (1), where $\overline{H}_{2n-1}(x;f)$ satisfies the following equalities if $f(x)$ is a continuous function in $[-1,1]$

$$\overline{H}_{2n-1}(x_\nu;f) = \overline{H}_{2n-1}(x_\nu) = f(x_\nu) = y_\nu, \quad (1 \leqq \nu \leqq n,\ n = 1, 2, \ldots) \tag{6}$$

and

$$\overline{H}'_{2n-1}(x_\nu;f) = \overline{H}'_{2n-1}(x_\nu) = \pi'_{2n-1}(x_\nu;F_\Delta) = y'_\nu \quad (1 \leqq \nu \leqq n,\ n = 1, 2, \ldots). \tag{7}$$

In (7) $F_\Delta(x) = F_\Delta(x;f)$ is defined as follows:

$$F_\Delta(x;f) \equiv \begin{cases} y_2 + \dfrac{y_1 - y_2}{h_2}(x - x_2) = f_1(x), & \text{for } x_1 \leqq x \leqq 1, \\ f_\nu(x) & \text{for } x_\nu \leqq x \leqq x_{\nu-1},\ 2 \leqq \nu \leqq n \\ y_n - \dfrac{y_{n-1} - y_n}{h_n} h_{n+1} + \dfrac{y_{n-1} - y_n}{h_n}(x+1) = f_{n+1}(x), & \\ & \text{for } -1 \leqq x \leqq x_n \end{cases} \tag{8}$$

where

$$f_2(x) = y_2 + \frac{y_1 - y_2}{h_2}(x - x_2) = f_1(x) \tag{9}$$

and

$$f_\nu(x) = y_\nu + \frac{y_{\nu-1} - y_\nu}{h_\nu}(x - x_\nu) + a_\nu(x - x_\nu)^2 + b_\nu(x - x_\nu)^3, \quad (3 \leqq \nu \leqq n) \tag{10}$$

furthermore

$$h_\nu = x_{\nu-1} - x_\nu > 0, \quad (2 \leqq \nu \leqq n), \quad h_{n+1} = x_n + 1 \tag{11}$$

and

$$a_\nu = -\frac{1}{h_\nu}\left(\frac{y_{\nu-2}-y_{\nu-1}}{h_{\nu-1}} - \frac{y_{\nu-1}-y_\nu}{h_\nu}\right), \quad (3 \leqq \nu \leqq n)$$

$$b_\nu = -\frac{a_\nu}{h_\nu}, \quad (3 \leqq \nu \leqq n). \tag{12}$$

Taking into consideration (9), (10), (11) and (12) we have from the definition (8) of $F_\Delta(x; f)$

$$F_\Delta(x_\nu; f) = f(x_\nu) = y_\nu \quad (1 \leqq \nu \leqq n), \tag{13}$$

$$f_\nu(x_{\nu-1}) = f_{\nu-1}(x_{\nu-1}) = y_{\nu-1}, \quad (2 \leqq \nu \leqq n+1) \tag{14}$$

and clearly

$$f'_\nu(x_{\nu-1}) = f'_{\nu-1}(x_{\nu-1}), \quad (2 \leqq \nu \leqq n+1). \tag{15}$$

It follows from (8)–(15) that

$$F_\Delta(x; f) \in C^{(1)}([-1,1]) \tag{16}$$

that is F_Δ is continuously differentiable in $[-1,1]$.

If $x = \cos\vartheta$ then $\varphi(\vartheta) \overset{\text{def}}{=} F_\Delta(\cos\vartheta)$ is an even function, so its Jackson mean $y_n(\vartheta, \varphi)$ [3] is a pure cosine polynomial of degree $2n-2$ and therefore obviously

$$y_n(\arccos x; \varphi) = \pi_{2n-1}(x; F_\Delta) \tag{17}$$

is an algebraic polynomial of degree $2n-2$, where

$$\pi_{2n-2}(x; F_\Delta) \equiv \frac{3}{\pi n(2n^2+1)} \int_0^{\frac{\pi}{2}} \{\varphi(u)_{u=\arccos x-2t} + \varphi(u)_{u=\arccos x+2t}\}\left(\frac{\sin nt}{\sin t}\right)^4 dt. \tag{18}$$

Since $F_\Delta \in C^{(1)}([-1,1]($ and $\varphi(u) = F_\Delta(\cos u)$, differentiating π_{2n-2} with respect to x and substituting x by x_ν we get

$$\pi'_{2n-2}(x_\nu; F_\Delta) \equiv -\frac{3}{\pi n(2n^2+1)} \frac{1}{\sqrt{1-x_\nu^2}} \times$$

$$\times \int_0^{\frac{\pi}{2}} \left\{\left(\frac{d\varphi}{du}\right)_{u=\arccos x_\nu-2t} + \left(\frac{d\varphi}{du}\right)_{u=\arccos x_\nu-2t}\right\}\left(\frac{\sin nt}{\sin t}\right)^4 dt. \tag{19}$$

The expression (19) figures in the equality (7).

We remark that the sequence of interpolation polynomials (5) satisfying the equalities (6) and (7) is really a *Hermite–Fejér type one.* Obviously, to construct the polynomial $\overline{H}_{2n-1}(x; f)$ of degree $2n-1$ at most we need the *discrete values* $y_\nu = f(x_\nu)$, $(\nu = 1, 2, \ldots, n, n = 1, 2, \ldots)$ *only.*

We prove the following theorem:

THEOREM. *If* $f(x) \in C([-1,1])$ *then the inequality*

$$|f(x) - \overline{H}_{2n-1}(x; f)| \leqq O(1)\omega\left(\frac{1}{n}; f\right), \quad (1 \leqq x \leqq 1) \tag{20}$$

is valid, where $\omega(\cdot\,; f)$ *denotes the modulus of continuity of* $f(x)$.

The estimation (20) shows that in $[-1,1]$ the rate of convergence of the sequence of interpolation polynomials $\overline{H}_{2n-1}(x; f)$ to the function $f(x)$ is the "Jackson order".

To prove our theorem we need the following lemmas:

LEMMA 1. If $f(x) \in C([-1,1])$ then the inequality

$$|f(x) - F_\Delta(x; f)| \leqq O(1)\omega\left(\frac{1}{n}; f\right), \quad (-1 \leqq x \leqq 1) \tag{21}$$

holds, where $\omega(\cdot\,; f)$ is the modulus of continuity of $f(x)$.

PROOF. First we consider the case $x \in [x_n, x_2]$. If $x \in [x_\nu, x_{\nu-1}]$, $\nu = 3, 4, \ldots$ $\ldots, n$; then we have by (8), (10), (11), (12) $(y_\nu = f_\nu(x_\nu),\ \nu = 1, 2, \ldots, n)$ and $a_\nu + b_\nu h_\nu = 0$ the estimation

$$\begin{aligned} |f(x) - F_\Delta(x; f)| &\equiv |f(x) - f_\nu(x)| \leqq |f(x_\nu) - y_\nu| + |y_{\nu-1} - y_\nu| + \\ &+ \max_{x_\nu \leqq x \leqq x_{\nu-1}} |a_\nu(x - x_\nu)^2 + b_\nu(x - x_\nu)^3| \leqq 2\omega(h_\nu; f) + \\ &+ \frac{2}{27}\left[\omega(h_\nu; f) + \frac{h_\nu}{h_{\nu-1}}\omega(h_{\nu-1}; f)\right]. \end{aligned} \tag{22}$$

It is easy to see form (1) and (11) that

$$h_\nu = x_{\nu-1} - x_\nu \leqq \begin{cases} \pi^2(\nu - 1)n^{-2} & \text{for } 2 \leqq \nu \leqq \left[\dfrac{n}{2}\right], \\ \pi^2(n - \nu + 1)n^{-2} & \text{for } \left[\dfrac{n}{2}\right] \leqq \nu \leqq n, \end{cases} \tag{23}$$

$$h_{\nu-1} = x_{\nu-2} - x_{\nu-1} \geqq \begin{cases} 4(\nu - 2)n^{-2} & \text{for } 3 \leqq \nu \leqq \left[\dfrac{n}{2}\right], \\ 4(n - \nu + 2)n^{-2} & \text{for } \left[\dfrac{n}{2}\right] + 1 \leqq \nu \leqq n, \end{cases}$$

$$\frac{h_\nu}{h_{\nu-1}} \leqq \frac{\pi^2}{2}, \quad (3 \leqq \nu \leqq n).$$

Using (22), (23), the monotonicity of $\omega(\cdot\,; f)$ and the well-known relation $\omega(\delta\lambda)\leqq\omega(\delta; f)(\lambda+1)$ we obtain

$$|f(x)-F_\Delta(x;\, f)|\leqq O(1)\omega\left(\frac{1}{n};\, f\right) \tag{24}$$

if $2\leqq\nu\leqq\left[\frac{n}{2}\right]$. A similar inequality is true, if $\left[\frac{n}{2}\right]+1\leqq\nu\leqq n$.

In the case $x\in[x_2, x_1]$ we get by (8), (9) and (23) inequality

$$|f(x)-F_\Delta(x; f)|\leqq|f(x)-y_2|+|y_1-y_2|\leqq O(1)\omega\left(\frac{1}{n}\right). \tag{25}$$

Finally, if $x\in[x_1, 1]$ or $x\in[-1, x_n]$ then from (8), (1), (23) and the relation $1-x_1 = 1+x_n$ follows

$$|f(x)-F_\Delta(x;\, f)|\leqq O(1)\omega\left(\frac{1}{n};\, f\right). \tag{26}$$

(24), (25) and (26) give the proof of Lemma 1.

LEMMA 2. If $\omega(\cdot\,; f)$ and $\omega(\cdot\,;\, F_\Delta)$ denote the moduli of continuity in $[-1,1]$ of $f(x)$ and $F_\Delta(x; f)$ respectively, then the inequality

$$\omega\left(\frac{1}{n};\, F_\Delta\right)\leqq O(1)\omega\left(\frac{1}{n};\, f\right) \tag{27}$$

holds.

PROOF. If $x', x''\in[-1,1]$ and $|x'-x''|\leqq\frac{1}{n}$, then we have by (21) the relation

$$|F_\Delta(x;\, f)-F_\Delta(x'';\, f)|\leqq|F_\Delta(x';\, f)-f(x')|+|f(x')-f(x'')|+$$
$$+|f(x'')-F_\Delta(x''; f)|\leqq O\,(1)\omega\left(\frac{1}{n};\, f\right).$$

Since

$$\omega\left(\frac{1}{n};\, F_\Delta\right)=\max_{|x'-x''|\leqq\frac{1}{n}}|F_\Delta(x'; f)-F_\Delta(x''; f)|$$

Lemma 2 is proved.

If $\pi_{2n-2}(x;\, F_\Delta)$ denotes the polynomial given in (18), then by the theorem of JACKSON [3] and the inequality (27) follows the estimation;

$$|F_\Delta(x;\, f)-\pi_{2n-2}(x;\, F_\Delta)|\leqq 6\omega\left(\frac{1}{n};\, F_\Delta\right)\leqq O(1)\omega\left(\frac{1}{n};\, f\right),\quad -1\leqq x\leqq 1. \tag{28}$$

Since $\pi_{2n-2}(x;\ F_\Delta)$ is a polynomial of degree $(2n-2)$ at most, therefore the following identity is satisfied:

$$\pi_{2n-2}(x;\ F_\Delta) \equiv \sum_{\nu=1}^{n} \pi_{2n-2}(x_\nu\ F_\Delta) h_\nu(x) + \sum_{\nu=1}^{n} \pi'_{2n-2}(x_\nu;\ F_\Delta) f_\nu(x), \tag{29}$$

where $h_\nu(x)$ an $\mathfrak{h}_\nu(x)$ denote the fundamental polynomials of interpolation of the first and the second kind for the abscissas (1), that is

$$h_\nu(x_\mu) = \begin{cases} 0 & \text{if } \nu \neq \mu, \\ 1 & \text{if } \nu = \mu, \end{cases} \quad h'_\nu(x_\mu) = 0,$$

$$\mathfrak{h}_\nu(x_\mu) = 0, \quad \mathfrak{h}'_\nu(x_\mu) = \begin{cases} 0 & \text{if } \nu \neq \mu \\ 1 & \text{if } \nu = \mu \end{cases}. \tag{30}$$

For the abscissas (1) L. FEJÉR [2] proved the inequality

$$h_\nu(x) \equiv \left[1 - \frac{x_\nu}{1-x_\nu^2}(x-x_\nu)\right] l_\nu^2(x) \geqq 0,$$

$$(-1 \leqq x \leqq 1;\ \nu = 1, 2, \ldots, n;\ n = 1, 2, \ldots) \tag{31}$$

where

$$l_\nu(x) = \frac{T_n(x)}{T'_n(x_\nu)(x-x_\nu)} \tag{32}$$

is the fundamental polynomial of Lagrange interpolation and

$$T_n(x) = \cos\,(n \arccos x).$$

The explicit form of the polynomial $\mathfrak{h}_\nu(x)$ is

$$\mathfrak{h}_\nu(x) = (x-x_\nu) l_\nu^2(x). \tag{33}$$

the following identity is well-known (see L. FEJÉR [2])

$$\sum_{\nu=1}^{n} h_\nu(x) \equiv 1. \tag{34}$$

Now we can turn to the proof of our theorem. On account of (30) the explicit form of the polynomial $\overline{H}_{2n-1}(x;f)$ satisfying (6) and (7) is

$$\overline{H}_{2n-1}(x;\ f) \equiv \sum_{\nu=1}^{n} y_\nu h_\nu(x) + \sum_{\nu=1}^{n} y'_\nu \mathfrak{h}_\nu(x). \tag{35}$$

Taking into consideration (35), (29), (31), (7), furthermore (21), (28), (34) and the fact that $y_\nu = F_\Delta(x_\nu; f)$ by (8), we get the estimation

$$|f(x)-\overline{H}_{2n-1}(x;\, f)| \leqq |f(x)-F_\Delta(x;\, f)| + |F_\Delta(x;\, f)-\pi_{2n-2}(x;\, F_\Delta)| +$$

$$+\sum_{\nu=1}^{n} |\pi_{2n-2}(x_\nu;\, F_\Delta)-y_\nu|\, h_\nu(x) \leqq O(1)\omega\left(\frac{1}{n};\, f\right)+O(1)\omega\left(\frac{1}{n};\, f\right)+$$

$$+O(1)\omega\left(\frac{1}{n};\, f\right)\sum_{\nu=1}^{n} h_\nu(x) = O(1)\omega\left(\frac{1}{n};\, f\right)$$

and the Theorem is proved.

References

[1] L. FEJÉR, Über Interpolation, *Nachrichten der Ges. der Wiss.* Göttingen, Math.–Phys., (1916), 66–91.

[2] L. FEJÉR, Lagrangesche Interpolation und die zugehörigen konjugierten Punkte, *Math. Annalen*, **106** (1932), 1–55.

[3] D. JACKSON, *Über die Genauigkeit der Annäherung stetiger Funktionen durch Polynome gegebenen Grades und trigonometrische Summen gegebener Ordnung*, Inaug. Diss., Göttingen, 1911.

MINIMALTETRAEDER BIKONJUGIERTER GITTER

Von

JOHANNES BÖHM und WALTER BÖRNER

Sektion Mathematik der Friedrich – Schiller – Universität Jena

(Eingegangen am 18. April 1983)

1. Einleitung

Bei Lagerungs- und Packungsproblemen spielen unter anderem bikonjugierte Punktsysteme eine besondere Rolle. Ist eine diskrete Punktmenge $\mathfrak{M}$ in einem n-dimensionalen euklidischen Raum vorgegeben, dann wird die Menge $\mathfrak{M}'$ der Eckpunkte der Dirichlet-Voronoischen Zellen von $\mathfrak{M}$ konjugiert zu $\mathfrak{M}$ genannt. Die Punktmenge $\mathfrak{M}$ heißt bikonjugiert, wenn $\mathfrak{M}'$ zu $\mathfrak{M}$ und $\mathfrak{M}$ zu $\mathfrak{M}'$ konjugiert sind; es gilt folglich für bikonjugierte Mengen $\mathfrak{M}$ die Beziehung $(\mathfrak{M}')' = \mathfrak{M}$. Unter Verwendung von Ergebnissen von E. S. FEDOROV [1] hat M. HOLLAI gezeigt, daß im dreidimensionalen euklidischen Raum genau drei Typen von bikonjugierten Punktgittern existieren, nämlich Quadergitter, Prismengitter und spezielle Tetraedergitter. Die L-Polyeder (Stützpolyeder) von bikonjugierten Tetraedergittern sind Tetraeder, die zwei einander gegenüberliegende rechte Keilwinkel besitzen.

Ein weiteres Anliegen in obigem Zusammenhang ist die Angabe von bikonjugierten Systemen von Einheitskugelpackungen. Dabei wird eine Einheitskugelpackung ϱ-System genannt, wenn die Stützkugelradien der Punktmenge $\mathfrak{M}_0$, die von den Einheitskugel-Mittelpunkten der Packung erzeugt wird, mindestens die Größe $\varrho+1$ besitzen. Das ϱ-System wird bikonjugiert genannt, wenn $\mathfrak{M}_0$ bikonjugiert ist.

Um Aussagen über maximale Dichte von bikonjugierten gitterförmigen ϱ-Systemen von Einheitskugelpackungen zu machen, sind Stützpolyeder der entsprechenden Punktgitter zu suchen mit minimalem Volumen bei vorgegebenem Stützkugelradius $R = \varrho+1$, mit Stützkugelmittelpunkt nicht außerhalb des Stützpolyeders und wegen der Packungseigenschaft von sich nicht durchdringenden Einheitskugeln mit Kantenlängen nicht kleiner als Zwei.

Der Quader- und der Prismentyp eines solchen ϱ-Systems sind ausführlich untersucht worden (vgl. M. HOLLAI [3]). Dagegen fehlte ein genaueres Studium des Tetraedertyps eines bikonjugierten gitterförmigen ϱ-Systems maximaler Dichte. Um einen solchen Typ vollständig beschreiben zu

können, genügt es, alle diejenigen Tetraeder im euklidischen Raum mit minimalem Volumen zu kennen, die folgende Eigenschaften haben:

(a) Es gibt zwei einander gegenüberliegende rechte Keilwinkel,
(b) es gibt keine stumpfen Keilwinkel,
(c) die kürzeste Kantenlänge ist nicht kleiner als 2,
(d) der Umkugelradius hat die vorgegebene Länge R.

Eigenschaft (b) resultiert aus der Tatsache, daß der Mittelpunkt der Umkugel des Tetraeders nicht außerhalb desselben liegen darf.

Das Ziel dieser Arbeit ist nun die vollständige Lösung dieser Minimumaufgabe. Es können sämtliche Minimaltetraeder mit den Eigenschaften (a) bis (d) und deren Volumen in Abhängigkeit von der Umkugelradiusgröße R hier angegeben werden.

Der erste der beiden Verfasser hatte 1979 nach dem Studium von zahlreichen Spezialfällen die Vermutung ausgesprochen, daß solche Minimaltetraeder gewisse wohlbestimmte symmetrische Tetraeder sind, die wegen der Forderung (b) jeweils in Abhängigkeit von der Umkugelradiusgröße R zu einem von zwei möglichen Tetraedertypen (dreikantengleiche bzw. Orthogonal-Tetraeder) gehören. Der Grenzfall liegt bei $R = \sqrt{3}$, und nur dieser kann gleichzeitig zu beiden Typen gerechnet werden. Für kleinere Umkugelradien kommen in dem Minimaltetraeder außer den Rechtwinkelkanten gleicher Größe drei weitere untereinander gleichgroße Tetraederkanten (dreikantengleiche Tetraeder) vor. Für größere Umkugelradien gibt es immer genau zwei inkongruente inhaltsgleiche Minimaltetraeder, die sich hinsichtlich ihrer Kantenlängen lediglich in der Länge einer Rechtwinkelkante unterscheiden. In einem solchen Minimaltetraeder kommt außer einem bzw. zwei Paaren gleichgroßer Kanten immer ein weiterer rechter Keilwinkel vor, so daß diese Minimaltetraeder stets Orthogonaltetraeder (Orthoscheme) sind. Diese angegebenen Beschreibungen reichen hin, die Mannigfaltigkeit aller Tetraeder beider Typen zu konstruieren. Es sei angemerkt, daß umgekehrt jedoch nicht alle Tetraeder dieser beiden Typen als minimale möglich sind. – Ein Beweis dieser Vermutung wurde dann vom ersten Verfasser auf elementar-geometrischem Wege durch Rückgang auf ein äquivalentes ebenes Kreisproblem geführt. Im folgenden zweiten Abschnitt wird die Beweisidee dafür skizziert. Daraus unmittelbar u. W. bisher noch nicht bekannte ableitbare Eigenschaften, Sätze und Folgerungen über Kreisvierecke sollen an anderer Stelle dargelegt werden. – Dem zweiten Verfasser ist es gelungen, koordinaten-geometrisch den Beweis der o. g. Aussagen zu führen. Dabei haben sich bemerkenswerte innergeometrische Beziehungen ergeben, so daß es für richtig gehalten wird, diesen weiteren Beweis hier im dritten Abschnitt ausführlich darzulegen.

2. Lösung der Aufgaben mit Hilfe eines äquivalenten ebenen Problems

$\mathfrak{T}$ sei die Menge der euklidischen Tetraeder mit den Eigenschaften (a) und (b). Ein beliebiges Tetraeder $T \in \mathfrak{T}$ besitzt ein eindeutig bestimmtes gemeinsames Lot der Länge l zu den beiden gegenüberliegenden Kanten, die

die Scheitelkanten der rechten Keilwinkel sind (Rechtwinkelkanten). T wird so in ein kartesisches Koordinatensystem eingebettet, daß der Mittelpunkt des genannten gemeinsamen Lotes dessen Ursprung darstellt, das Lot selbst in die y-Achse fällt und eine Rechtwinkelkante parallel zur x-Achse verläuft. Die Orthogonalprojektion T' von T auf die xz-Ebene (Ebene ε_0 genannt) ergibt dann ein Kreisviereck. Der Mittelpunkt M' seines Umkreises fällt mit dem Mittelpunkt M der Umkugel von T zusammen. Wegen der Eigenschaft (b) liegt M nicht außerhalb von T und $M' = M$ nicht außerhalb von T'. In dieser kanonischen Lage können dann für das Tetraeder T sowie für seine Projektion T' – abgesehen von einer Orientierung – vier charakteristische Parameter l, α, p_1, p_2 mit o. B. d. A. $0<\alpha\leqq\frac{\pi}{2}$, $0<p_1\leqq p_0$, $p_0 = \min\left(l, p_2, \frac{l^2}{p_2}\right)$, $0<p_2$ abgelesen werden, die T und T' eindeutig beschreiben. Dabei ist α die Größe des orientierten Winkels zwischen den beiden Vierecksdiagonalen von T', und $p_i' = p_i \operatorname{cosec} \alpha$ $(i = 1, 2)$ stellen die Größe der Diagonalenabschnitte dar, die den Winkel der Größe α erzeugen. Abgesehen von der Orientierung können T und T' einander eindeutig zugeordnet werden, für ihren Inhalt gilt

$$\text{(1)} \qquad \operatorname{vol}(T) = \frac{l}{3} \operatorname{vol}_2(T').$$

Für den ersten Teil der Untersuchung wird zur Normierung ähnlicher Tetraeder die Menge $\mathfrak{T}$ auf die Menge $\mathfrak{T}^*$ mit allen $T\in\mathfrak{T}$ eingeschränkt, deren Projektionen T' den Umkreisradius $R' = 1$ haben. Das ermöglicht einen übersichtlichen Vergleich der Vierecksinhalte. Nun wird $\mathfrak{T}^*$ in Klassen $\mathfrak{T}_l^*$ mit Tetraeder gleicher Lotlänge l $(0<l\leqq 1)$ (Lotklassen genannt) zerlegt, die ihrerseits in Unterklassen $\mathfrak{T}_{l,\alpha}^*$ zerfallen, zu denen alle Tetraeder $T\in\mathfrak{T}_l^*$ gehören, in deren Projektion T' die Vierecksdiagonalen sich unter einem Winkel der Größe α, $0<\alpha\leqq\frac{\pi}{2}$ schneiden. Diese Normierung, Klassen- und Unterklassenbildung kann auf $\mathfrak{T}'$ übertragen werden und ergibt die Mengen $\mathfrak{T}^{*\prime}$, $\mathfrak{T}_l^{*\prime}$ und $\mathfrak{T}_{l,\alpha}^{*\prime}$ (vgl. Abb. 1). In der Ebene ε_0 gelten bezüglich der Menge $\mathfrak{T}_l^{*\prime}$ die folgenden Hilfssätze:

Hilfssatz 1. *Die Kreislinie um M' mit Radius $\varrho = \sqrt{1-\frac{l^2}{\sin^2\alpha}}$ ist genau die Menge der Diagonalenschnittpunkte der Vierecke aus $\mathfrak{T}_{l,\alpha}^{*\prime}$.*

Hilfssatz 2. *Zwei Kreisvierecke aus $\mathfrak{T}^{*\prime}$ mit paarweise gleichgroßen Seiten sind inhaltsgleich und gehören zu ein und derselben Lotklasse.*

Darum bietet sich für die weitere Betrachtung eine Beschreibung der Vierecke $T'\in\mathfrak{T}_l^{*\prime}$ durch die Angabe des Diagonalenschnittpunktes von T' durch Polarkoordinaten (ϱ, t) an, wofür M' der Pol und die Halbgerade mit dem Anfangspunkt M' und demselben Richtungssinn wie die positive x-Achse die Polarachse p sind (vgl. Abb. 1). Bei Beachtung des Zusammen-

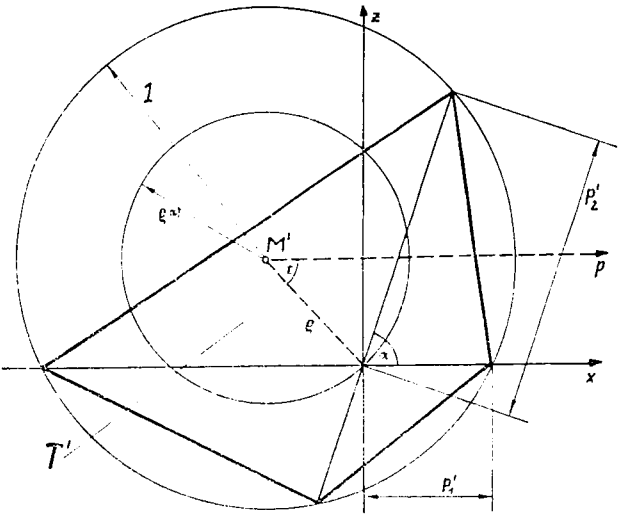

Abb. 1.

hangs zwischen ϱ und der Winkelgröße α kann dann jedes $T' \in \mathfrak{T}_l^{*\prime}$ außer durch l durch die Parameter α und t charakterisiert werden. Wegen der Möglichkeit der Einschränkung der ursprünglichen Parameterwerte auf die oben angegebenen Bereiche, was hinsichtlich Inhaltsuntersuchungen auch durch Hilfssatz 2 noch einmal legitimiert wird, genügt es hier, α und t im Bereich $\mathbf{B}_0$ mit

$$\arcsin l = \alpha_0 \leqq \alpha \leqq \frac{\pi}{2} \text{ und } -\left(\frac{\pi}{2} - \frac{\alpha}{2}\right) \leqq t \leqq \frac{\alpha}{2} \tag{2}$$

zu wählen. Für den Inhalt von $T' = T'(\alpha, t) \in \mathfrak{T}_l^{*\prime}$ ergibt sich

$$\begin{aligned}(\mathrm{vol}_2(T'(\alpha, t))^2 = {} & \frac{(\sin^2\alpha - l^2)^2}{2\sin^2\alpha}\cos(4t - 2\alpha) + \\ & + \frac{2(\sin^4\alpha - l^4)\cos\alpha}{\sin^2\alpha}\cos(2t - \alpha) + \\ & + 3\frac{(\sin^2\alpha + l^2)^2}{2\sin^2\alpha} - (\sin^2\alpha - l^2)^2 - 2l^2.\end{aligned} \tag{3}$$

Für die Vierecke $T' = T'(\alpha, t)$ aus $\mathfrak{T}_l^{*\prime}$ mit je zwei gleichlangen (gegenüberliegenden) Seiten der Länge $m_{T'}$ ist entweder $t = -\left(\frac{\pi}{2} - \frac{\alpha}{2}\right)$ oder $t = \frac{\alpha}{2}$. Sie werden zu der Menge $\mathfrak{D}_l^{(1)}$ bzw. $\mathfrak{D}_l^{(2)}$. zusammengefaßt. Genau die

zwei Vierecke $D_l^{(i)} = D_l^{(i)}(\alpha_l, t) \in \mathfrak{D}_l^{(i)} (i = 1, 2)$ haben drei gleichlange Seiten, und es gilt bei diesen zwischen l und α der Zusammenhang

$$l^2 = (1+2\cos\alpha_1)(1-\cos\alpha_1)^2 \tag{4}$$

bzw.

$$l^2 = (1-2\cos\alpha_2)(1+\cos\alpha_2)^2; \tag{5}$$

α_i $(i = 1, 2)$ sind die einzigen Winkelgrößen mit $0 < \alpha_i \leqq \frac{\pi}{2}$, die für $0 < l \leqq 1$ die jeweiligen Gleichungen (4) bzw. (5) erfüllen. Für die Größe der drei gleichlangen Seiten ergibt sich $m^2_{D_l^{(1)}} = 2(1-\cos\alpha_2)$ bzw. $m^2_{D_l^{(2)}} = 2(1+\cos\alpha_2)$. Auf der Menge $\mathfrak{T}_l^{*\prime}$ besitzt $D_l^{(1)}$ die längste kleinste und $D_l^{(2)}$ die kürzeste längste Vierecksseite der Größe $m_{D_l^{(1)}}$ bzw. $m_{D_l^{(2)}}$ Die zu den Vierecken $D_l^{(i)}$ gehörenden Tetraeder sind spezielle dreikantengleiche Tetraeder. Die Vierecke aus $\mathfrak{D}_l^{(i)}$ sind durch weitere Extremaleigenschaften ausgezeichnet, wie sich aus (3) ableiten läßt:

HILFSSATZ 3. *Auf der Menge $\mathfrak{T}_{l,\alpha}^{*\prime}$ nehmen die Vierecke $T' \in \mathfrak{D}_l^{(i)} \cap \mathfrak{T}_{l,\alpha}^{*\prime}$ $(i = 1, 2)$ extremalen Flächeninhalt an, und zwar bei $i = 2$ maximalen Inhalt, bei $i = 1$, $\cos\alpha \geqq 1-l$ minimalen Inhalt und bei $i = 1$, $0 \leqq \cos\alpha < 1-l$ (relativ) maximalen Inhalt.*

Wegen der letzten Aussage in Hilfssatz 3 muß es für α mit $0 \leqq \cos\alpha < 1-l$ und $t \in \left]-\left(\frac{\pi}{2}-\frac{\alpha}{2}\right), \frac{\alpha}{2}\right[$ ein Viereck in $\mathfrak{T}_{l,\alpha}^{*\prime}$ mit minimalem Inhalt geben. Für dieses ist

$$\cos(2t-\alpha) = -\frac{\sin^2\alpha + l^2}{\sin^2\alpha - l^2}\cos\alpha\,(> -1).$$

$\mathfrak{D}_l^{(0)}$ sei die Menge dieser letztgenannten Vierecke aus $\mathfrak{T}_l^{*\prime}$ zuzüglich des Grenzfalles $A_l = T'\left(\alpha_s, -\left(\frac{\pi}{2}-\frac{\alpha_s}{2}\right)\right)$ mit $\cos\alpha_s = 1-l$. Folglich besteht der Durchschnitt $\mathfrak{D}_l^{(0)} \cap \mathfrak{D}_l^{(1)}$ nur aus dem einen Element A_l, und es ist für $l < 1$ stets $\alpha_1 < \alpha_2, \alpha_s$. In $\mathfrak{T}_l^{*\prime}$ gibt es wegen der Einschränkung auf den Bereich $\mathbf{B}_0$ (vgl. (2)) keine weiteren Extremwerte bezüglich des Flächeninhalts als die hier angegebenen.

SATZ 1. *Auf der Menge $\mathfrak{T}_l^{*\prime}$ nehmen genau die Vierecke $\mathfrak{D}_l^{(1)}$ den minimalen und $\mathfrak{D}_l^{(2)}$ den maximalen Flächeninhalt an.*

Gleichzeitig erhält man Monotonieaussagen, die zusammengefaßt werden zu

HILFSSATZ 4. *Die Funktion $(\mathrm{vol}_2(T'))^2$ ist auf der Menge $\mathfrak{D}_l^{(1)}$ mit $\arcsin l = \alpha_0 \leqq \alpha \leqq \alpha_1$ eine streng monoton fallende Funktion und auf der Menge $\mathfrak{D}_l^{(1)}$ mit $\alpha_1 \leqq \alpha \leqq \frac{\pi}{2}$ sowie auf der Menge $\mathfrak{D}_l^{(0)}$ mit $\alpha_s \leqq \alpha \leqq \frac{\pi}{2}$ eine streng monoton wachsende Funktion in α.*

Die Menge $\mathfrak{T}^{*\prime}_{l,\alpha_0}$ – im Spezialfall bei $l = 1$ handelt es sich um die Menge $\mathfrak{T}^{*\prime}_{1,\frac{\pi}{2}}$ – ist einelementig und fällt mit der Menge $\mathfrak{D}_l^{(1)} \cap \mathfrak{D}_l^{(2)}$ zusammen. Die Funktionswerte der Inhaltsfunktion von $T' \in \mathfrak{D}_l^{(1)} \cap \mathfrak{T}^{*\prime}_{l,\alpha_0} = \mathfrak{T}^{*\prime}_{l,\alpha_0}$ einerseits und von $T' \in \mathfrak{D}_l^{(0)} \cap \mathfrak{T}^{*\prime}_{l,\frac{\pi}{2}}$ andererseits stimmen infolge von Hilfssatz 2 überein. Ebenso sind die Funktionswerte der Inhaltsfunktion für die Vierecke $T'\left(\frac{\pi}{2}, -\frac{\pi}{4}\right) \in \mathfrak{D}_l^{(1)}$ und $T'\left(\frac{\pi}{2}, \frac{\pi}{4}\right) \in \mathfrak{D}_l^{(2)}$ gleich.

Wegen der Eigenschaft (b) sind die Vierecke $D_l^{(1)}$ nur für $\frac{1}{2} \leqq l^2 \leqq 1$ als minimale brauchbar. Für ihren Inhalt ergibt sich

$$\min_{T' \in \mathfrak{T}_l^{*\prime}} (\mathrm{vol}_2(T')) = \mathrm{vol}_2(D_l^{(1)}) = 2 \sin^3 \alpha_1.$$

Die Diagonalenschnittpunkte markanter Kreisvierecke aus $\mathfrak{T}_l^{*\prime}$ für ein festes l mit $\frac{1}{2} \leqq l^2 < 1$ liegen etwa so, wie in Abbildung 2 skizziert ist (dort für $l^2 = 0{,}648$); ihre Bezeichnung wird mit der für die zugehörigen Vierecke bzw. Vierecksmengen identifiziert.

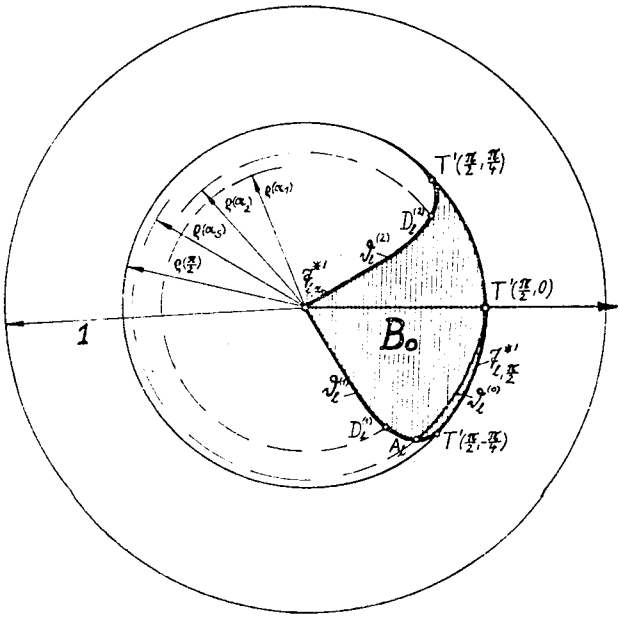

Abb. 2.

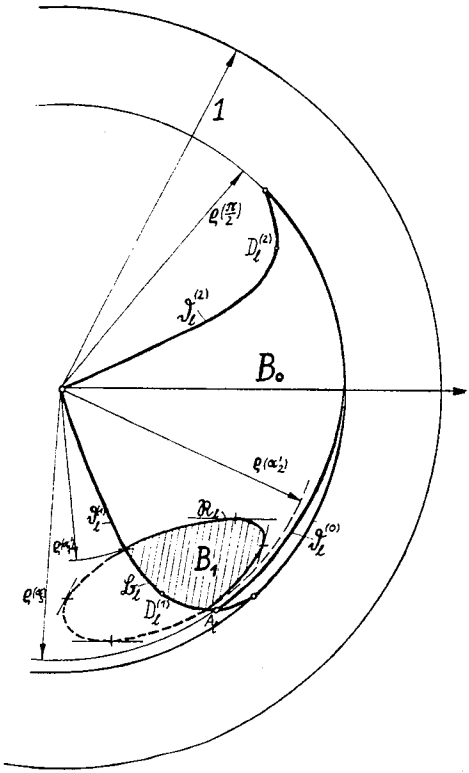

Abb. 3.

Für $l^2 < \frac{1}{2}$ liegt der Mittelpunkt des Umkreises von $D_l^{(1)}$ außerhalb $D_l^{(1)}$. Darum sind bei diesen Lotlängen weitere Untersuchungen erforderlich.

Die Menge der für die vorliegenden Betrachtungen wegen der Forderung (b) nicht zulässigen Vierecke aus $\mathfrak{T}_l^{*\prime}$ werde mit $\mathfrak{B}_l$ bezeichnet, die Menge der Vierecke aus $\mathfrak{T}_l^{*\prime}$, deren Umkreismittelpunkt jeweils auf einer ihrer Seiten liegt, so daß diese Seite ein Durchmesser des Umkreises ist, sei $\mathfrak{R}_l$. Die dazugehörigen Tetraeder sind Orthoscheme. Die Menge der Diagonalenschnittpunkte der Vierecke aus $\mathfrak{B}_l$ ergibt den konvexen Bereich $\mathbf{B}_1$ (vgl. Abb. 3; dort ist $l^2 = 0{,}446$ gewählt). Der Rand dieses Bereiches setzt sich aus den Diagonalenschnittpunkten aller Vierecke der Mengen $\mathfrak{B}_l := \mathfrak{D}_l^{(1)} \cap \mathfrak{B}_l$ und $\mathfrak{R}_l$ zusammen. Zu $\mathfrak{B}_l$ gehören genau alle Vierecke $T'\left(\alpha, -\left(\frac{\pi}{2} - \frac{\alpha}{2}\right)\right) \in \mathfrak{D}_l^{(1)}$, bei denen für α die Beziehung $g > 0$ gilt mit

$$g = 2 \cos \alpha (1 - \cos \alpha) - l^2. \tag{6}$$

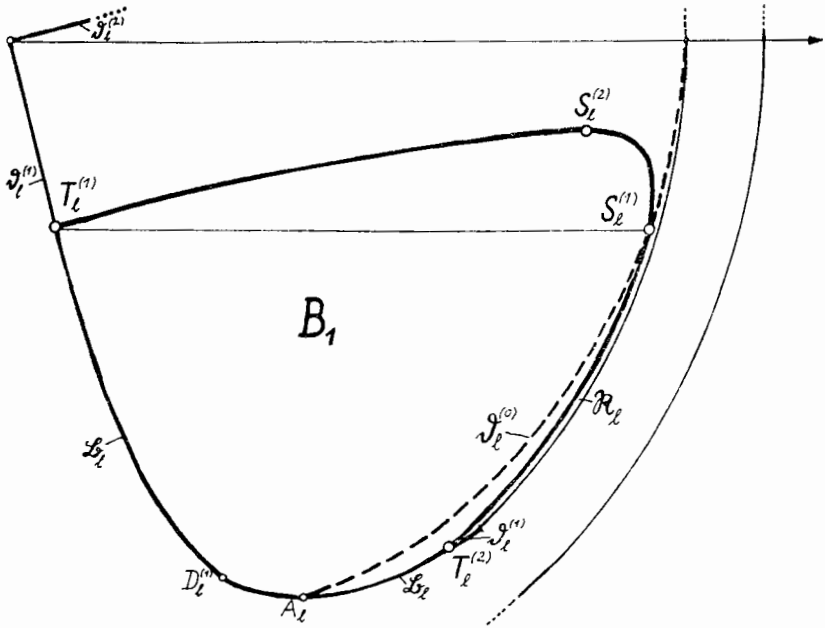

Abb. 4.

$\left\{T_l^{(1)}\left(\alpha_1', -\left(\frac{\pi}{2}-\frac{\alpha_1'}{2}\right)\right), T_l^{(2)}\left(\alpha_2', -\left(\frac{\pi}{2}-\frac{\alpha_2'}{2}\right)\right)\right\}$ ist für $0<l^2<\frac{1}{2}$ die zweielementige Menge $\mathfrak{D}_l^{(1)}\cap\mathfrak{R}_l$, wobei α_i' $(i = 1, 2)$ die Lösungen der Gleichung $g = 0$ sind und $0<\alpha_1'<\alpha_2'<\frac{\pi}{2}$ gelte. Für alle l mit $\frac{1}{2}\leqq l^2\leqq 1$ ist $\mathfrak{B}_l = \emptyset$, und für $0<l^2<\frac{1}{2}$ enthält $\mathfrak{B}_l$ jeweils $D_l^{(1)}$ aber niemals $D_l^{(2)}$ (vgl. Abb. 4; dort für $l^2 = 0{,}21$ skizziert. Der Diagonalenschnittpunkt von $A_l = T'\left(\alpha_s, -\left(\frac{\pi}{2}-\frac{\alpha_s}{2}\right)\right)$ liegt genau für $0<l<\frac{2}{3}$ in $\mathbf{B}_1$). Auf Grund der Lage der Diagonalenschnittpunkte der Vierecke mit extremalem Inhalt in Bezug auf die Menge $\mathfrak{T}_l^{*\alpha}$ und der in Hilfssatz 4 angegebenen Monotonieverhältnisse muß für $l^2<\frac{1}{2}$ ein zulässiges Viereck mit minimalem Inhalt in $\mathfrak{R}_l$ liegen, so daß es für die weiteren Betrachtungen bei festem l mit $l^2<\frac{1}{2}$ ausreicht, nur Vierecke auf der Menge $\mathfrak{R}_l$ zu betrachten und dort diejenigen mit minimalem Inhalt zu bestimmen.

Es sei jetzt $l^2 < \frac{1}{2}$. Die Menge $\mathfrak{R}_{l,\alpha} := \mathfrak{R}_l \cap \mathfrak{T}^{*\prime}_{l,\alpha}$ besteht für $g \geqq 0$ aus genau einem Element, sonst ist sie gleich der leeren Menge. Um den Inhalt des Vierecks T', das das Element der einelementigen Menge $\mathfrak{R}_{l,\alpha}$ sein möge, einfach beschreiben zu können, wird ein Parameter γ, $0 < \gamma < \frac{\pi}{2}$, eingeführt. Dieser kann bei dem Viereck in kanonischer Lage als die Größe des Winkels interpretiert werden, der gebildet wird einerseits von der parallel zur Polarachse (x-Achse) verlaufenden Vierecksdiagonalen und andererseits von der Vierecksseite, die mit einem Durchmesser des Umkreises zusammenfällt (vgl. Abb. 5). Zwischen α und γ besteht die Bindung

$$G := \cos\alpha \sin\gamma \operatorname{sind}(\alpha-\gamma) - \frac{1}{4} l^2 = 0.$$

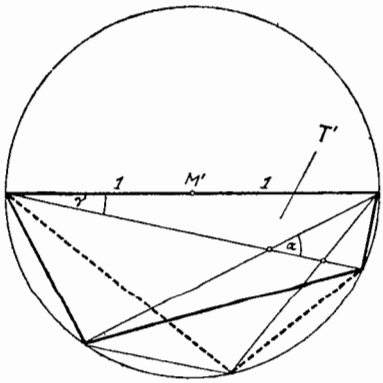

Abb. 5.

Dann ergibt sich

$$\mathrm{vol}_2(T') = 2 \sin\alpha \cos\gamma \cos(\alpha-\gamma).$$

Zur Bestimmung der gesuchten Extremwerte für den Vierecksinhalt wird die Funktion

$$H := \frac{1}{2} \mathrm{vol}_2(T') + \lambda G$$

hinsichtlich Extremwerte auf dem Bereich $\mathfrak{R}_l$ untersucht. Es ergeben sich genau die drei möglichen Stellen $\gamma = \frac{\alpha}{2}$, $\gamma = \frac{\pi}{2} - \alpha$ und $\gamma = 2\alpha - \frac{\pi}{2}$. Eine Auswertung dieses Ergebnisses zeigt, daß alle drei Fälle durch Seitentausch von Vierecken entsprechend Hilfssatz 2 ineinander überführt werden können. Darum reicht es aus, etwa nur den ersten Fall zu diskutieren: Für $\gamma = \frac{\alpha}{2}$

erhält man aus $G = 0$ die Beziehung $g = 0$, woraus die Lösungen α_1' und α_2' mit $0<\alpha_1'<\alpha_1<\alpha_2'<\frac{\pi}{2}$ bestimmt werden können. Die dazugehörigen Vierecke $T_l^{(i)} = T_l^{(i)}\left(\alpha_i', -\left(\frac{\pi}{2}-\frac{\alpha_i'}{2}\right)\right)$ $(i = 1, 2)$ sind infolge $g = 0$ die beiden Elemente von $\mathfrak{D}_l^{(1)} \cap \mathfrak{R}_l$, und ihr Inhaltsvergleich ergibt wegen $\mathrm{vol}_2(T_l^{(i)}) = \sin\alpha_i'(1+\cos\alpha_i')$, daß $\mathrm{vol}_2(T_l^{(1)}) < \mathrm{vol}_2(T_l^{(2)})$ gilt.

Die anderen beiden erhaltenen Werte von γ geben zu weiteren vier Kreisvierecken Anlaß, die zwar nicht zu $\mathfrak{D}_l^{(1)}$, aber von denen genau zwei, $S_l^{(1)}$ und $S_l^{(2)}$, zu $\mathfrak{R}_l$ gehören, die, wie bereits erwähnt, entsprechend Hilfssatz 2 durch Seitentausch aus $T_l^{(1)}$ bzw. $T_l^{(2)}$ hervorgehen und somit mit diesen jeweils inhaltsgleich sind. Im einzelnen gilt, daß für $\alpha_i'' = \frac{\pi}{2} - \frac{\alpha_i'}{2}$ die Menge $\mathfrak{R}_{1,\alpha_i''}$ aus dem genau einen Element $S_l^{(i)}$ besteht (vgl. Abb. 4; es ergibt sich, daß stets gilt $S_l^{(1)} \notin \mathfrak{D}_l^{(0)}$. Der Schnittpunkt zwischen den mit $\mathfrak{D}_l^{(0)}$ und $\mathfrak{R}_l$ in dieser Abbildung bezeichneten Kurvenbögen, der nur bei $0<l\leqq\frac{2}{3}$ existiert, muß dann auf Grund der Lage der Extremwerte bezüglich der Mengen $\mathfrak{T}_{l,\alpha}^{*\prime}$ für $l<\frac{2}{3}$ immer auf demjenigen Bogen zwischen $T_l^{(2)}$ und $S_l^{(1)}$ liegen, der $T_l^{(1)}$ nicht enthält). Da es keine weiteren Extremstellen auf $\mathfrak{R}_l$ gibt, kann somit geschlossen werden, daß genau die beiden Vierecke $T_l^{(1)}$ und $S_l^{(1)}$ minimalen und gleichen Inhalt auf $\mathfrak{R}_l$ annehmen. Darum gilt insgesamt

SATZ 2. *Bei* $0<l^2<\frac{1}{2}$ *nehmen genau die Vierecke* $T_l^{(1)}$ *und* $S_l^{(1)}$ *auf* $\mathfrak{U}_l := \mathfrak{T}_l^{*\prime} \setminus \mathfrak{V}_l$ *den minimalen Flächeninhalt an.*

Für den Wert des Inhalts gilt

$$\min_{T'\subset\mathfrak{U}_l} (\mathrm{vol}_2(T')) = \mathrm{vol}_2(T_l^{(1)}) = \mathrm{vol}_2(S_l^{(1)}) = \sin\alpha_1'(1+\cos\alpha_1')$$

$$\text{mit } 0<l^2<\frac{1}{2}.$$

Da $T_l^{(1)}$ und $S_l^{(1)}$ zu $\mathfrak{D}_l^{(1)}$ gehören, haben sie jeweils zwei Seiten gleicher Länge $m_{T_l^{(1)}} = m_{S_l^{(1)}}$ mit $m_{T_l^{(1)}}^2 = 2(1-\cos\alpha_1')$. $T_l^{(1)}$ und $S_l^{(1)}$ besitzen auf der Menge $\mathfrak{U}_l$ die längste kleinste Vierecksseite der Größe $m_{T_l^{(1)}}$. Es genügt darum für das weitere, nur eines dieser beiden Vierecke, etwa $T_l^{(1)}$, zu betrachten.

Die Sätze 1 und 2 lassen sich gemäß (1) auf die ihnen zugeordneten Tetraeder übertragen, so daß es jetzt genügt, nur die den Vierecken $D_l^{(1)}$ zugeordneten dreikantengleichen Tetraeder und die den Vierecken $T_l^{(1)}$ zugeordneten Orthoscheme zu betrachten. Ein Inhaltsvergleich dieser Tetraeder bei verschiedenem l kann im Sinne der vorgelegten Aufgabe erst durchgeführt werden, wenn diese gleichen Umkugelradius besitzen. Darum werden

durch geeignete Ähnlichkeitsabbildungen die den Vierecken $D_l^{(1)}$ und $T_l^{(1)}$ zugeordneten Tetraeder auf solche mit einem Umkugelradius von ein und derselben Größe R abgebildet. Die Lotlängen $l\,(0<l\leqq 1)$ der Originaltetraeder dienen als Parameter zur Beschreibung der Bildtetraeder. Für das zugehörige Volumen V und die Länge k der kleinsten Kante der jeweiligen Bildtetraeder, die aus den Kreisvierecken $D_l^{(1)}$ bzw. $T_l^{(1)}$ hervorgegangen sind, erhält man in Abhängigkeit von dem Parameter $l = \mathrm{l}(\alpha)$

(7)
$$V^2 = \mathrm{V}^2(\mathrm{l}(\alpha)) = \begin{cases} \left(\dfrac{16}{3}\right)^2 R^6 \cdot \dfrac{(1-\cos\alpha)^5(1+2\cos\alpha)}{(5-5\cos\alpha+2\cos^2\alpha)^3} & \text{für } \dfrac{1}{2}\leqq l^2\leqq 1 \\ \left(\dfrac{4}{3}\right)^2 R^6 \cdot \dfrac{\cos\alpha(1-\cos\alpha)^2}{(2-\cos\alpha)^3} & \text{für } 0<l^2<\dfrac{1}{2}, \end{cases}$$

(8)
$$k^2 = \mathrm{k}^2(\mathrm{l}(\alpha)) = \begin{cases} 4R^2 \cdot \dfrac{(1-\cos\alpha)(3-2\cos\alpha)}{5-5\cos\alpha+2\cos^2\alpha} & \text{für } \dfrac{1}{2}\leqq l^2\leqq 1 \\ 4R^2 \cdot \dfrac{1-\cos\alpha}{2-\cos\alpha} & \text{für } 0<l^2<\dfrac{1}{2}. \end{cases}$$

Die Funktionen V und k sind stetig in α und in l, insbesondere auch an der Stelle $l^2 = \dfrac{1}{2}$ bzw. für den dazugehörigen Wert $\alpha = \dfrac{\pi}{3}$, sowie streng monoton wachsende Funktionen auf dem Intervall $[0, 1]$ für l bzw. auf $\left[0, \dfrac{\pi}{2}\right]$ für α. Um für die Tetraeder mit dem geforderten minimalen Volumen die Eigenschaft (c) zu gewährleisten, sind darum dem Umkugelradius der Größe R genau diejenigen bei den oben genannten Ähnlichkeitsabbildungen entstandenen Tetraeder als solche minimalen Volumens zuzuordnen, für die die Kantenfunktion k in (8) den Wert 2 annimmt. Denn das sind gerade diejenigen Tetraeder, die auf den zum Vergleich zugelassenen Mengen von Tetraedern gleicher Lotlänge die größte kleinste Kante besitzen, die dann hier die Länge 2 hat. Demgemäß ergibt sich aus (8) der Zusammenhang

(9)
$$\cos\alpha = \begin{cases} \dfrac{1}{4(R^2-1)}\left(5R^2-5-\sqrt{(R^2-1)(R^2+15)}\right) & \text{für } \dfrac{5}{3}\leqq R^2\leqq 3 \\ \dfrac{R^2-2}{R^2-1} & \text{für } 3<R^2. \end{cases}$$

Folglich kann zu jedem Umkugelradius R mit $R^2\in\left[\dfrac{5}{3}, \infty\right[$ aus (9) ein wohl-

bestimmtes α und daraus aus (4) bei $\frac{5}{3} \leqq R^2 \leqq 3$ bzw. bei $3 < R^2$ aus (6) (dort für $g = 0$) ein wohlbestimmtes l gefunden werden. Genau die Tetraeder mit dem Umkugelradius R, deren oben beschriebene kanonische Projektion zu den Kreisvierecken $D_l^{(1)}$ bzw. $T_l^{(1)}$ und $S_l^{(1)}$ mit dem aus (9) berechenbaren $l(\alpha)$ ähnlich sind, nehmen auf der Menge aller Tetraeder mit den Eigenschaften (a), (b), (c) und (d) minimales Volumen an. Für den Wert $V_{\min}$ des Volumens erhält man aus (7) und (9)

$$V_{\min}^2(R) = \tag{10}$$

$$= \begin{cases} \frac{1}{9}[2(R^4+11R^2+12)\cdot\sqrt{(R^2-1)(R^2+15)} - \\ \qquad -(2R^6+36R^4+114R^2-88)] & \text{für } \frac{5}{3} \leqq R^2 \leqq 3 \\ & \left(bzw.\ \frac{1}{2} \leqq l^2 \leqq 1\right) \\ & \text{für } 3 \leqq R^2 \\ \frac{16}{9}(R^2-2) & \left(bzw.\ 0 < l^2 \leqq \frac{1}{2}\right). \end{cases}$$

Durch Radizieren ergibt sich dann daraus die gesuchte Funktion von R für das Volumen der Tetraeder minimalen Volumens mit den geforderten Eigenschaften (vgl. die zusammenfassende Darstellung am Ende von Abschnitt 3 hinsichtlich Gestalt und Volumen dieser Tetraeder).

Es sei angemerkt, daß sich (9) andererseits auch eindeutig nach R^2 auflösen läßt. Man erhält dann als Zusammenhang zwischen Umkugelradius R und Parameter α eines Tetraeders von minimalem Volumen und mit den Eigenschaften (a), (b), (c) und (d)

$$R^2(\alpha) = \begin{cases} \frac{5-5\cos\alpha+2\cos^2\alpha}{(1-\cos\alpha)(3-2\cos\alpha)} & \text{für } \frac{\pi}{3} \leqq \alpha \leqq \frac{\pi}{2}, \\ \frac{2-\cos\alpha}{1-\cos\alpha} & \text{für } 0 < \alpha < \frac{\pi}{3}. \end{cases} \tag{9a}$$

Zusammen mit (7) ergibt sich daraus eine Darstellung von $V_{\min}^2$ als Funktion des Parameters α (bzw. $\cos\alpha$), der in dem dem Tetraeder kanonisch zugeordneten Kreisviereck die oben angegebene geometrische Bedeutung hat, in der Gestalt

$$V_{\min}^2(\alpha) = \tag{10a}$$

$$= \begin{cases} \left(\frac{16}{3}\right)^2 \frac{(1-\cos\alpha)^2(1+2\cos\alpha)}{(3-2\cos\alpha)^3} & \text{für } \frac{\pi}{3} \leqq \alpha \leqq \frac{\pi}{2} \\ \left(\frac{4}{3}\right)^2 \frac{\cos\alpha}{1-\cos\alpha} & \text{für } 0 < \alpha \leqq \frac{\pi}{3}. \end{cases}$$

Weitere Untersuchungen lassen erkennen, daß in (9a) und (10a) R^2 und $V^2_{\min}$ in dem angegebenen Intervall $\left]0, \frac{\pi}{2}\right]$ stetige und streng monoton wachsende Funktionen von α sind. Daraus kann dann geschlossen werden, daß die zugehörigen Umkehrfunktionen $\cos\alpha$ bzw. $V^2_{\min}$ in (9) und (10) ebenfalls stetige und streng monoton wachsende Funktionen von R^2 im Intervall $\frac{5}{3} \leqq R^2 < \infty$ sind.

Als spezielle Werte, die für weitere praktische Berechnungen von besonderer Wichtigkeit sind, erhält man aus (9), (10) bzw. (10a):

$$V^2_{\min}(R) =$$

$$= \begin{cases} \frac{256}{243} & \text{für } R = \sqrt{\frac{5}{3}} \quad \textit{bzw.} \quad \alpha = \frac{\pi}{2} \\ \frac{4}{3} & \text{für } R = R_k \quad \textit{bzw.} \quad \alpha = \alpha_k \\ \frac{16}{9} & \text{für } R = \sqrt{3} \quad \textit{bzw.} \quad \alpha = \frac{\pi}{3}, \end{cases}$$

wobei α_k die kleinste Wurzel der Gleichung

$$152 \cos^3\alpha - 300 \cos^2\alpha + 162 \cos\alpha - 17 = 0$$

mit dem Näherungswert

$$\cos\alpha_k = 0{,}13751968\ldots \text{ bzw. } \alpha_k = 82{,}0956537\ldots{}^\circ$$

darstellt. Für den zugehörigen Umkugelradius R_k erhält man dann

$$R_k^2 = a = 1{,}85098238\ldots \text{ bzw. } R_k = 1{,}3605081\ldots,$$

was Anlaß zu einem ϱ-System mit $\varrho = \sqrt{a} - 1 = 0{,}3605081\ldots$ gibt.

3. Lösung der Aufgabe auf koordinatengeometrischem Wege

Im folgenden werden die inhaltskleinsten Tetraeder mit den Eigenschaften (a) bis (d) – unabhängig von Abschnitt 2. – mit Mitteln der analytischen Geometrie bestimmt. Außerdem wird in 3.1. eine anschauliche Beschreibung aller Tetraeder mit der Eigenschaft (a) mittels eines Kegelschnittbüschels gegeben.

3.1. Tetraeder mit zwei gegenüberliegenden rechten Keilwinkeln

Zunächst werden die Eigenschaften (b) bis (d) nicht beachtet und Tetraeder mit der Eigenschaft (a) untersucht.

In einem kartesischen Koordinatensystem sei

$$A \sim (0, 0, 0),\ B \sim (0, 0, 1),\ C \sim (c_1, 0, c_3),\ D \sim (0, d_2, d_3) \tag{11}$$

mit $c_1 > 0$ und $d_2 > 0$ (vgl. Abb. 6).

Jedes Tetraeder mit der Eigenschaft (a) ist ähnlich zu einem geeigneten derartigen Tetraeder $ABCD$, welches längs der Kante CD einen rechten Keilwinkel hat. Es soll untersucht werden, wie C und D in den Koordinatenebenen gewählt werden müssen, damit längs CD ein rechter Keilwinkel liegt.

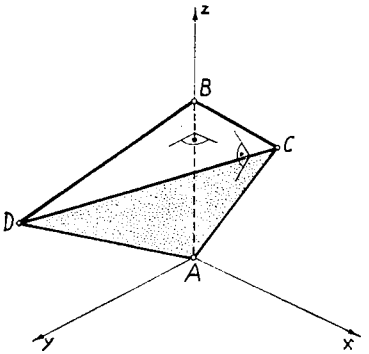

Abb. 6.

Längs CD liegt genau dann ein rechter Keilwinkel, wenn die Normalenvektoren der Ebenen ADC und BDC senkrecht sind. In Koordinaten bedeutet dies $d_2^2(1 + c_3(c_3 - 1)c_1^{-2}) + d_3(d_3 - 1) = 0$. Aus dieser Gleichung folgt: Bei festgehaltenem C ergeben genau die Punkte D der in der yz-Ebene liegenden Kurve

$$y^2(1 + \lambda) + z(z - 1) = 0 \tag{12}$$

eine Rechtwinkelkante CD, und andererseits führen die Punkte C, für die $c_3(c_3 - 1)c_1^{-2}$ konstant gleich λ ist, also Punkte der in der xz-Ebene liegenden Kurve

$$z(z - 1) = \lambda x^2 \tag{13}$$

auf dieselbe Kurve (12). Faßt man in (12) und (13) die Zahl λ als Parameter eines Büschels von Kurven zweiter Ordnung auf, so ergibt sich

LEMMA 1. *Das Tetraeder $ABCD$ gemäß* (11) *hat längs CD genau dann einen rechten Keilwinkel, wenn C und D auf zugeordneten Kurven der beiden Büschel von Kurven zweiter Ordnung*

$$(1 + \lambda')x^2 + z(z - 1) = 0,$$

$$(1 + \lambda)y^2 + z(z - 1) = 0$$

liegen, wobei die Büschel durch die Beziehung $\lambda' = -\lambda - 1$ aufeinander abgebildet sind.

Die Büschel (12) und (13) sind offenbar kongruent, sie haben A und B als (doppelte) reelle Grundpunkte und bestehen aus Ellipsen, einem Kreis, Hyperbeln und einem Parallelenpaar. Macht man die wegen (11) allein interessierenden Halbebenen $x>0$, $y=0$ und $y>0$, $x=0$ durch eine Drehung einer der Halbebenen um die z-Achse komplanar, so ergibt sich das in Abb. 7 gezeigte Bild. Gemäß Gleichung (12) gehören die Kurven folgendermaßen zu den Parameterwerten:

$\lambda < -1$: Hyperbeln
$\lambda = -1$: Parallelenpaar
$-1 < \lambda < 0$: Ellipsen außerhalb des Büschelkreises
$\lambda = 0$: Kreis
$\lambda < 0$: Ellipsen innerhalb des Büschelkreises.

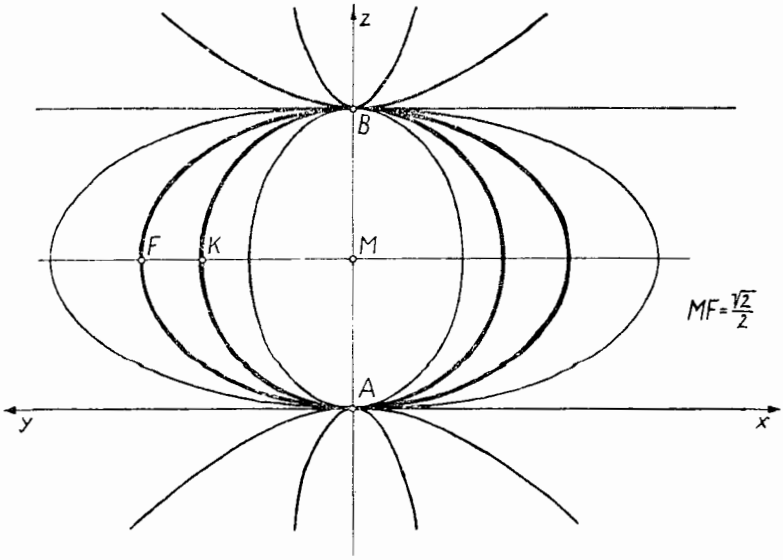

Abb. 7.

Die in Lemma 1 genannte Büschelabbildung (die in Abb. 7 als involutorische Selbstabbildung aufgefaßt werden kann) bildet die Hyperbeln auf die Ellipsen innerhalb des Büschelkreises, das Parallelenpaar auf den Büschelkreis und die Ellipsen außerhalb des Büschelkreises wieder auf Ellipsen außerhalb des Büschelkreises ab. Die Ellipse mit dem Parameterwert $\lambda = -1/2$ (Scheitel $F \sim (0, \sqrt{1/2}, 1/2)$) wird auf die dazu kongruente Ellipse abgebildet, sie möge Fixellipse genannt werden.

Hieraus ergibt sich auch

LEMMA 2. *Das Tetraeder $ABCD$ gemäß* (11) *mit der Eigenschaft* (a) *ist genau dann nicht stumpfwinklig, wenn C und D nicht außerhalb des Büschelparallelenpaares und nicht innerhalb des Büschelkreises des Büschels* (13) *bzw.*

(12) *liegen. Es ist ein Orthoschem, wenn einer der Punkte C und D auf dem Kreis und einer auf dem Parallelenpaar liegt.*

BEWEIS. Die zweite Behauptung ergibt sich aus dem Satz des Thales, denn liegt z. B. D auf dem Kreis, so wird AD senkrecht zu DB, und C muß auf der x-Achse oder der Parallelen dazu durch B liegen, also wird AC oder BC senkrecht zur Ebene ADB, d. h., $ABCD$ ist ein Orthoschem. Liegt aber D innerhalb des Büschelkreises, so liegt C auf einer Büschelhyperbel, so daß entweder ($c_3 < 1$) längs BD oder ($c_3 < 0$) längs AD der Keilwinkel stumpf ist. Liegt D außerhalb des Büschelkreises und innerhalb des Parallelenpaares, so C ebenfalls, und es gibt keine stumpfen Keilwinkel.

Ferner gilt

LEMMA 3. *Hat ein Tetraeder die Eigenschaft* (a), *so ist eine der vier Kanten, die von den beiden gegenüberliegenden Rechtwinkelkanten verschieden sind, kürzeste Kante des Tetraeders.*

BEWEIS. Das betrachtete Tetraeder habe gemäß (11) die Eckpunkte A, B, C, D und Rechtwinkelkanten längs AB und CD. Liegt D auf einer Büschelhyperbel, ohne Beschränkung der Allgemeinheit sei $d_3 < 1$, so liegt C im Kreis, also ist $BC < AB$; außerdem ist das Dreieck DBC stumpfwinklig mit stumpfem Winkel bei B, so daß $BC < DC$ ist. Liegt aber D auf oder zwischen den Büschelparallelen, so C ebenfalls; das Tetraeder ist nicht stumpfwinklig. Wäre in diesem Fall die Rechtwinkelkante AB eine kürzeste, so wären AC und BC beide größer oder gleich 1, also läge C außerhalb, folglich D innerhalb der Fixellipse, was aber $AD < 1$ zur Folge hat, so daß AB doch nicht kürzeste sein kann. Wäre die Rechtwinkelkante CD kürzeste, so könnten durch Ähnlichkeitstransformation CD in AB und AB in CD überführt werden und dieselben Überlegungen angestellt werden.

3.2. Umformulierung der Aufgabenstellung

Statt der Eigenschaft (c) wird zunächst die folgende Eigenschaft (c′) in Betracht gezogen, dabei sei b eine beliebige positive reelle Zahl:

(c′) Die kürzeste Kantenlänge ist b.

Es sei $M^{R,b}$ die Menge aller Tetraeder mit den Eigenschaften (a), (b), (c′), (d), und es soll zunächst das minimale Volumen der Tetraeder aus der Menge $M^{R,b}$ bestimmt werden. M sei die Menge der Tetraeder $ABCD$ gemäß (11) mit den Eigenschaften (a) und (b), also den Rechtwinkelkanten AB und CD, $AB = 1$. Zu jedem Tetraeder $T \in M^{R,b}$ existiert eine Ähnlichkeitsabbildung φ, so daß $\varphi(T) \in M$ ist. Ist r die kürzeste Kantenlänge von $\varphi(T)$, so ist wegen (c′) der Längenveränderungsfaktor von φ gleich r/b, und $\varphi(T)$ hat wegen (d) den Umkugelradius Rr/b. Es gilt dann für das zu bestimmende minimale Volumen

$$\min\{\mathrm{vol}\, T : T \in M^{R,b}\} = \min\{b^3 r^{-3}\, \mathrm{vol}\, \varphi(T) : T \in M^{R,b}\} = \tag{14}$$
$$= b^3 \min\{k^{-3}\, \mathrm{vol}\, X : X \in M \text{ und } X \text{ hat den Umkugelradius } Rk/b\},$$
wobei $k = \min\{AC, BC, AD, BD\}$ ist.

Folglich ist für die Funktion k^{-3} vol X in einem bestimmten Bereich das Minimum zu ermitteln.

3.3. Rechnerische Darstellung der zu minimierenden Funktion

Zunächst gilt

LEMMA 4. *ABCD sei ein Tetraeder gemäß* (11). *Sind von ihm die Punkte A, B, D und der Umkugelradius u gegeben, so liegt C auf dem Kreis*

$$x^2 - 2\sqrt{u^2 - \frac{1}{4} - \left(\frac{d_2^2 + d_3(d_3-1)}{2d_2}\right)^2}\cdot x + z(z-1) = 0. \tag{15}$$

BEWEIS. Der Umkugelmittelpunkt hat als erste Koordinate x_0 die erste Koordinate des Umkreismittelpunktes von ABC, man berechnet leicht $x_0 = (2c_1)^{-1}(c_1^2 + c_3(c_3-1))$. Analog berechnet man für die zweite Koordinate y_0 den Wert $(2d_2)^{-1}(d_2^2 + d_3(d_3-1))$, die dritte Koordinate ist 1/2. Da A auf der Umkugel liegt, ergibt die Umkugelgleichung die Beziehung $x_0^2 + y_0^2 + 1/4 = u^2$, hieraus folgt, wenn man $x = c_1$, $z = c_2$ setzt, die Gleichung (15).

Im folgenden soll vorausgesetzt werden, daß im Tetraeder $ABCD$ gemäß (11) die Kante AD eine kürzeste ist. Wegen Lemma 3 ist das keine Einschränkung der Allgemeinheit. Wegen (14) sind dann nur solche Tetraeder von Interesse, die den Umkugelradius $R \cdot AD/b$ haben. Es gilt

LEMMA 5. *Das Tetraeder ABCD gemäß* (11) *mit* $d_3(d_3-1) \neq 0$ *hat genau dann die Eigenschaft* (a) *und den Umkugelradius* $R \cdot AD/b$, *wenn C in der Weise von D abhängt, daß für die Koordinaten* $(c_1, 0, c_3)$ *von C gilt:*

$$c_1 = \frac{d_3}{d_3(1-d_3)}\sqrt{(4R^2b^{-2})d_2^4 + ((4R^2b^{-2}-2)d_3^2 + 2d_3 - 1)d_2^2 - d_3^2(1-d_3)^2}, \tag{16}$$

$$c_3 = \frac{1}{2}\left(1 \pm \sqrt{1 + 4c_1^2(d_3(1-d_3)d_2^{-2} - 1)}\right) \quad \text{mit } c_1 \text{ aus (16).} \tag{17}$$

BEWEIS. Es seien A, B, D gegeben. D liegt auf einer Büschelkurve (12) mit einem Parameterwert λ. Es ist $d_2^2(1+\lambda) + d_3(d_3-1) = 0$, also $\lambda + 1 = d_3(1-d_3)d_2^{-2}$.

$ABCD$ hat wegen Lemma 1 genau dann die Eigenschaft (a), wenn C auf der Kurve (13) liegt. Wegen Lemma 4 hat es den Umkugelradius $R \cdot AD/b$ genau dann, wenn C auf dem Kreis $x^2 - 2wx + z(z-1) = 0$ liegt, wo w die in (15) vorkommende Wurzel mit $u^2 = R^2b^{-2}(d_2^2 + d_3^2)$ ist. Aus beidem folgt $c_1 = 2w(1+\lambda)^{-1}$, und das ergibt Gleichung (16). Aus Gleichung (13) ergibt sich eine quadratische Gleichung für c_3 mit der Lösung (17), wenn man für λ den angegebenen von d_2 und d_3 abhängenden Wert nimmt und $x = c_1$ setzt.

ERGÄNZUNG. Ist $d_3(d_3-1) = 0$, d. h., liegt D auf den Büschelparallelen, so ergibt jeder Punkt C des Büschelkreises ein Tetraeder mit dem Umkugelradius $R \cdot AD/b$, wobei $d_2 = (4R^2b^{-2} - 1)^{-1/2}$ sein muß. Denn es ist dann der oben genannte Parameterwert gleich -1, hieraus folgt $w = 0$, und das ergibt für d_2 diesen Wert.

Statt der nach (14) zu minimierenden Funktion k^{-3} vol $X = \frac{1}{6} c_1 d_2 k^{-3}$ wird im folgenden das Quadrat ihres Sechsfachen, also die Funktion

$$f = k^{-6} c_1^2 d_2^2 \tag{18}$$

betrachtet und minimiert. Wird AD als kürzeste Kante vorausgesetzt, so ist

$$f = c_1^2 d_2^2 (d_2^2 + d_3^2)^{-3}. \tag{19}$$

Führt man $s := k^2 = d_2^2 + d_3^2$ als neue Variable ein, so läßt sich f als Funktion von s und d_3 darstellen, falls $d_3(d_3 - 1) \neq 0$ ist. Zunächst ergibt sich aus (16) die Gleichung

$$c_1^2 = s(s - d_3^2) d_3^{-2} (1 - d_3)^{-2} \big((4R^2 b^{-2} - 1)s - 4R^2 b^{-2} d_3^2 + 2d_3 - 1\big), \tag{20}$$

und hieraus folgt mit (19) die Gleichung

$$f = \frac{(s - d_3^2)^2 \big((4R^2 b^{-2} - 1)s - 4R^2 b^{-2} d_3^2 + 2d_3 - 1\big)}{d^2 (1 - d_3)^2 s^2}. \tag{21}$$

Für die partielle Ableitung nach s ergibt sich

$$\begin{aligned} \frac{\partial f}{\partial s} = \frac{s - d_3^2}{d_3^2 (1 - d_3)^2} \big(& (4R^2 b^{-2} - 1)s^2 + (4R^2 b^{-2} - 1) d_3^2 s + \\ & + 2d_3^2 (-4R^2 b^{-2} d_3^2 + 2d_3 - 1) \big). \end{aligned} \tag{22}$$

3.4. Abgrenzung des Bereiches für die zu minimierende Funktion f

Die Festlegung der nach (14) zu betrachtenden Tetraeder soll mittels des Punktes D erfolgen. Zu jedem (geeigneten) Punkt D gibt es dann nach Lemma 5 genau zwei (eventuell übereinstimmende) Tetraeder X aus M mit dem Umkugelradius Rk/b, die gemäß (11) D als einen Eckpunkt haben. Wegen Eigenschaft (b) und Lemma 2 liegt D nicht außerhalb der Büschelparallelen und nicht innerhalb des Büschelkreises. Da AD als kürzeste Kante vorausgesetzt wird, müssen die drei Ungleichungen

$$AD \leqq BD, \tag{23}$$

$$AD \leqq AC, \tag{24}$$

$$AD \leqq BC \tag{25}$$

gelten. (23) ist gleichbedeutend mit $d_3 \leqq 1/2$. Aus (24) und (25) folgt: Ist S der von A und B verschiedene in der xz-Ebene liegende Scheitel der dem Punkt D gemäß Lemma 1 zugeordneten Büschelellipse, so muß $AS \geqq AD$ sein, denn andernfalls kann weder (24) noch (25) gelten. Formuliert man diese Ungleichung unter Benutzung von (12) und (13) in Koordinaten, so ergibt sich, da S durch D eindeutig bestimmt ist, eine Ungleichung für d_2 und d_3, nämlich

$$\left(d_2^2 + d_3^2 - \frac{1}{2} d_3\right)\left(d_2^2 + d_3^2 - \frac{1}{2} d_3 - \frac{1}{2}\right) \leqq 0.$$

Hierdurch ist für D ein konzentrischer Kreisring $\left(\text{Mittelpunkt } y = 0, z = \frac{1}{4}\right)$ festgelegt, von dem der innere Kreis innerhalb des Büschelkreises liegt, so daß nur der äußere Kreis

$$d_2^2 = -d_3^2 + \frac{1}{2}d_3 + \frac{1}{2} \tag{26}$$

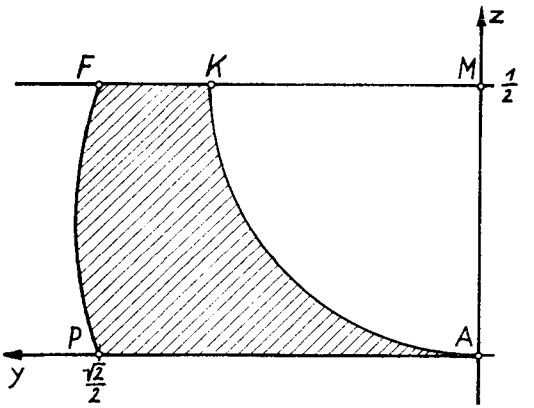

Abb. 8.

eine Einschränkung für D liefert. Insgesamt sind also nur Punkte D des in Abb. 8 dargestellten Bereiches zu betrachten, er wird durch die auf den Büschelparallelen liegenden Strecken FK und AP, den auf dem Büschelkreis liegenden Bogen KA und den auf dem Kreis (26) liegenden Bogen PF begrenzt. Punkte D außerhalb dieses Bereiches ergeben Tetraeder, bei denen AD nicht kürzeste Kante ist oder die stumpfwinklig sind. Andererseits erfüllt nicht jeder Punkt D dieses Bereiches mit dem nach Lemma 5 zu ihm gehörenden Punkt C die Ungleichungen (24) und (25). Zur Umformulierung dieser Ungleichungen in Koordinaten ist zu berücksichtigen, daß die beiden nach (17) zu festem D gehörenden Punkte C_1 und C_2 (die sich durch unterschiedliche Wahl des Vorzeichens in (17) ergeben) symmetrisch zur Mittelsenkrechten $z = \frac{1}{2}$ von AB liegen, daß also die Längen von AC_1 und BC_2 gleich sind, so daß (24) und (25) beide erfüllt sind, wenn man nur (24) mit demjenigen Punkt C fordert, für den in (17) das untere Vorzeichen genommen wurde. Die Ungleichung $AD^2 \leqq AC^2$ ergibt unter Verwendung der Formeln (16) und (17) nach einer hier nicht ausgeführten Zwischenrechnung die gleichbedeutende Ungleichung

$$(4R^2b^{-2} - 1)d_2^4 + \left((4R^2b^{-2} - 1)d_3^2 + 2d_3 - 1\right)d_2^2 + 2d_3^2(d_3 - 1) \geqq 0. \tag{27}$$

Der Rand des durch (27) gegebenen Gebietes der yz-Ebene schneidet die Geraden $z = \text{const}$ in zwei Punkten, deren y-Koordinate durch die aus (27) für den Fall der Gleichheit sich ergebende quadratische Gleichung in d_2^2 bestimmt ist. Man rechnet leicht nach, daß von diesen beiden Punkten höchstens derjenige mit der größeren y-Koordinate in den relevanten Bereich fallen kann, daß also von der quadratischen Gleichung nur die Lösung

$$d_2^2 = \frac{1}{2(4R^2b^{-2}-1)}\Big(-(4R^2b^{-2}-1)d_3^2-2d_3+1+$$

$$(28)\qquad +\sqrt{(4R^2b^{-2}-1)^2d_3^4-4(4R^2b^{-2}-1)d_3^3+2(12R^2b^{-2}-1)d_3^2-4d_3+1}\Big)$$

zu berücksichtigen ist. Der Bereich **B**, in dem das Minimum der Funktion f zu bestimmen ist, ergibt sich als Durchschnitt des oben beschriebenen Bereiches (vgl. Abb. 8) mit dem Bereich (27). Die durch (28) gegebene Kurve $d_2 = d_2(d_3)$ ist, soweit sie in dem in Abb. 8 beschriebenen Bereich verläuft, Teil des Randes von **B**. Abb. 9 zeigt zwei charakteristische Beispiele des Verlaufs dieser Kurve und der Gestalt von **B**. Hinsichtlich der Begrenzung von **B** gibt es genau zwei Fälle: Entweder besteht die Begrenzung aus einer Strecke auf der Geraden $z = \frac{1}{2}$, einem Kreisbogenstück des Büschelkreises, einem Stück der Kurve (28), einer Strecke auf der y-Achse und einem Stück des Kreises (26) (Abb. 9a), oder sie besteht aus einer Strecke auf der Geraden $z = \frac{1}{2}$, einem Stück der Kurve (28) und einem Stück des Kreises (26) (Abb. 9b). Der erste Fall tritt bei $4^2R^{-2} > 3$ ein, der zweite bei $4R^2b^{-2} \leqq 3$; das ergibt sich, wenn man in (28) $d_3 = \frac{1}{2}$ bzw. 0 setzt und so die Lage der Endpunkte des relevanten Teils der Kurve (28) bestimmt.

Hinsichtlich der unteren Grenze für R gilt

LEMMA 6. *Der kleinstmögliche Umkugelradius eines Tetraeders mit den Eigenschaften* (a), (b) (c′) *ist* $\frac{b}{2}\sqrt{\frac{5}{3}}$.

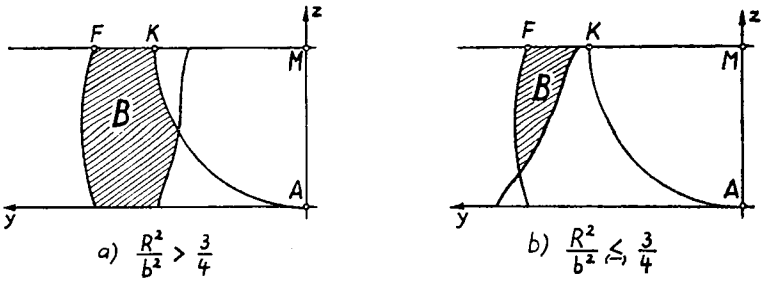

Abb. 9.

BEWEIS. Zu jedem Tetraeder mit den Eigenschaften (a), (b), (c′) existiert ein ähnliches in der Menge M, für das der Eckpunkt D im Bereich **B** liegt. Für $D \in \mathbf{B}$ gilt die Ungleichung (27), also ist

$$4R^2b^{-2} - 1 \geqq \frac{(1-2d_3)d_2^2 + 2d_3^2(1-d_3)}{(d_2^2+d_3^2)d_2^2},$$

wegen (26) ist $d_2^2 + d_3^2 \leqq \frac{3}{4}$, also

$$4R^2b^{-2} - 1 \geqq \frac{4}{3} \cdot \frac{(1-2d_3)d_2^2 + 2d_3^2(1-d_3)}{d_2^2}.$$

Ferner ist

$$\frac{(1-2d_3)d_2^2 + 2d_3^2(1-d_3)}{d_2^2} \geqq \frac{1}{2},$$

denn diese Ungleichung ist gleichbedeutend mit $(1-4d_3)d_2^2 + 4d_3^2(1-d_3) \geqq 0$, und dies ist für $0 \leqq d_3 \leqq \frac{1}{4}$ offensichtlich erfüllt, während für $\frac{1}{4} < d_3 \leqq \frac{1}{2}$ auf Grund von (26) die Abschätzung

$$(1-4d_3)d_2^2 + 4d_3^2(1-d_3) \geqq (1-4d_3)\left(\frac{1}{2} + \frac{1}{2}d_3 - d_3^2\right) + 4d_3^2(1-d_3) =$$

$$= (1-d_3)\left(\frac{1}{2} - d_3\right) \geqq 0$$

möglich ist. Folglich ist $4R^2b^{-2} - 1 \geqq \frac{4}{3} \cdot \frac{1}{2}$, also $4R^2b^{-2} \geqq \frac{5}{3}$. Der Wert $\frac{5}{3}$ wird bei $d_2^2 = d_3 = \frac{1}{2}$ auch tatsächlich angenommen.

3.5. Bestimmung der Lage des Minimums von f

Es soll gezeigt werden, daß das Minimum von f nicht im Inneren von B, sondern auf dem Rande angenommen wird, und zwar auf dem in Abb. 9 dargestellten "rechten" Rand, genauer:

LEMMA 7. *Im Bereich* **B** *ist* (19) *für konstantes* d_3 $(\neq 0)$ *eine monoton steigende Funktion von* d_2.

BEWEIS. a) $4R^2b^{-2} \geqq 3$.

Im Bereich **B** ist $s > d_2 > d_2^2$ und $d_3 \leqq \frac{1}{2}$. Setzt man dies in (22) ein, ergibt sich

$$\frac{\partial f}{\partial s} \geqq \frac{1-d_3^2}{1-d_3}(8R^2b^{-2}d_3 + 4R^2b^{-2} - 3) > 0.$$

Auf Grund der Definition von s unterscheiden sich $\frac{\partial f}{\partial s}$ und $\frac{\partial f}{\partial d_2}$ nur um einen positiven Faktor, also gilt die Behauptung.

b) $4R^2b^{-2}<3$.

In (22) ist der Vorfaktor positiv. Ferner entnimmt man (22), daß $\frac{\partial f}{\partial s}$ für festes d_3 höchstens eine positive Nullstelle s_0 hat, für $s>s_0$ ist $\frac{\partial f}{\partial s}$ positiv. Es wird gezeigt, daß s_0 kleiner ist als der sich aus (28) ergebende Wert s_1 für s des zum selben festen d_3 gehörenden ("rechten") Randpunktes von **B**. Da sich $\frac{\partial f}{\partial s}$ und $\frac{\partial f}{\partial d_2}$ nur um einen positiven Faktor unterscheiden, folgt die Positivität von $\frac{\partial f}{\partial d_2}$ in **B**, also die Behauptung. Zur Abkürzung werde $p:=4R^2b^{-2}-1$ und $y:=d_3$ gesetzt. Aus (22) folgt

$$s_0=\frac{1}{2p}\left(-py^2+\sqrt{9p^2y^4+8py^4-16py^3+8py^2}\right).$$

Ersetzt man in (27) bzw. (28) d_2^2 durch $s-d_3^2$ bzw. $s-y^2$, so ergibt sich

$$s_1=\frac{1}{2p}\left(py^2-2y+1+\sqrt{(py^2-2y+1)^2+4py^2}\right).$$

Es ist zu zeigen, daß $s_0<s_1$, also

$$2py^2-2y+1+\sqrt{(py^2-2y+1)^2+4py}>\sqrt{9p^2y^4+8py^4-16py^3+8y^2} \tag{29}$$

$$\text{für } \frac{2}{3}\leqq p<2 \text{ und } 0<y\leqq\frac{1}{2}$$

gilt. Es ist

$$2p^2-2y+1=\left(\sqrt{2p}y-\frac{1}{\sqrt{2p}}\right)^2+\frac{1}{2p}(2p-1)>0,$$

so daß (29) quadriert werden kann, es ergibt sich

$$\begin{aligned}&-2(p^2+2p)y^4+2py^3+(p+4)y^2-4y+1+\\&+(2p^2-2y+1)\sqrt{(py^2-2y+1)^2+4py^2}>0\end{aligned} \tag{30}$$

als zu beweisende Ungleichung. Für $0<y\leqq\frac{1}{4}$ ergeben die ersten fünf Summanden etwas Positives, so daß für diese y die Ungleichung und erst recht

(29) gilt. Die linke Seite von (30) wird verkleinert, wenn man die Wurzel durch $2\sqrt{py}$ ersetzt. Es genügt demzufolge zu zeigen

$$(31)\qquad \begin{gathered} -2(p^2+2p)y^4+(2p+4p\sqrt{p})y^3+(p-4\sqrt{p}+4)y^2+(2\sqrt{p}-4)y+1>0 \\ \text{für } \frac{1}{4}<y\leqq\frac{1}{2} \text{ und } \frac{2}{3}\leqq p<2. \end{gathered}$$

Für konstantes y ist die linke Seite eine monoton steigende Funktion von p, denn die partielle Ableitung nach p ist

$$\frac{1}{2p}(y^3p(-8y\sqrt{p}+12)+(4y^3(-2y+1)+2y^2)\sqrt{p}+2y(-2y+1)),$$

sie ist wegen $-8y\sqrt{p}+12\geqq-8\cdot\frac{1}{2}\sqrt{2}+12>0$ positiv. Es genügt also, (31) für $p = 2/3$ zu beweisen, d.h. die Ungleichung

$$(32)\qquad -34\sqrt{3}y^4+(12\sqrt{3}+24\sqrt{2})y^3+(42\sqrt{3}-36\sqrt{2})y^2+(18\sqrt{2}-36\sqrt{3})y+9\sqrt{3}>0.$$

Die Ableitung der linken Seite von (32) ist kleiner als

$$(33)\qquad -235y^3+165y^2+44y-36,$$

aus der Entwicklung von (33) an der Stelle $y = 1/2$ gewinnt man die Darstellung von (33) in der Form

$$-\frac{17}{8}-\frac{131}{4}(1/2-y)-\frac{1}{2}(1/2-y)^2(374-235(1-2y)),$$

aus der man ablesen kann, daß (33) für $1/4<y\leqq 1/2$ negativ ist. Also ist die linke Seite von (32) monoton fallend. Da sie an der Stelle $y = 1/2$ den positiven Wert $\frac{7}{8}\sqrt{3}+3\sqrt{2}$ hat, ist die Ungleichung (32) richtig, und die Behauptung ist bewiesen.

In Lemma 7 ist wie in (21) der Fall $d_3 = 0$ ausgeschlossen worden. Dieser Fall kann aber außer acht bleiben, denn es muß dann C auf dem Büschelkreis liegen, und aus Symmetriegründen erhält man diese Tetraeder auch schon, wenn D auf dem Büschelkreis liegt.

Auf Grund von Lemma 7 und der Begrenzung von **B** gibt es für die Stellen minimaler Funktionswerte zwei Möglichkeiten:

1. f wird in einem Punkt der Kurve (28) minimal,
2. f wird in einem Punkt des Büschelkreises minimal.

Der Fall 1. tritt ein, wenn $4R^2b^{-2}\leqq 3$ ist, bei $4R^2b^{-2}>3$ sind zunächst beide Fälle denkbar (vgl. Abb. 9).

Zur Bestimmung des minimalen Funktionswertes und der zugehörigen Tetraeder genügt es nun in Falle 1., Punkte D mit $d_3 = 1/2$ zu betrachten. Zum Beweis dieser Behauptung sei D_0 ein beliebiger Punkt auf der Kurve (28) mit minimalem Funktionswert. Nach Lemma 5 gibt es dazu zwei (möglicherweise gleiche) Punkte C_1, C_2, die symmetrisch zur Geraden $z = 1/2$ liegen, es möge C_1 die kleinere z-Koordinate haben. Da die Koordinaten von D_0 die Gleichung (28) erfüllen, sind AD_0, AC_1 und BC_2 gleich lang. Das Tetraeder $T_1 := ABC_1D_0$ hat dann zwei zusammenhängende kürzeste Kanten, nämlich AB und AC_1. Bei der Ähnlichkeitsabbildung β, die D_0C_1 auf AB, A in die Halbebene $x = 0, y > 0$, B in die Halbebene $y = 0, x > 0$ abbildet, ergibt sich (Abb. 10) ein Tetraeder T_1' aus der Menge M, dessen einer

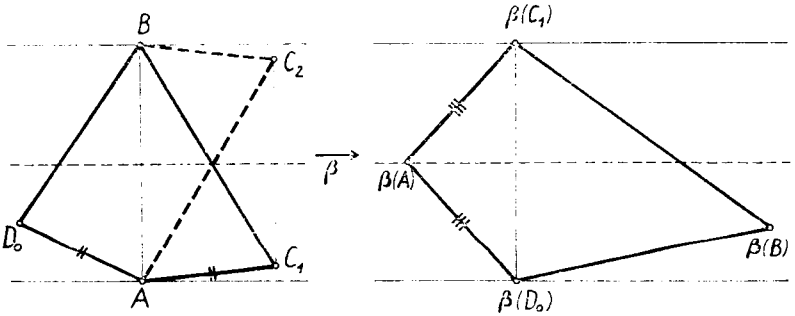

Abb. 10.

Eckpunkt $\beta(A)$ auf der Geraden $z = 1/2$ liegt, das den Umkugelradius $b^{-1}R$ $\beta(A)$ $\beta(D_0)$ hat, und der Funktionswert $f(T_1')$ ist derselbe wie bei T_1, weil die Funktion $f = (k^{-3} \cdot \mathrm{vol}\, X)^2$ offenbar invariant gegen Ähnlichkeitsabbildung ist. Es ergibt sich also ein zum Minimaltetraeder T_1 ähnliches Tetraeder T_1', wenn man die Minimaltetraeder nur längs $z = 1/2$ bestimmt, so daß T_1 selbst nicht explizit bestimmt zu werden braucht. Das Tetraeder $T_2 := ABC_2D_0$ kann ebenfalls aus T_1' gewonnen werden. Denn es ergibt sich T_2 aus T_1, indem man den Punkt C_1 an $z = 1/2$ spiegelt und die anderen Eckpunkte festläßt, so daß man es aus T_1' herleiten kann, indem man auf T_1' die Abbildung β^{-1} anwendet und anschließend den in der Ebene $y = 0$ liegenden Punkt C_1 an der Geraden $z = 1/2$ spiegelt. (Es kann natürlich auch gleich der Punkt $\beta(C_1)$ an der in der Ebene $\beta(ABC_1)$ liegenden Mittelsenkrechten zur Strecke $\beta(A)$ $\beta(B)$ gespiegelt werden, das neue Tetraeder hat dann aber zum Koordinatensystem nicht mehr die Lage gemäß (11); vgl. den Schluß von 3.6. Die Tetraeder T_1 und T_1' gehören zu den in 2. betrachteten Vierecken $S_1^{(1)}$, und T_2 gehört zu den Vierecken $T_l^{(1)}$, vgl. Satz 2.) Damit ist bewiesen, daß man alle minimalen Tetraeder mit Eckpunkt D auf der Kurve (28) aus den minimalen Tetraedern mit $d_3 = 1/2$ herleiten kann.

3.6. Bestimmung der minimalen Tetraeder mit kürzester Kantenlänge b

Erster Fall: $4R^2b^{-2}<3$.

Für das minimale Tetraeder $ABCD$ ist zunächst D als Schnittpunkt der Kurve (28) mit der Geraden $z = 1/2$ zu bestimmen. Setzt man in (28) $d_3 = 1/2$, ergibt sich

$$d_2^2 = \frac{1}{8}\left(-1+\sqrt{\frac{4R^2+15b^2}{4R^2-b^2}}\right). \tag{34}$$

Damit ist D bestimmt. Aus (16) und (17) berechnet man für die Koordinaten des zugehörigen Punktes C (wobei man sich in (17) aus Symmetriegründen z. B. auf das untere Vorzeichen beschränken kann)

$$c_1^2 = \frac{1}{2}\left(-1+\sqrt{\frac{4R^2+15b^2}{4R^2-b^2}}\right)-\frac{b^2}{4R^2-b^2}, \tag{35}$$

$$c_3 = \frac{1}{2}\left(1-\sqrt{\frac{13}{2}-\frac{5}{2}\sqrt{\frac{4R^2+15b^2}{4R^2-b^2}}+\frac{4b^2}{4R^2-b^2}}\right). \tag{36}$$

Für die kürzeste Kantenlänge AD bzw. BD ergibt sich

$$k = \sqrt{d_2^2+\frac{1}{4}} = \sqrt{\frac{1}{8}\left(1+\sqrt{\frac{4R^2+15b^2}{4R^2-b^2}}\right)}, \tag{37}$$

dies ist, da D auf der Kurve (28) liegt, zugleich die Länge von AC. Die Kante BC ist für $4R^2b^{-2}>5/3$ länger als AD, für $4R^2b^{-2} = 5/3$ haben die vier Kanten AD, BD, AC, BC alle die Länge $\frac{\sqrt{3}}{2}$. Für die Kantenlänge CD ergibt sich $CD^2 = c_1^2+d_2^2+\left(c_3-\frac{1}{2}\right)^2 = 1$, die beiden Rechtwinkelkanten sind also gleich lang. Die am Schluß von 3.5. erwähnten Tetraeder T_1 und T_2 sind im vorliegenden Fall kongruent (Abb. 11). Das so erhaltene Tetraeder gehört zur Menge M (vgl. (14)). Das minimale Volumen des entsprechenden Tetraeders aus der Menge $M^{R,b}$ bekommt man durch Rücktransformation. Der Längenveränderungsfaktor ist dabei b/k, wobei k die kürzeste Kantenlänge (37) von $ABCD$ ist, und das Volumen $\frac{1}{6}c_1d_2$ von $ABCD$ ist mit $\left(\frac{b}{k}\right)^3$ zu multiplizieren; c_1 und d_2 sind dabei aus (34) und (35) zu nehmen. Es ergibt sich

$$\min_{T\in M^{R,b}} \operatorname{vol} T = \tag{38}$$

$$= \frac{1}{24\sqrt{2}}\sqrt{\sqrt{4R^2-b^2}\sqrt{4R^2+15b^2}-4R^2-b^2}\cdot$$
$$\cdot\left(4R^2+7b^2-\sqrt{4R^2-b^2}\sqrt{R^2+15b^2}\right).$$

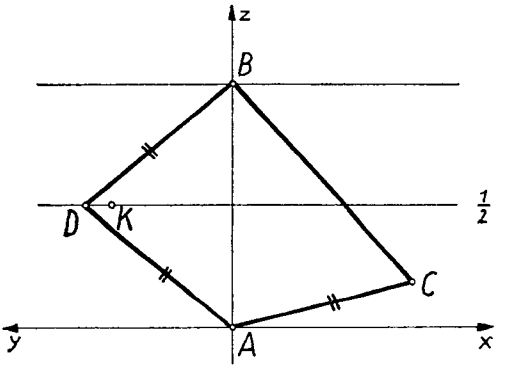

Abb. 11.

Zweiter Fall: $4R^2b^{-2} \geqq 3$.

Auf Grund von Lemma 7 kann das Minimum von f auf Punkten des Büschelkreises oder der Kurve (28) angenommen werden. Um den Verlauf von f längs des Büschelkreises zu untersuchen, ist in (21) $d_3 = s$ zu setzen (denn der Büschelkreis hat die Gleichung $y^2+z(z-1) = 0$, also gilt $s = d_2^2+d_3^2 = d_3$), das ergibt

$$f(s) = -4R^2b^{-2}+(4R^2b^{-2}+1)s^{-1}-s^{-2},$$

und in (27) ist $d_3 = s$ sowie $d_2^2 = d_3-d_3^2 = s(1-s)$ zu setzen, das ergibt $s(1-s)((4R^2b^{-2}-1)s-1) \geqq 0$, wegen $1-s>0$ also $s \geqq (4R^2b^{-2}-1)^{-1}$, so daß f für Werte s aus dem Intervall $\left[(4R^2b^{-2}-1)^{-1}, \frac{1}{2}\right]$ zu untersuchen ist. Man rechnet leicht nach, daß

$$f\left(\frac{1}{4R^2b^{-2}-1}\right) = f\left(\frac{1}{2}\right) = 4R^2b^{-2}-2$$

ist, und daß im Inneren des Intervalls die Funktionswerte größer sind. Das Minimum wird also auf dem Schnittpunkt des Büschelkreises mit der Geraden $z = 1/2$ bzw. mit der Kurve (28) angenommen; es genügt aber, die Gerade $z = 1/2$ zu betrachten. Dieser Schnittpunkt D hat die Koordinaten $\left(0, \frac{1}{2}, \frac{1}{2}\right)$. Für den zugehörigen Punkt C, der auf den Büschelparallelen liegen muß, ergibt sich aus (20), wo man $s = d_3 = 1/2$ zu setzen hat, $c_1 = b^{-1}\sqrt{2R^2-b^2}$, und es ist c_3 z. B. gleich 0. Das Tetraeder $ABCD$ ist demnach ein Orthoschem mit dem total-orthogonalen Kantenzug $BDAC$ (Abb. 12). Die kürzesten Kanten sind AD und BD mit der Länge $\sqrt{\frac{1}{2}}$. Die Kante AC ist für $4R^2b^{-2} > 3$ länger, für $4R^2b^{-2} = 3$ ist sie ebenfalls kürzeste Kante. Die Rechtwinkel-

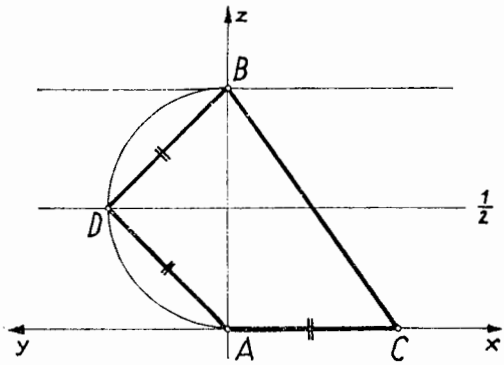

Abb. 11.

kante DC hat die Länge $\sqrt{\frac{4R^2-b^2}{2b^2}}$, sie ist für $4R^2b^{-2}>3$ länger als die Rechtwinkelkante AB, für $4R^2b^{-2}=3$ gleichlang mit AB.

Das zweite minimale, zu $ABCD$ nur im Falle $4R^2b^{-2}=3$ kongruente Tetraeder $AB'CD$ ergibt sich durch Übergang von B zu B' mittels Spiegelung an der in der Ebene BCD gelegenen Mittelsenkrechten zu DC. Es ist ebenfalls ein Orthoschem, es hat den total-orthogonalen Kantenzug $DACB'$ und die kürzesten Kanten AD und CB'. Die Rechtwinkelkanten AB' und DC sind gleich lang, wie sich sofort aus der Kongruenz der rechtwinkligen Dreiecke DAC und ACB' ergibt. Das minimale Volumen der entsprechenden Tetraeder aus $M^{R,b}$ berechnet man aus dem Volumen $\frac{1}{6}c_1d_2=\frac{1}{12b}\sqrt{2R^2-b^2}$ von $ABCD$ durch Multiplikation mit $(b\sqrt{2})^3$, es ergibt sich

$$\min_{T\in M^{R,b}} \operatorname{vol} T = \frac{1}{6}b^2\sqrt{4R^2-2b^2}. \tag{39}$$

3.7. Bestimmung der minimalen Tetraeder mit kürzester Kantenlänge 2

Zur Bestimmung des minimalen Volumens von Tetraedern mit den Eigenschaften (a), (b), (c), (d) ist $\min\{\operatorname{vol} X : X\in M^{R,b} \text{ und } b\geq 2\}$ zu bestimmen. Es gilt

LEMMA 8. *Für konstantes R ist* vol X *mit $X\in M^{R,b}$ eine monoton wachsende Funktion von b.*

BEWEIS. Es werden die partiellen Ableitungen nach b der in (38) und (39) angegebenen Volumenfunktionen betrachtet.

Erster Fall: $5/3\leq 4R^2b^{-2}<3$, Formel (38).

Es ist nach (38)

$$(\mathrm{vol}\,T)^2 = \frac{1}{1152}(W - 4R^2 - b^2)(4R^2 + 7b^2 - W)^2$$

mit

$$W = \sqrt{(4R^2 - b^2)(4R^2 + 15b^2)} = \sqrt{16R^2 + 56R^2b^2 - 15b^4}.$$

Hieraus folgt

$$\frac{\partial}{\partial b}(\mathrm{vol}\,T)^2 = \frac{16b}{384W}(4R^2 + 7b^2 - W)(24R^2 + 38R^2b^2 - 15b^4 + (b^2 - 6R^2)W).$$

Auf der rechten Seite dieser Gleichung sind die beiden ersten Faktoren offensichtlich positiv, und aus der Voraussetzung

$$b^2 < \frac{12}{5}R^2 \text{ folgt } 16b^4(12R^2 - 5b^2)(10R^2 - 3b^2) > 0,$$

und hieraus

$$(24R^4 + 38R^2b^2 - 15b^4)^2 > (b^2 - 6R^2)^2(16R^4 + 56R^2b^2 - 15b^4),$$

so daß auch der dritte Faktor positiv wird, nur für $b^2 = \frac{12}{5}R^2$ ist er 0. Also ist vol T in b monoton wachsend.

Zweiter Fall: $4R^2b^{-2} \geqq 3$, Formel (39).

Es ist nach (39)

$$\frac{\partial}{\partial b}(\mathrm{vol}\,T)^2 = \frac{b^3}{9}(4R^2 - 3b^2) \geqq 0$$

für $4R^2b^{-2} \geqq 3$, also ist vol T in b monoton wachsend.

Auf Grund von Lemma 8 sind die inhaltskleinsten Tetraeder mit den Eigenschaften (a), (b), (c), (d) diejenigen in 3.6. ermittelten minimalen Tetraeder, bei denen $b = 2$ ist. Hieraus ergibt sich folgender

HAUPTSATZ. (I) *Für* $R = \sqrt{\frac{5}{3}}$ *gibt es bis auf Kongruenz genau ein Tetraeder mit den Eigenschaften* (a) *bis* (d) *und minimalem Volumen. Es hat zwei senkrechte windschiefe Rechtwinkelkanten der Länge* $\frac{4\sqrt{3}}{3}$, *die übrigen vier Kanten haben alle die Länge* 2 (Abb. 13), *und das Volumen ist* $\frac{16\sqrt{3}}{27}$.

(II) Zu jedem R *mit* $\sqrt{\frac{5}{3}} < R < \sqrt{3}$ *gibt es bis auf Kongruenz genau ein Tetraeder mit den Eigenschaften* (a) *bis* (d) *und minimalem Volumen. Es hat drei Kanten der Länge* 2, *die restlichen Kanten sind länger, die Rechtwinkelkanten*

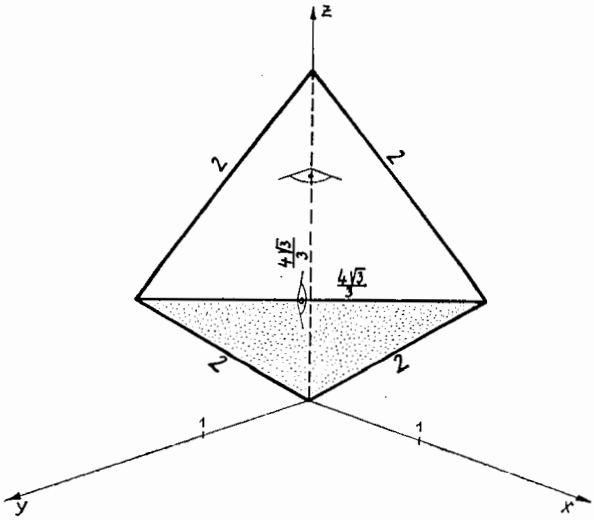

Abb. 13.

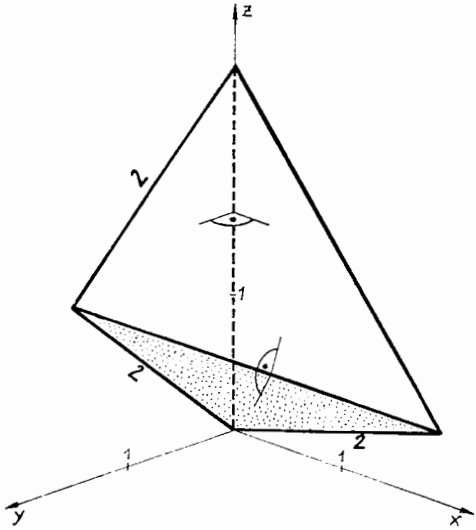

Abb. 14.

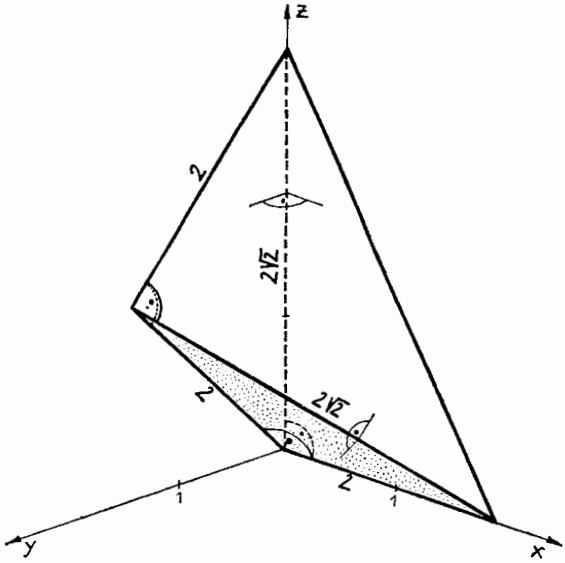

Abb. 15.

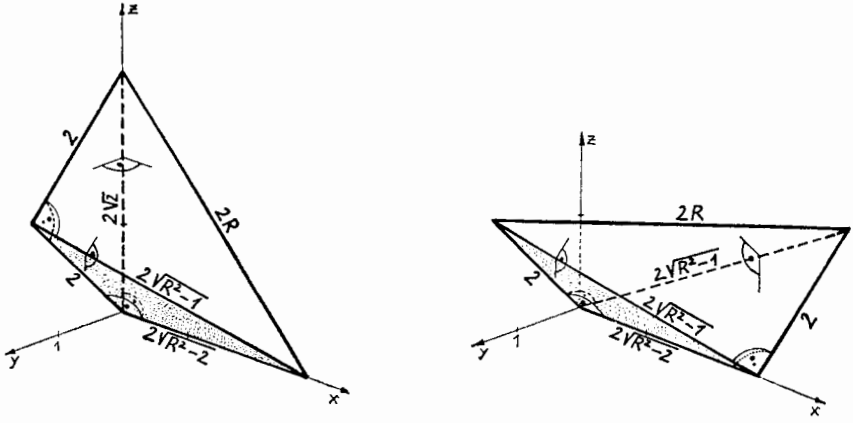

Abb. 16.

sind gleich lang. Das Tetraeder ist ähnlich zum Tetraeder ABCD mit Koordinaten nach (11), (34), (35), (36) *mit* $b = 2$ *und* $d_3 = \frac{1}{2}$ (Abb. 14). *Das Volumen berechnet sich nach der Formel*

$$\frac{\sqrt{2}}{6}\sqrt{\sqrt{R^2-1}\,\sqrt{R^2+15}-R^2-1\cdot(R^2+7-\sqrt{R^2-1}\,\sqrt{R^2+15})} \quad (\text{vgl. } (10)).$$

(III) *Für* $R = \sqrt{3}$ *gibt es bis auf Kongruenz genau ein Tetraeder mit den Eigenschaften* (a) *bis* (d) *und minimalem Volumen. Es ist ein Orthoschem mit den Eckpunkten* $P_1 \sim (0, 0, 0)$, $P_2 \sim (0, 0, 2\sqrt{2})$, $P_3 \sim (0, \sqrt{2}, \sqrt{2})$, $P_4 \sim (2, 0, 0)$ *und dem total-orthogonalen Kantenzug* $P_2P_3P_1P_4$, *dessen drei Kanten alle die Länge* **2** *haben, die restlichen Kanten sind länger, und die gegenüberliegenden Rechtwinkelkanten sind gleich lang* (Abb. 15). *Das Volumen beträgt* $\frac{4}{3}$.

(IV) *Für* $R > \sqrt{3}$ *gibt es bis auf Kongruenz genau zwei Tetraeder mit den Eigenschaften* (a) *bis* (d) *und minimalem Volumen. Es sind Orthoscheme, die durch die total-orthogonalen Kantenzüge mit den Längenfolgen* $(2, 2, 2\sqrt{R^2-2})$ *und* $(2, 2\sqrt{R^2-2}, 2)$ *charakterisiert sind. Die gegenüberliegenden Rechtwinkelkanten sind im ersten Fall unterschiedlich lang, beim zweiten Tetraeder sind sie gleich lang* (Abb. 16). *Das Volumen beträgt* $\frac{4}{3}\sqrt{R^2-2}$ (*vgl.* (10)).

Literatur

[1] Федоров, Е. С., *Симметрия правильных систем фигур*, СПб., 1980.
[2] Hollai, M., Doppelgitterförmige Lagerungen inkongruenter Kreise und Kugeln, *Annales Univ. Sci. Budapest*, Sectio Math., **18** (1975), 75–86.
[3] Sz. Hollai, M., Die bikonjugierten Punktsysteme. *Annales Univ. Sci. Budapest*, Sectio Math., 28 (1985), 279–281.

УЗКАЯ УПАКОВКА ОБЛАСТЕЙ ГИПЕРЦИКЛОВ И ГИПЕРСФЕР В ГИПЕРБОЛИЧЕСКИХ ПЛОСКОСТИ И ПРОСТРАНСТВЕ

БУЙ ВАН ЗУНГ

Кафедра Начертательной и Проективной Геометрии
Университета им. Л. Этвеша, Будапешт

(Поступило 9.9.1983.)

Пусть a_i и α_i — прямая и плоскость в гиперболических плоскости $\mathbf{H}^2$ или пространстве $\mathbf{H}^3$ соответственно. Множества точек P, расстояние которых от a_i в $\mathbf{H}^2$ или α_i в $\mathbf{H}^3$ меньше t, назвём областьями гиперцикла или гиперсферы и обозначим через $\mathbf{H}^2_{a_i t}$ или $\mathbf{H}^3_{\alpha_i t}$, где $t \geqq 0$ вещественное число. Прямая a_i или плоскость α_i называеться основой областей гиперцикла $\mathbf{H}^2_{a_i t}$ или гиперсферы $\mathbf{H}^3_{\alpha_i t}$.

Множество областей $\{\mathbf{H}^2_i\}_t$ или $\{\mathbf{H}^3_i\}_t$ образует упаковку, если любие два из них не имеют общую точку. Очевидно, если $\{\mathbf{H}^2_i\}_t$ или $\{\mathbf{H}^3_i\}_t$ образует упаковку, то их основы не пересекают друг друга и они параллельны, если $t = 0$.

Узость $e^2(t)$ или $e^3(t)$ данной упаковки областей $\{\mathbf{H}^2_i\}_t$ или $\{\mathbf{H}^3_i\}_t$ в $\mathbf{H}^2$ или в $\mathbf{H}^3$ определяется по формулой

$$e^n(t) = \sup \varrho, \quad n = 2, 3 \tag{1}$$

где точная верхняя грань берется по всем радиусам кругов или шаров, расположенных в не покрытых частях плоскости $\mathbf{H}^2$ или пространства $\mathbf{H}^3$.

Наша основная задача решить нижнюю грань узости $e^2(t)$ или $e^3(t)$ и упаковку областей $\{\mathbf{H}^2_i\}_t$ или $\{\mathbf{H}^3_i\}_t$, для которых этот минимум достигается.

Вопрос об узости упаковок поставил Л. Фейеш Тот [2]. Несколько результатов находятся в работах [1]—[7]. И. Вермеш [8] занимался плотностью упаковки областей $\{\mathbf{H}^2_i\}_t$.

В нашей работе мы найдем узкую упаковку областей $\{\mathbf{H}^2_i\}_t$ в $\mathbf{H}^2$ для всех $t \geq 0$ и исследуем функцию $e^2(t)$. Мы даем нижнюю оценку $\bar{e}^3(t)$ для узости упаковки областей $\{\mathbf{H}^3_i\}_t$ и найдем значения t и упаковки $\{\mathbf{H}^3_i\}_t$, для которых эта оценка точна (бесконечно много раз), далее исследуем функцию $\bar{e}^3(t)$.

1. Сначала мы конструируем правильную упаковку $\mathbf{H}^*$ областей $\{\mathbf{H}_i^2\}_t$ в $\mathbf{H}^2$, затем мы докажем, что $\mathbf{H}^*$ дает узкую упаковку областей $\{\mathbf{H}_i^2\}_t$ для всех t.

Посмотрим правильный треугольник идеальными вершинами A, B, C, в котором расстояние любых двух сторон равно $2t$ (идеальной точкой называется общая точка прямых, перпендикулярных данной прямой в $\mathbf{H}^2$ или данной плоскости в $\mathbf{H}^3$), области $\mathbf{H}^2_{AB,t}$, $\mathbf{H}^2_{AC,t}$ и $\mathbf{H}^2_{BC,t}$ попарно касаются (рис. 1.).

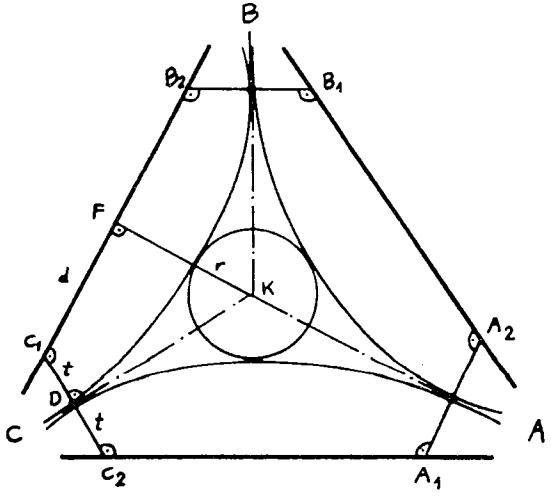

Рис. 1.

Пусть r — радиус круга, касающегося областей $\mathbf{H}^2_{AB,t}$, $\mathbf{H}^2_{AC,t}$ и $\mathbf{H}^2_{BC,t}$ и находящегося в прямоугольном шестиугольнике, в котором $B_1B_2 = A_1A_2 = C_1C_2$ и $AB \perp B_1B_2 \perp BC$ и. т. д. Очевидно, что

$$r = R - t, \tag{2}$$

где $R = KF$ — радиус вписанного круга треугольника ABC. Из четырёхугольника Ламберта FC_1DK

$$ch\, R = \frac{\operatorname{ch} t}{\sin\frac{\pi}{3}} = \frac{2}{\sqrt{3}} \operatorname{ch} t. \tag{3}$$

Так,

$$r = \operatorname{Arch}\left(\frac{2}{\sqrt{3}} \operatorname{ch} t\right) - t. \tag{4}$$

(В нашей работе мы используем естественный или абсолютный измеритель отрезков.)

Отобразим прямоугольный шестиугольник $A_1A_2B_1B_2C_1C_2$ и области $\mathbf{H}^2_{it}$ на стороны шестиугольника, потом продалжаем это отображение и получим разбиение плоскости $\mathbf{H}^2$ на шестиугольники и упаковки областей $\{\mathbf{H}^2_i\}_t$. Легко видеть, что узость упаковки областей $\{\mathbf{H}^2_i\}_t$ равна (4).

2. Теорема 1. *Узость упаковки областей* $\{\mathbf{H}^2_i\}_t$ *в* $\mathbf{H}^2$ *не меньше, чем*

$$\operatorname{Arch}\left(\frac{2}{\sqrt{3}}\operatorname{ch} t\right)-t,\ t\geq 0. \tag{6}$$

Равенство достигается в упаковке областей $\{\mathbf{H}^2_i\}_t$, *определенной в пункте* **1.**

Доказательство. Рассмотрим упаковку множества областей $\{\mathbf{H}^2_i\}_t$. Очевидно, что основы областей не пересекаются. Пусть $\mathbf{H}^2_{a_jt}$ — одна из $\{\mathbf{H}^2_i\}_t$ и a_j её основа.

Множество точек P, расстояние которые от a_j больше, чем от основ других областей $\mathbf{H}^2_{a_kt}\in\{\mathbf{H}^2_i\}_t$ $(k\neq j)$, назвём областью Дирихле $\mathbf{D}^2_j$, принадлежащей области $\mathbf{H}^2_{a_jt}$ (см. ещё в [8]). (Например, область Дирихле $\mathbf{D}_j$ в **1.** — многоугольник, симметрический на a_j. Число его вершин бесконечно, его угла равны $\frac{2\pi}{3}$ и расстояние его сторон от a_j равно t.) Пусть P — одна из вершин области $\mathbf{D}_j$, и $k\geq 3$ — её степени, т. е. число областей $\mathbf{D}_i$, вершиной которых является P, равно k. Из определения областей $\mathbf{D}_i$ следует, что расстояние точки P от основ равно R, и от других основ больше R. Эти основы (прямые) определяют k-угольник идеальными или бесконечными вершинами. (Вершина бесконечна только тогда, когда $t=0$.)

Пусть $A_1, A_2, \ldots, A_k$ его вершины и $A_{i1}A_{i2}$ — расстояние сторон, принадлежащих вершине A_i. Очевидно $A_{i1}A_{i2}\geq 2t$ и $PA_{i1}=PA_{i2}$ (рис. 2.), т. е. P находится на оси симметрии прямых a_{i-1} и a_i. (Может случится, что P не находится в прямоугольном $2k$-угольнике $A_{11}A_{12}A_{21}\ldots A_{i1}A_{i2}\ldots A_{k2}$.)

Пусть $2\gamma_i=\sphericalangle(E_{i-1}PE_i)$, где $PE_i\perp a_i$ и $E_0=E_k$ $(i=1,2,\ldots,k)$. Предположим, что $\gamma_2\leq\min(\gamma_1\ldots\gamma_i)$. Легко видеть, что

$$\gamma_2\leq\frac{\pi}{3} \tag{7}$$

и равенство только тогда, когда $k=3$ и треугольник $A_1A_2A_3$ правилен. Из четырёх-угольника Ламберта $B_2A_{22}E_2P$,

$$\operatorname{ch} R=\operatorname{ch} PE_2=\frac{\operatorname{ch} B_2A_{22}}{\sin\gamma_2}. \tag{8}$$

Из (7) и неравенства $\frac{1}{2}A_{21}A_{22}=B_2A_{22}\geq t$ следует

$$\operatorname{ch} R\geq\frac{\operatorname{ch} t}{\sin\sqrt{3}}=\frac{2}{\sqrt{3}}\operatorname{ch} t. \tag{9}$$

Из определения точки P следует, что существует круг радиуса $(R-t)$ с центром в точке P, который не пересекает ни одну из $\{\mathbf{H}^2_i\}_t$, т. е. узость упаковки областей $\{\mathbf{H}^2_i\}_t$ не меньше (6), и равенство только тогда, когда $\gamma_2 = \frac{\pi}{3}$ и $B_2A_{22} = t$, т. е. треугольник $A_1A_2A_3$ правилен. Равенство поступает в упаковке $\mathbf{H}^*$ (см. **1.**).

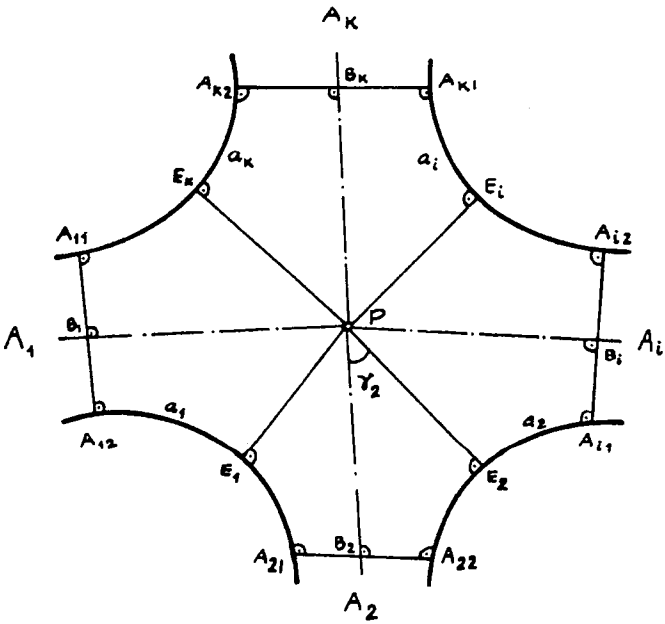

Рис. 2,

Легко видеть, что если P бесконечная или идеальная точка, то существует круг радиуса r, не пересекающий ни одну из $\{\mathbf{H}^2_i\}_t$ и $r > \operatorname{Arch}\left(\frac{2}{\sqrt{3}}\operatorname{ch} t\right) - t$.

Теорема 2. *Узость узкой упаковки областей $\{\mathbf{H}^2_i\}_t$ сторого убывающая функция от t. Значение узости меняется от $\ln\frac{3}{\sqrt{3}}$ до $\ln\frac{2}{\sqrt{3}}$ и равно $\ln\sqrt{3}$, если $t = 0$ и $\ln\frac{2}{\sqrt{3}}$, если $t \to \infty$.*

Доказательство. Из теоремы 1 мы знаем, что узость $\bar{e}(t)$ узкой упаковки областей $\{\mathbf{H}^2_i\}_t$

$$\bar{e}(t) = \operatorname{Arch}\left(\frac{2}{\sqrt{3}}\operatorname{ch} t\right) - t. \tag{10}$$

Из (10)

$$\bar{e}'(t) = \frac{\frac{2}{\sqrt{3}}\operatorname{sh} t}{\sqrt{\frac{4}{3}\operatorname{ch}^2 t - 1}} - 1 = \frac{\operatorname{sh} t}{\sqrt{\operatorname{ch}^2 t - \frac{3}{4}}} - 1.$$

Так как

$$\sqrt{\operatorname{ch}^2 t - \frac{3}{4}} > \sqrt{\operatorname{ch}^2 t - 1} = \operatorname{sh} t,$$

так

$$\frac{sh\, t}{\sqrt{\operatorname{ch}^2 t - \frac{3}{4}}} < 1,$$

поэтому $\bar{e}'(t) < 0$, т. е. $e(t)$ сторого убывающая функция.

Из (10)

$$\bar{e}(0) = \operatorname{Arch}\frac{2}{\sqrt{3}} = \ln\sqrt{3} \sim 0{,}549312\ldots. \tag{11}$$

Уравнение (10) может быть приведено к виду:

$$\bar{e}(t) = \ln\left(\frac{2}{\sqrt{3}}\operatorname{ch} t + \sqrt{\frac{4}{3}\operatorname{ch}^2 t - 1}\right) - t.$$

После превращения

$$\bar{e}(t) = \ln\frac{1}{\sqrt{3}} + \ln\left(1 + e^{-2t} + \sqrt{1 + e^{-4t} - e^{-2t}}\right),$$

поэтому

$$\lim_{t\to\infty} \bar{e}(t) = \ln\frac{2}{\sqrt{3}} \sim 0{,}143841\ldots. \tag{12}$$

Теорема 2 доказана.

3. Рассмотрим четыре гиперплоскости α_i ($i = 0, 1, 2, 3$), пусть расстояние между ними попарно равно $2t$ (рис. 3.). Области $\mathbf{H}^3_{\alpha_i t}$, принадлежащие плоскостям α_i попарно касаются. Пусть $A_i^k A_k^i$ — прямая, перпендикулярная α_i и α_k ($i \neq k$; i, $k = 0, 1, 2, 3$), где $A_i^k \in \alpha_i$. Так как $A_i^k A_k^i$, $A_i^l A_j^i$, $A_i^l A_l^i \perp \alpha_l$, поэтому рассмотренные 3 прямые имеют общую идеальную точки A_i. Тетраэдр идеальными вершинами $A_0 A_1 A_2 A_3$ правилен и тело $A_0^1 A_0^2 A_0^3 A_1^0 A_1^2 A_1^3 A_0^2 A_2^1 A_2^3 A_3^0 A_3^1 A_2^3$ правильный усечённый тетраэдр.

Пусть P — центр вписанного шара, касающегося плоскостей α_i, $E_i = PA_i \cap \alpha_i$ и $PE_i = R$, далее F_{ik} — середина отрезка $A_i^k A_k^i$, а $r(t)$ — радиус шара, касающегося четырёх областей $\{\mathbf{H}^3_{\alpha_i}\}_t$. Так

$$r(t) = R - t. \tag{13}$$

Из четырёхугольника Ламберта $E_0A_0^1F_{01}P$

(14) $$\operatorname{ch} PE_0 = \frac{\operatorname{ch} A_0^1F_{01}}{\sin v^*} = \frac{\operatorname{ch} t}{\sqrt{1-\cos^2 v^*}},$$

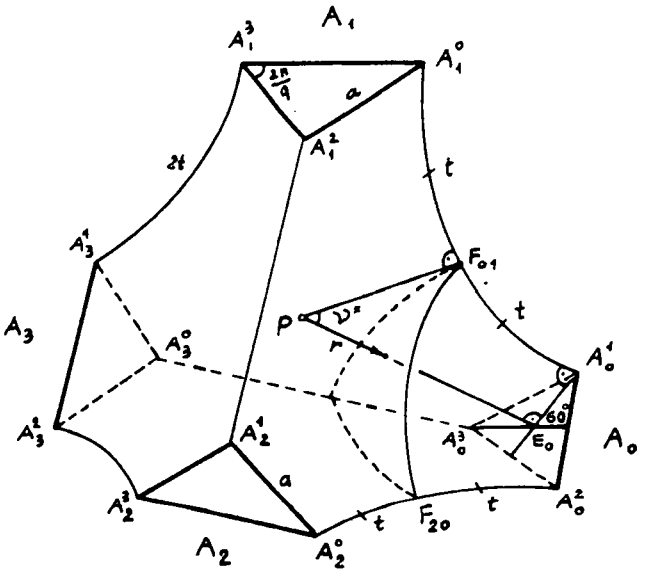

Рис. 3.

(15) $$\cos v^* = \operatorname{sh} A_0^1E_0 \cdot \operatorname{sh} A_0^1E_{01} = \operatorname{sh} A_0^1E_0 \cdot \operatorname{sh} t,$$

где

(16) $$v^* = \sphericalangle (E_{01}PE_0).$$

Из прямоугольного шестиугольника $A_0^1A_1^0A_1^2A_2^1A_2^0A_0^2$

$$\operatorname{ch} A_0^1A_0^2 = \frac{\operatorname{ch} 2t + \operatorname{ch}^2 2t}{\operatorname{sh}^2 2t}.$$

Из этого следует, что

(17) $$\operatorname{sh}\frac{A_0^1A_0^2}{2} = \frac{1}{2\operatorname{sh} t}.$$

Из (17) и правильного треугольника $A_0^1A_0^2A_0^3$

(18) $$\operatorname{sh} A_0^1E_0 = \operatorname{sh}\frac{A_0^1A_0^2}{2} \cdot \frac{1}{\sin 60^\circ} = \frac{1}{\sqrt{3}\operatorname{sh} t}.$$

Из (15) и (18)

$$\cos v^* = \frac{1}{\sqrt{3}}, \tag{19}$$

поэтому

$$\operatorname{ch} PE_i = \operatorname{ch} R = \sqrt{\frac{3}{2}} \operatorname{ch} t. \tag{20}$$

Так, что

$$r(t) = R - t = \operatorname{Arch}\left(\sqrt{\frac{3}{2}} \operatorname{ch} t\right) - t. \tag{21}$$

Лемма 3.1. Пусть α_i — плоскость, перпендикулярная ребрам A_iA_k ($k = 0, 1, 2, 3$; $k \neq i$) правильного тетраэдра идеальными вершинами $A_0A_1A_2A_3$, $A_i^kA_i^lA_i^j$ — вершинный треугольник, принадлежащий вершине A_i и $A_i^kA_k^i = 2t > 0$. В $\mathbf{H}^2$ для бесконечно много значений t существует нормальное разбиение на правильные треугольники $A_i^kA_i^lA_i^j$ где $i, j, k, l = 0, 1, 2, 3$ и разные числа.

Доказательство. Грань $A_0^1A_1^0A_1^2A_2^1A_2^0A_0^2$ усечённого тетраэдра A_i^k ($i \neq k$; $i, k = 0, 1, 2, 3$) — прямоугольный шестиугольник (рис. 3.). На основе тригонометрии прямоугольного шестиугольника [9] следует, что

$$\operatorname{ch} 2t = -\operatorname{ch}^2 2t + \operatorname{sh}^2 2t + \operatorname{ch} a,$$

где

$$a = A_1^0A_1^2 = A_2^1A_2^0 = A_0^2A_0^1.$$

Отсюда вытекает, что

$$\operatorname{ch} a = \frac{\operatorname{ch} 2t}{\operatorname{ch} 2t - 1}. \tag{22}$$

Из треугольника $A_0^1A_0^2A_0^3$

$$\cos\frac{2\pi}{q} = \frac{\operatorname{ch}^2 a - \operatorname{ch} a}{\operatorname{sh}^2 a} = \frac{\operatorname{ch} a}{\operatorname{ch} a + 1}, \tag{23}$$

где $\frac{2\pi}{q}$ — угол этого треугольника, $p > 6$ вещественное число. Из (22) и (23) дается, что

$$\cos\frac{2\pi}{q} = \frac{\operatorname{ch} 2t}{2\operatorname{ch} 2t - 1}. \tag{24}$$

Так как $0 < t < \infty$, то

$$1 > \cos\frac{2\pi}{q} > \frac{1}{2},$$

т. е. для любого целого числа $q>6$ существует t такое, что выполняется (24).

С другой стороны, известно, что если $q>6$ — целое число, то существует нормальное разбиение плоскости $\mathbf{H}^2$ на правильные треугольники с углами, равными $\frac{2\pi}{q}$.

Так лемма 3.1. доказана.

Лемма 3.2. Если $t\geq 0$, вещественное число и в уравнении (24) $q(t) \geq 7$ — целое число, то существует нормальное разбиение пространства $\mathbf{H}^3$ на правильные усечённые тетраэдры $A_i^k (i\neq k;\ i,\ k = 0, 1, 2, 3)$.

Доказательство. Рассмотрим правильный усечённый тетраэдр вершинами A_i^k и обозначим через $\mathbf{T}_e(t)$. Из ограничения для t, данного в Лемме 3.2., следует, что отображениями одно за другим относительно граней, получаем множество $\{\mathbf{T}(t)\}$, элементы которого не имеют внутренней точки. Для доказательства, что в $\mathbf{H}^3$ множество $\{\mathbf{T}(t)\}$ образует разбиение, достаточно показать, что любая точка пространства принадлежит по крайней мере одному из $\{\mathbf{T}(t)\}$.

Обозначим плоскость грани A_i^j, A_i^k, A_i^l, (где i, j, k, $l = 0, 1, 2, 3$ и не равны между собой) через $\alpha_{i(e)}$, а плоскость грани, противоположную от $\alpha_{i(e)}$ через $\beta_{i(e)}$, расстояние между ими через x (рис. 4.).

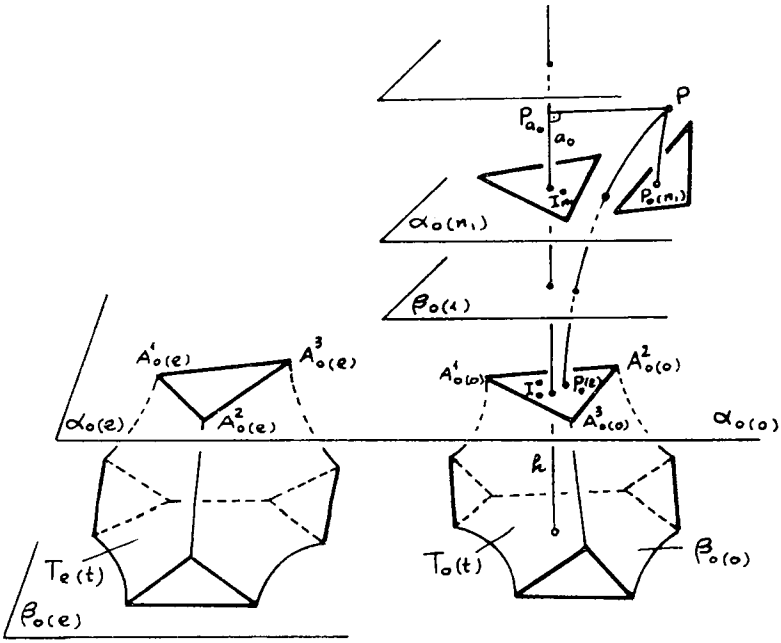

Рис. 4.

Пусть P — произвольная точка в $\mathbf{H}^3$, $PP_{i(e)}$ — расстояния точки P от плоткостей $\alpha_{i(e)}$. Предположим, что $PP_{0(e)} = \min(PP_{i(e)})$; $i = 0, 1, 2, 3$.

В связи с условием леммы следует, что существует разбиение плоскости $\alpha_{0(e)}$ на правильные треугольники с углами, равными $\frac{2\pi}{q}$ $(q \geq 7)$. Поэтому, $P_{0(e)}$ находится в одном, например, $A^1_{0(0)}A^2_{0(0)}A^3_{0(0)}$, из этих треугольников, принадлежащем к $\mathbf{T}_0(t)$ из $\{\mathbf{T}(t)\}$.

Пусть a_0 — ось симметрии $\mathbf{T}_0(t)$, перпендикулярная к плоскости $\alpha_{0(0)}(A^1_{0(0)}A^2_{0(0)}A^3_{0(0)})$, и P_{a_0} — проекция точки P на прямую a_0, далее $I^0_0 = a_0 \cap \alpha_{0(0)}$. Отображим $\mathbf{T}_0(t)$ на $\alpha_{0(0)}$, потом на $\beta_{0(1)}$, где $\beta_{0(1)}$ — отражение $\beta_{0(0)}$, и. т. д.

Пусть $I^0_{n_1}$ или $\alpha_{0(n_1)}$ — отражение точки I^0_0 или плоскости $\alpha_{0(0)}$ после $2n_1$ отображений. Очевидно, что существует целое число n_1 такое, что $P_{a_0} \in I^0_{n_1} I^0_{n_1+1}$.

Пусть $PP_{0(n_1)}$ — расстояние точки P от плоскости $\alpha_{0(n_1)}$. Совершенно так же как и выше изложенное рассуждение, существует разбиение плоскости $\alpha_{0(n_1)}$ на правильные треугольники. Поэтому $P_{0(n_1)}$ находится по крайней мере в одном из этих треугольников, принадлежающем например, к $\mathbf{T}_{n_1,0}(t)$ из $\{\mathbf{T}(t)\}$.

Пусть a_1 — ось симметрии $\mathbf{T}_{n_1,0}(t)$, перпендикуляная к $\alpha_{0(n_1)}$, далее P_{a_1} — проекция точки P на a_1 и $a_1 \cap \alpha_{0(n_1)} = I^{n_1}_0$. Аналогично предыдущему мы видим, что существует целое число n_2 такое, что после $2n_2$, отображений $P_{a_1} \in I^{n_1}_{n_2} I^{n_1}_{n_2+1}$, где $I^{n_1}_{n_0}$ — отраждение точки $I^{n_1}_0$.

Согласно изложенному, мы получим, что

$$PP_{0(n_1)} < PP_{0(e)} - 2n_1 \cdot h$$

$$PP_{0(n_2)} < PP_{0(n_1)} - 2n_2 \cdot h < PP_{0(e)} - 2(n_1 + n_2)h$$

$$\cdots\cdots\cdots$$

$$PP_{0(n_m)} < \ldots < PP_{0(e)} - 2(n_1 + n_2 + \ldots + n_m)h. \tag{25}$$

Из неравенства (25) следует, что после несколько отображений получается некоторое тело $\mathbf{T}_i(t)$, содержащее в себе точку P.

Итак Лемма 3.2. доказана.

4. Из лемма 3.2. видно, что для всех t столько, что в уравнении (24) $q(t)$ представляет собой целое число, существует нормальное разбиение пространтсва $\mathbf{H}^3$ на правильные усечённые тетраэдры $A^k_i (i \neq k;\ i, k = 0, 1, 2, 3)$, где $A^k_i A^i_k = 2t$.

Пусть α_i — плоскости треугольной грани этого тела. Посмотрим области гиперсфер $\{\mathbf{H}^3_i\}_t$ с основами α_i. Легко видеть, что множество областей $\{\mathbf{H}^3_i\}_t$ образует упаковку в $\mathbf{H}^3$ и значение узости этой упаковки выражено формулой (21).

Обозначим через $\mathbf{H}^{**}$ эту упаковку областей $\{\mathbf{H}^3_i\}_t$.

Теорема 3. *Узость упаковки областей* $\{\mathbf{H}^3_i\}_t$ *не меньше чем*

$$\operatorname{Arch}\left(\sqrt{\frac{3}{2}}\operatorname{ch} t\right)-t. \tag{26}$$

Равенства выпольняется, если в уравнении (24) $q(t)$ *представляет собой целое число и упаковка будет узкой упаковкой определенной в* **4.**

Доказательство. Рассмотрим упаковку множества $\{\mathbf{H}^3_i\}_t$. Пусть $\mathbf{H}^3_{\alpha_j t}$ — одна из областей $\{\mathbf{H}^3_i\}_t$ и α_j — её основа. Множества точек P, расстояние которых от α_j не больше чем от основ другие областей $\mathbf{H}^3_{\alpha_k t}\in\{\mathbf{H}^3_i\}_t$ $(j\neq k)$ называется областью Дирихле $\mathbf{D}^3_j$, принадлежащей области $\mathbf{H}^3_{\alpha_j t}$.

Пусть Q — одно из вершин области $\mathbf{D}^3_j$. Из определения области $\mathbf{D}^3_i$ следует, что расстояние точки Q от k основ равно $R(k\geq 4)$ и от других основ больше R. Выбираем из k областей $\{\mathbf{H}^3_k\}_t$ основами α_0, α_1, α_2, α_3, не перпендикуляными одновременно к ни одной плоскости. Расстояние $A^k_i A^i_k$ двух плоскостей из них не меньше $2t$, где $A^k_i\in\alpha_i$. Прямые $A^k_i A^i_k$ определяют тетраэдр идеальными вершинами A_0, A_1, A_2, A_3 и выпуклая оболочка точек A^k_i — усечённый тетраэдр (рис. 5.).

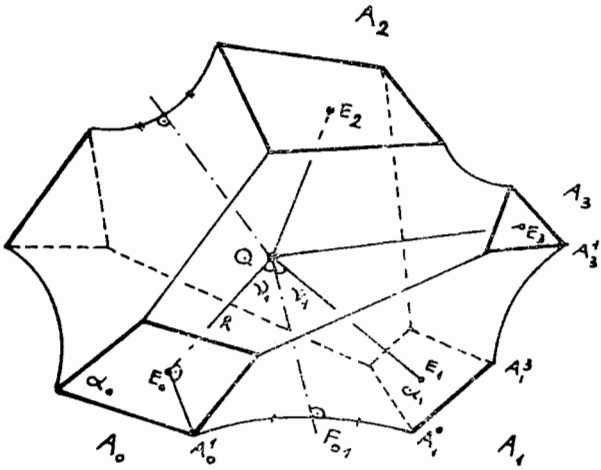

Рис. 5.

Пусть $E_i=\alpha_i\cap OA$, $QE_i=R$, далее $2\nu_1=\sphericalangle(E_0QE_1)$ наименьший угол из углов $\sphericalangle(E_iQE_j)$ $(i\neq j;\ i,\ j=0,\ 1,\ 2,\ 3)$.

Легко видеть, что

$$\nu_1\leq\nu^*, \tag{27}$$

где ν^* определён по формулами (15), (16) и (19) при случае, когда точки $E_i(i=0,\ 1,\ 2,\ 3)$ образуют правильный тетраэдр (см. **3.**).

Из четырёхугольника Ламберта $A_0^1F_{01}QE_0$,

$$\operatorname{ch} R = \operatorname{ch} QE_0 = \frac{\operatorname{ch} A_0^1F_{01}}{\sin \nu_1}. \tag{28}$$

Так как $\nu_1 \leq \nu^*$ и $A_0^1F_{01} \geq t$, следует

$$\operatorname{ch} R \geq \frac{\operatorname{ch} t}{\sin \nu^*} = \sqrt{\frac{3}{2}} \operatorname{ch} t. \tag{29}$$

Согласно узложенному, существует шар с радиусом $(R-t)$, не пересакающий ни одну из областей $\{\mathbf{H}_i^3\}_t$, т. е. узость упаковки областей $\{\mathbf{H}_i^3\}_t$ не меньше (26).

Заметим, что равенство имеет место только тогда, когда $\nu_1 = \nu^*$, $A_0^1F_{01} = t$, т. е. усечённый тетраэдров A_i^k стает правильным.

Как известно, сущесвует разбиение пространства $\mathbf{H}^3$ на правильные усечённые тетраэдры, если в уравнении (24) $q(t)$ представляет собой целое число. Отсюда следует, что равенство выполняется в случае упаковки $\mathbf{H}^{**}$.

Теорема 3 доказана.

Теорема 4. *Оценка $\bar{e}(t)$ для узости упаковки областей $\{\mathbf{H}_i^3\}_t$, данная в формуле* (26), *сторого убывающая функция.*

Доказательство. Из (26)

$$\bar{e}(t) = \operatorname{Arch}\left(\sqrt{\frac{3}{2}} \operatorname{ch} t\right) - t. \tag{30}$$

Отсюда

$$\bar{e}'(t) = \frac{\sqrt{\frac{3}{2}} \operatorname{sh} t}{\sqrt{\frac{3}{2} \operatorname{ch}^2 t - 1}} - 1 = \frac{\operatorname{sh} t}{\sqrt{\operatorname{ch}^2 t - \frac{2}{3}}} - 1.$$

Легко видеть (см. доказательство теоремы 2), что $\bar{e}'(t) < 0$, т. е. $\bar{e}(t)$ сторого убывающая функция переменного t.

$$\lim_{t \to \infty} \bar{e}(t) = \lim_{t \to \infty} \operatorname{Arch}\left(\sqrt{\frac{3}{2}} \operatorname{ch} t\right) - t = \ln \sqrt{\frac{3}{2}},$$

$$\bar{e}(0) = \operatorname{Arch} \sqrt{\frac{3}{2}} \sim 0{,}658478\ldots.$$

Литература

[1] BEZDEK, A.: Remark on the closest packings of convex discs, *Studia Sci. Math. Hung.* **15** (1980), 283–285.
[2] FEJES TÓTH, L.: Close packing and loose covering with balls, *Publ. Math. Debrecen*, **23** (1976), 323–326.
[3] FEJES TÓTH, L.: Remarks on the closest packing of convex discs, *Comment. Math. Helvitici*,, **53** (1978), 536–541.
[4] ХОРВАТ, Е. (HORVÁTH J.): Узкая решетчатая упаковка шаров в решетках первого типа, *Annales Univ. Sci. Budapest*, Sect. Math., **20** (1977), 191–194.
[5] ХОРВАТ, Е.: Об узкой решетчатой упаковке единичных шаров в пространстве E^n, *Труды МИАН СССР* **152** (1980), 216–231; English transl. in *Proc. Steklov Inst. of Math.*, 1982 no. 2., 237–254.
[6] ХОРВАТ, Е.: Об узости упаковки шаров, рыхлости покрытия шарами и k-узости множества точек в n-мерных пространствах постоянной кривизный, *Studia Sci. Math. Hung.*, (под. печ.).
[7] LINHART, J.: Closest packing and closest coverings by translates of a convex disc, *Studia Sci. Math. Hung.*, **13** (1978), 157–162.
[8] VERMES, I.: Ausfüllungen der hyperbolischen Ebene durch kongruente Hyperzykelbereiche, *Period. Math. Hung.*, **10** (1979), 217–229.
[9] НЕСТОРОВИЧ, Н. М.: *Геометрические построения в плоскости Лобачевского*, Государственное издательство технико–теоретической литературы, Москва–Ленинград 1951.

ON AN INVERSE FUNCTION THEOREM OF HALKIN

By

B. M. GARAY

Technical University of Budapest

(Received April 7, 1982)

In [3], HALKIN has proved the following

THEOREM: Let X be a finite dimensional real normed space. Let $x_0 \in X$, $U \subset X$ be a neighbourhood of x_0, $f: U \to X$ a continuous function which is differentiable at x_0, and suppose that $Df(x_0)$ is invertible. Then there exists a neighbourhood W of $f(x_0)$ and a function $h: W \to U$ such that $h(f(x_0)) = x_0$, $f(h(y)) = y$ for $y \in W$, h is differentiable at $f(x_0)$ and $Dh(f(x_0)) = [Df(x_0)]^{-1}$.

In [4], SZILÁGYI has given a new proof for Halkin's theorem. He has formulated the conjecture that Halkin's theorem does not remain valid if X is to be taken infinite dimensional. The aim of this paper is to verify Szilágyi's conjecture.

Let us observe first that the conclusion of Halkin's theorem implies that W being a neighbourhood of $f(x_0)$ is contained in the range of f. Consequently, $f(x_0)$ is an interior point of the range of f.

Let $(X, \|\cdot\|)$ be an infinite dimensional real Hilbert space. As usual, the zero element of X is denoted by $\mathbf{0}_X$. Fix an $e \in X$ such that $e \neq \mathbf{0}_X$. The orthogonal complement of $\{\lambda e | \lambda \in \mathbf{R}\}$ is denoted by M. Denote by P_M the orthogonal projection of X onto M and by P_0 the orthogonal projection onto $\{\lambda e | \lambda \in \mathbf{R}\}$.

Let $K = \{x \in X | \ \|P_M x\| < \|P_0 x\|^2\}$. Since $\{\lambda e | \lambda \in \mathbf{R} \setminus \{\mathbf{0}\}\} \subset K$ cuts every neighbourhood of $\mathbf{0}_X$, it follows that

(1) $\qquad \mathbf{0}_X$ is not an interior point of $X \setminus K$.

In order to prove Szilágyi's conjecture, we define a function $f: X \to X$ satisfying the following conditions:

(2) $\qquad f(\mathbf{0}_X) = \mathbf{0}_X$.

(3) $\qquad f$ maps X onto $X \setminus K$.

(4) $\qquad f$ is continuous on X.

(5) $\qquad f$ is Frechet differentiable at $\mathbf{0}_X$ and the derivative is I, the identity on X.

Let B_M and S_M denote $\{x \in X | \ \|x\| \leq 1\} \cap M$ and $\{x \in X | \ \|x\| = 1\} \cap M$, respectively. It is well-known that S_M is a retract of B_M [1], i. e. there exists a continuous function $r: B_M \to S_M$ such that for $x \in S_M$, $r(x) = x$.

As usual, the closure of K is denoted by $\overline{K}$. We define a function $g : \overline{K} \to X$ as follows

$$g(x) = \begin{cases} P_0x + \|P_0x\|^2\, r(P_Mx/\|P_0x\|^2) & \text{if } x \in \overline{K}\setminus\{\mathbf{0}_X\} \\ \mathbf{0}_X & \text{if } x = \mathbf{0}_X. \end{cases}$$

The definition makes sense since $x \in \overline{K}\setminus\{\mathbf{0}_X\}$ implies $P_0x \neq \mathbf{0}_X$, $P_Mx/\|P_0x\|^2 \in B_M$.

Before defining f, we make some simple remarks. Let us observe first that g is continuous on $\overline{K}$. For $x \in \overline{K}\setminus\{\mathbf{0}_X\}$, it is a simple consequence of the continuity of P_0, P_M, r. The continuity at $\mathbf{0}_X$ follows from

$$\|g(x) - x\| \leq 2\|x\|^2, \quad x \in \overline{K}. \tag{6}$$

Inequality (6) can be checked by a simple computation. In fact, $x \in \overline{K}\setminus\{\mathbf{0}_X\}$ implies

$$\|g(x) - x\| = \|P_0x - x + \|P_0x\|^2\, r(P_Mx/\|P_0x\|^2)\| \leq$$
$$\leq \|-P_Mx\| + \|P_0x\|^2 \leq 2\|P_0x\|^2 \leq 2\|x\|^2.$$

On the other hand, $\|g(\mathbf{0}_X) - \mathbf{0}_X\| = 0$.

A direct computation shows that

$$g(x) = x \text{ provided that } x \in \overline{K}\setminus K. \tag{7}$$

In fact, $x \in \overline{K}\setminus(K \cup \{\mathbf{0}_X\})$ implies $P_Mx_0/\|Px\|^2 \in S_M$, which, in turn, implies

$$g(x) = P_0x + \|P_0x\|^2 P_Mx/\|P_0x\|^2 = P_0x + P_Mx = x$$

Now we are in a position to define f. For $x \in X$, let

$$f(x) = \begin{cases} g(x) & \text{if } x \in K \\ x & \text{if } x \in X\setminus K. \end{cases}$$

We have to show that f satisfies conditions (2), (3), (4), (5).

(2) follows directly from the definition of f.

In order to prove (3) we distinguish two cases according as $x \in K$ or $x \in X\setminus K$. If $x \in K$, then $P_0f(x) = P_0x$ and $P_Mf(x) = \|P_0x\|^2\, r(P_Mx/\|P_0x\|^2)$. Since $r(P_Mx/\|P_0x\|^2) \in S_M$, it follows that $f(x) \in \overline{K}\setminus K \subset X\setminus K$. The case $x \in X\setminus K$ is, of course, trivial.

Clearly f is continuous on K and on $X\setminus K$, separately. Since g is continuous on $\overline{K}$, the continuity of f on X is a simple consequence of (7). Thus f satisfies condition (4).

Using inequality (6), it follows directly from the definition of f that

$$\text{for all } x \in X,\ \|f(x) - x\| \leq 2\|x\|^2. \tag{8}$$

5) is a simple consequence of inequality (8).

In virtue of (4) and (5), the conditions of Halkin's theorem are fulfilled apart from the one assuring X to be finite dimensional. (x_0 is to be taken for $\mathbf{0}_X$, U can be taken for X.) On the other hand, (1), (2) and (3) imply that the conclusion of Halkin's theorem is not true. Thus, we have arrived at the following

PROPOSITION. *Halkin's theorem does not remain valid if X is allowed to be infinite dimensional.*

On an infinite dimensional version of Halkin's theorem and on its applications in the theory of functional differential equations, see [2].

ACKNOWLEDGEMENT. The author is indebted to T. SZILÁGYI for valuable discussions.

References

[1] BORSUK, K.: *Theory of retracts*, PWN, Warsaw, 1967.
[2] GARAY, B. M.: Controllably periodic perturbations of autonomous functional differential equations in *Colloquia Math. Soc. J. Bolyai 30*, Qualitative theory of differential equations, Szeged, 1979, 267–276.
[3] HALKIN, H.: Implicite functions and optimisation problems without continuous differentiability of the data, *SIAM J. of Control*, 12 (1974), 229–236.
[4] SZILÁGYI, T.: Inverse function theorem – without assuming continuous differentiability, *Annales Univ. Sci. Budapest*, Sect. Math., **20** (1977), 107–110.

ADDED IN PROOF. The author was informed in 1986 that, independently from Halkin's original work [3], W. BARIT and G. R. WOOD [Differentiable retracts and a modified inverse function theorem, *Bull. Austral. Math. Soc.* **18** (1978), 37–43.] have proved a version of Halkin's inverse function theorem. Their paper contains an infinite-dimensional counterexample similar to the one of the present paper.

APPROXIMATION BY VILENKIN POLYNOMIALS

By

S. FRIDLI

Department II. for Analysis of the L. Eötvös University, Budapest

(Received September 24, 1982)

Introduction. In this paper we prove an approximation theorem for the Vilenkin systems, which is in the so-called "bounded" case a generalization of a theorem proved for the Walsh system by H. J. WAGNER in [7]. We follow the method of paper [7]. Our statement is proved for those Vilenkin systems, which satisfy a certain condition. We remark that this condition was introduced earlier by T. S. QUEK and LEONARD Y. H. YAP [4] and it is satisfied not only for the "bounded", but many "unbounded" Vilenkin systems too. The concept of a derivative due to C. W. ONNEWEER [2] plays an important role in our paper.

§. 1. In this section we introduce some notations and definitions. For all sequences

$$(1) \qquad m = (m_0, m_1, \ldots, m_k, \ldots), \quad (2 \leq m_k, \ m_k \in \mathbf{N}: = \{0, 1, 2, \ldots\})$$

of natural numbers let us denote by Z_{m_k} the m_k^{th} discrete cyclic group, i.e.

$$Z_{m_k}: = \{0, 1, \ldots, m_k - 1\}, \quad (k \in \mathbf{N})$$

and we define the group G_m as the direct product of the groups Z_{m_k}. Then G_m is a compact Abelian group, and his elements are of the from $x = (x_0, x_1, \ldots, \ldots, x_k, \ldots)$ with $x_k \in Z_{m_k}$, $(k \in \mathbf{N})$. For $x, y \in G_m$ their sum is got by adding the kth coordinates of x and y modulo m_k, $(k \in \mathbf{N})$. The symbol of the addition is $\dot{+}$, the inverse of $\dot{+}$ let $\dot{-}$.

Let $\hat{G}_m = \{\psi_n | n \in \mathbf{N}\}$ denote the character group of G_m. We enumerate the elements of $\hat{G}_m$ as follows. For $k \in \mathbf{N}$ and $x \in G_m$ let r_k be the function given by

$$r_k(x): = \exp \frac{2\pi i x_k}{m_k} \quad (x \in G_m, i: = \sqrt{-1}, k \in \mathbf{N}).$$

Define the sequence $(M_k)_{k\in\mathbf{N}}$ by $M_0 := 1$ and $M_k := m_0\, m_1 \ldots m_{k-1}$, $(k\in\mathbf{P} := \{1, 2, \ldots\})$, then each $n\in\mathbf{N}$ has a unique representation of the form

$$n = \sum_{k=0}^{\infty} n_k M_k$$

where $0 \leq n_k < m_k$ $(n_k \in \mathbf{N})$. Now we define the function ψ_n by

$$\psi_n := \prod_{k=0}^{\infty} (r_k)^{n_k}, \quad (n\in\mathbf{N}).$$

We remark that $\hat{G}_m$ is a complete orthonormal system with respect to the normalized Haar measure dx on G_m [6].

For $f\in L_1(G_m)$ we define its Fourier coefficients by

$$\hat{f}(n) := \int_{G_m} f(t)\overline{\psi_n(t)}\, dt.$$

$D_n = \sum_{k=0}^{n-1} \psi_k$ $(n\in\mathbf{N})$ is the so-called Dirichlet kernel. It is known [6] that

$$D_{M_n}(x) = \begin{cases} M_n, & (x_k = 0,\ 0 \leq k < n) \\ 0 & \text{otherwise} \end{cases}$$

We denote the spaces $L^p(G_m)$ $(1 \leq p < \infty)$ or $C(G_m)$ (the space of the continuous functions on G_m) by $X(G_m)$. If $f, g\in L^1(G_m)$, then the convolution $f*g$ is defined by $(f*g)(x) = \int_{G_m} f(t)\, g(x \dot{-} t)\, dt$, $(x\in G_m)$. We define the integrated modulus of continuity as follows

$$\omega(f, X, \delta) := \sup_{\lambda(y)<\delta} \|f(\cdot \dot{+} y) - f(\cdot)\|_X, \quad (\delta > 0, f\in X(G_m)),$$

where

$$\lambda(y) := \sum_{k=0}^{\infty} \frac{y_k}{M_{k+1}}, \quad (y\in G_m).$$

Let Lip (α, X) be the set of those functions for which

$$\|f(\cdot) - f(\cdot \dot{+} y)\|_X = O(\lambda(y)^\alpha) \quad (\lambda(y) \to 0,\ \alpha > 0, f\in X(G_m))$$

holds.

C. W. Onneweer [2] has given the following generalization of the concept of a derivative of P. L. Butzer and H. J. Wagner: the function $c\in X(G_m)$ has a (strong) derivative

$$d^{[1]}f = df \in X(G_m) \text{ if } \lim_n \|df - d_n f\|_X = 0$$

where

$$(d_n f)(x) := \sum_{j=0}^{n-1} M_j \sum_{k=0}^{m_j-1} k\, m_j^{-1} \sum_{l=0}^{m_j-1} \overline{r_j(l\, e_j)^k} f(x \dotplus l\, e_j)$$

$$(x \in G_m,\ n \in \mathbf{P},\ e_j := (\overset{0}{\smile}{0}, \overset{1}{\smile}{0}, \ldots, 0, \overset{j}{\smile}{1}, 0, \ldots) \in G_m,\ (e_j := \overset{1}{\smile}{e_j} \dotplus \overset{2}{\smile}{e_j} \dotplus \ldots \dotplus \overset{l}{\smile}{e_j}).$$

The operator of the derivative d is linear and closed in $X(G_m)$ [2].

A good property of d is easily verified:

$$d\psi_n = n\psi_n, \qquad (n \in \mathbf{N}).$$

The nth, $(n \in \mathbf{P})$ derivative $d^{[n]} f$ of f is defined by induction. Let us denote by P_n the set of the Vilenkin polynomials of order $n \in \mathbf{P}$, i. e.

$$P_n := \{f \in L^1(G_m) \mid \hat{f}(k+n) = 0,\ k \in \mathbf{N}\},$$

and for an arbitrary $f \in X(G_m)$ we define $E_n(f, X)$, $(n \in \mathbf{P})$ as follows

$$E_n(f, X) := \inf_{p \in P_n} \|f - p\|_X.$$

It is well-known [6] that there exists $p_{n,f} \in P_n$, for which $E_n(f, X) = \|f - p_{n,f}\|_X$, $(f \in X(G_n))$ holds. The set of Vilenkin polynomials $(\bigcup_{n \in P} P_n)$ is dense in $X(G_m)$ [6], consequently $\lim_n E_n(f, X) = 0$ $(f \in X(G_m))$.

Further we say "bounded" Vilenkin system if $\limsup m < \infty$ and" unbounded" Vilenkin system otherwise.

§. 2. Let us denote by $\mathcal{H}$ the set of the sequences m with the property (1), for which the folloving condition holds:

(2) for every $\varepsilon > 0$ there exists $n_0 \in \mathbf{N}$ such that $m_n < M_n^{\varepsilon}$ for all $n > n_0$, $(n \in \mathbf{N})$.

Next we give examples for some sequences m belonging to $\mathcal{H}$.

(i) If m is bounded, then $m \in \mathcal{H}$.
(ii) If $m_n = O(n^k)$ (where k is an arbirary but fixed natural number), or $m_n = O(n!)$, $(n \in \mathbf{N})$, then $m \in \mathcal{H}$.

From the definition of $\mathcal{H}$ it follows that for all $m \in \mathcal{H}$ and $\varepsilon > 0$ are the following assertions true:

(3)
(i) $$\sum_{n=0}^{\infty} \frac{m_n}{M_n^{\varepsilon}} < \infty,$$

(ii) $$n = O_{\varepsilon}\left(M_{K(n)}^{1+\varepsilon}\right) \quad \text{and}$$

(iii) $$M_{K(n)}^{-1} = O_{\varepsilon}(n^{-1+\varepsilon})$$

where $K(n)\in\mathbf{P}$ denotes the number for which $M_{K(n)}\leq n<M_{K(n)+1}$ holds. The next theorem is the main result of this paper. We shall prove the equivalence of five conditions for those functions of $X(G_m)$, which are to be good approximated in $X(G_m)$ by Vilenkin polynomials.

THEOREM. *If $m\in\mathcal{H}$, $f\in X(G_m)$, $r\in\mathbf{N}$, $\alpha>0$ and $n\to\infty$, then the following conditions are equivalent:*

(i) $$E_n(f,X)=O_\varepsilon(n^{-r-\alpha+\varepsilon}),\ (\varepsilon>0),$$

(ii) $$d^{[r]}f\in X(G_m)\ \textit{and}\ d^{[r]}f\in \mathrm{Lip}\,(\alpha-\varepsilon,X),\ (0<\varepsilon<\alpha),$$

(iii) $$d^{[r]}f\in X(G_m)\ \textit{and}\ \omega(d^{[r]}f,X,n^{-1})=O_\varepsilon(n^{-\alpha+\varepsilon}),\ (\varepsilon>0),$$

(iv) $$d^{[r]}f\in X(G_m),\ 0\leq s\leq r\quad \textit{and}$$
$$\|d^{[s]}f-d^{[s]}p_{n,f}\|_X=O_\varepsilon(n^{-r-\alpha+s+\varepsilon}),\ (\varepsilon>0),$$

(v) $$\|d^{[s]}p_{n,f}\|_X=O_\varepsilon(n^{s-r-\alpha+\varepsilon}),\ (s>r+\alpha,\varepsilon>0).$$

This theorem shows the applicability of the concept of the derivative in the approximation theory. We remark that the above theorem is valid also in the case $\varepsilon=0$ for the bounded Vilenkin system. (For the Walsh system see [7].)

§. 3. To the proof of the theorem we need some lemmas.

LEMMA 1. (see e. g. [1]) Let $\hat{G}_m$ be an arbitrary Vilenkin system. Then for every $f\in X(G_m)$

$$E_{M_n}(f,X)\leq\omega(f,X,M_n^{-1})\leq 2E_{M_n}(f,X),\ (n\in\mathbf{N}).$$

LEMMA 2. If the function f belongs to $X(G_m)$ and $r\in\mathbf{N}$ such that $d^{[r]}f\in X(G_m)$, then

(i) $$\omega(f,X,\delta)=O\left[\left(\frac{\log M_n}{M_n}\right)^r\omega(d^{[r]}f,X,\delta)\right],$$

where $M_n\leq\delta^{-1}<M_{n+1}$, $\delta\to 0$.

(ii) If m is bounded, then

$$\omega(f,X,\delta)=O(\delta^r\omega(d^{[r]}f,X,\delta)),\ (\delta\to 0).$$

PROOF. Define the functions $w_r^n\in L^2(G_m)$, $(n\in\mathbf{N})$ by its Vilenkin-Fourier coefficiens as follows:

$$(w_1^n)^\wedge(k)=\begin{cases}0 & \text{if } 0\leq k<M_n\\ \dfrac{1}{k} & \text{if } k\geq M_n,\end{cases}\quad (k\in\mathbf{N}).$$

Furthermore the functions w_r^n, $(r\in\mathbf{P},\ n\in\mathbf{N})$ are defined inductively:

$$w_r^n:=w_{r-1}^n*w_1^n.$$

It is obvious (by the well-known equality $(f*g)^\wedge(k) = \hat{f}(k)\cdot\hat{g}(k)$) that

$$(w_r^n)^\wedge(k) = \begin{cases} 0 & \text{if } 0 \le k < M_n \\ \dfrac{1}{k^r} & \text{if } k \ge M_n \ (k\in\mathbf{N}, n\in\mathbf{N}, r\in P). \end{cases}$$

The relation

$$4) \qquad f(x)-f(x\dot{+}h) = (w_r^n * d^{[r]}f)(x)-(w_r^n * d^{[r]}f)(x\dot{+}h)$$

$\left(\text{a.e. } x\in G_m,\ \lambda(h)<\dfrac{1}{M_n}\right)$ is to be proved by comparison of the Fourier coefficients, applying Fubini's theorem and $(d^{[r]}\hat{f})^\wedge(k) = k^r f(k)$, $(k\in\mathbf{N})$.
By means of the convolution theorem it follows from (4) that

$$\omega(f, X, \delta) = \omega(w_r^n * d^{[r]}f, X, \delta) \le \|w_r^n\|_1 \cdot \omega(d^{[r]}f, X, \delta).$$

Since in the "bounded" case $\|w_1^n\|_1 = O(M_n^{-r})$ holds ([3] lemma 2), thus $\|w_r^n\|_1 = O(M_n^{-r})$, $(n\in\mathbf{N}, r\in\mathbf{P})$. Applying that $M_n^{-1} = O(\delta^r)$ the proof of part (ii) is complete.

Now we consider part (i). Let $w_{1,k}^n = w_1^n * D_{M_k}$ for an arbitrary $k\in\mathbf{N}$, then $\lim\limits_k \|w_1^n - w_{1,k}^n\|_1 = 0$ (see [3] lemma 1). Applying the Abel transformation we obtain

$$\|w_1^n\|_1 = \lim_k \|w_{1,k}\|_1 = O\left[\lim_k\left(\sum_{l=M_n}^{M_k-1}\frac{\log l}{l^2}+\frac{1}{M_k}+\frac{1}{M_n-1}\right)\right] = O\left(\frac{\log M_n}{M_n}\right).$$

From this $\|w_r^n\|_1 = O\left[\left(\dfrac{\log M_n}{M_n}\right)^r\right]$ follows. Lemma 2 is proved.

COROLLARY 1. If $m\in\mathcal{H}$, then

$$\left(\frac{\log M_n}{M_n}\right)^r = O_\varepsilon(\delta^{r-\varepsilon}), \quad (\varepsilon>0),$$

therefore

$$\omega(f, X, \delta) = O_\varepsilon[\delta^{r-\varepsilon}\omega(d^{[r]}f, X, \delta)], \quad (\varepsilon>0).$$

COROLLARY 2. Let $f\in X(G_m)$ and $d^{[r]}f\in X(G_m)$, $(r\in\mathbf{P})$.

(i) If $m\in\mathcal{H}$, then $E_n(f, X) = O_\varepsilon(n^{-r+\varepsilon}\|d^{[r]}f\|_X)$, $(\varepsilon>0)$.

(ii) If $\overline{\lim}\, m<\infty$, then $E_n(f, X) = O(n^{-r}\|d^{[r]}f\|_X)$.

LEMMA 3. The following Bernstein type statement is true.

(i) If $m\in\mathcal{H}$ and $p\in P_n$, then

$$\|d^{[r]}p\|_X = O_\varepsilon(n^{r+\varepsilon}\|p\|_X), \quad (n\in\mathbf{N}, r\in\mathbf{P}, \varepsilon>0).$$

(ii) If $\overline{\lim}\, m < \infty$ and $p \in P_n$, then

$$\|d^{[r]}p\|_X = O(n^r \|p\|_X), \quad (n \in \mathbf{N}, r \in \mathbf{P}). \tag{5}$$

PROOF. It is enough to prove the statements only in the case $r = 1$, as for $r > 1$ they follow from this by induction. Since

$$|r_j(l\, e_j)^k| = 1 \quad (j, l, k \in \mathbf{N}, 0 \le l < m_j, 0 \le k < m_j),$$

thus we have that

$$\|dp\|_X = \left\| \sum_{j=0}^{K(n)} M_j \sum_{l=0}^{m_j-1} k m_j^{-1} \sum_{l=1}^{m_j-1} \overline{r_j(l\, e_j)^k}\, p(\cdot \dot{+} l\, e_j) \right\|_X \le$$

$$\le \|p\|_X \sum_{j=0}^{K(n)} m_j^2 M_j.$$

First we consider part (ii) of Lemma 3. In this case $m_j = O(1)$ and $\sum_{j=0}^{K(n)} M_j < 2n$, consequently (ii) is proved.

Now we prove part (i). It is obvious that

$$\sum_{j=0}^{K(n)} m_j^2 M_j = n^{1+\varepsilon} \sum_{j=0}^{K(n)} \frac{m_j^2 M_j}{n^{1+\varepsilon}} \le n^{1+\varepsilon} \sum_{j=0}^{K(n)} \frac{m_j^2}{M_j^\varepsilon}, \quad (\varepsilon > 0).$$

By (2) (i) we get $\sum_{j=0}^{\infty} \frac{m_j^2}{M_j^\varepsilon} < \infty$, thus the relation (i) is true. This proves our statements.

We remark that the relation (5) is not valid, if the Vilenkin system has an unbounded structure, namely $D_{M_n} \in P_{M_n} (n \in \mathbf{N})$ but $\|dD_{M_n}\|_1 \ne O(M_n)$ [5] in this case.

Now we begin to prove the theorem.

PROOF of (i)$\Rightarrow$(iv). We define for an arbitrary but fixed $f \in X(G_m)$ the sequence of functions $(U_k)_{k \in \mathbf{N}}$ as follows:

$$U_k := \begin{cases} p_{M_{k+1}, f} - p_{M_k, f} & \text{if } k = 2, 3, \ldots \\ p_{M_k, f} & \text{if } k = 1 \end{cases}$$

By (i) we have that

$$\|U_k\|_X \le \|p_{M_{k+1}, f} - f\|_X + \|f - p_{M_k, f}\|_X \le 2E_{M_k}(f, X) = O_\varepsilon(M_k^{-r-\alpha+\varepsilon}),$$
$$(\varepsilon < 0).$$

We have, applying Lemma 3 (i), the estimation

$$\|d^{[s]}U_k\|_X = O_\varepsilon(M_k^{s-r-\alpha+\varepsilon}), \quad (\varepsilon > 0). \tag{6}$$

Since $p_{M_{n,f}} = \sum_{k=1}^{n-1} U_k$, therefore by (6) we get for $\varepsilon < \alpha$ that

$$\|d^{[s]}p_{M_l,f} - d^{[s]}p_{M_n,f}\|_X = \left\| \sum_{k=n}^{l-1} d^{[s]}U_k \right\|_X = O_\varepsilon(M_k^{s-r-\alpha+\varepsilon}), \quad (n,l \in \mathbf{N},\ l > n).$$

From this it is to be seen that $(d^{[s]}p_{M_n,f})_{n\in\mathbf{N}}$ is a Cauchy sequence, and since the operator d is closed, consequently there exists $d^{[s]}f \in X(G_m)$ for which

$$\lim_n \|d^{[s]}f - d^{[s]}p_{M_n,f}\|_X = 0. \tag{7}$$

Now we consider the formula

$$\|d^{[s]}f - d^{[s]}p_{n,f}\|_X, \quad (n \in \mathbf{N}).$$

By reason of the definition of U_k we have

$$\|d^{[s]}f - d^{[s]}p_{n,f}\|_X \leq \|d^{[s]}f - d^{[s]}p_{M_{K(n)+1},f}\|_X + \|d^{[s]}p_{n,f} - d^{[s]}p_{M_{K(n)+1},f}\|_X \leq$$

$$\leq \|d^{[s]}f - d^{[s]}p_{M_{l,f}}\|_X + \|\sum_{K=l(n)+1}^{l} d^{[s]}U_k\|_X + \|d^{[s]}p_{n,f} - d^{[s]}p_{M_{K(n)+1},f}\|_X$$

where $l \in \mathbf{N}$, $l > K(n)+1$. In this estimation there tends the first part by (7) to zero, furthermore

$$\lim_l \left\| \sum_{k=K(n)+1}^{l} d^{[s]}U_k \right\|_X = O_\varepsilon(M_{K(n)+1}^{s-r-\alpha+\varepsilon})$$

is valid.

On the other hand

$$\|p_{n,f} - p_{M_{K(n)+1},f}\|_X \leq 2E_n(f, X) = O_\varepsilon(n^{-r-\alpha-\varepsilon})$$

hence according to Lemma 3 (i) and the relation

$$M_{K(n)+1} = O_\varepsilon(n^{1+\varepsilon})$$

we obtain

$$\|d^{[s]}p_{n,f} - d^{[s]}p_{M_{K(n)+1},f}\|_X = O_\varepsilon(n^{s-r-\alpha+\varepsilon}).$$

From the above facts we have for all $\varepsilon > 0$

$$\|d^{[s]}f - d^{[s]}p_{n,f}\|_X = O_\varepsilon(n^{s-r-\alpha+\varepsilon})$$

and (i)$\Rightarrow$(iv) is proved.

COROLLARY 3. Let $f \in X(G_m)$ and $r \geq 1$, then from (i) follows the existence of $d^{[s]}f \in X(G_m)$ such that $E_n(d^{[s]}f, X) = O_\varepsilon(n^{s-r-\alpha+\varepsilon})$, $(\varepsilon < 0)$, $(0 \leq s \leq r)$.

PROOF of (iv)$\Rightarrow$(i). This is a simple consequence of Corollary 2.

PROOF of (i)$\Rightarrow$(v). For $\|d^{[s]}p_{n,f}\|_X$ $(s \in \mathbf{N})$ we get easily the estimation

$$\|d^{[s]}p_{n,f}\|_X \leq \|d^{[s]}p_{M_{K(n)},f}\|_X + \|d^{[s]}p_{n,f} - d^{[s]}p_{M_{K(n)},f}\|_X.$$

By means of the method used in the proof of (i)$\Rightarrow$(iv) we obtain

$$\|d^{[s]}p_{n,f}-d^{[s]}p_{M_{K(n)},f}\|_X = O_\varepsilon(n^{s-r-\alpha+\varepsilon}),\ (\varepsilon>0),$$

and according to (6) we have

$$\|d^{[s]}p_{M_{K(n)},f}\|_X = \left\|\sum_{k=1}^{K(n)-1} d^{[s]}U_k\right\|_X = O_\varepsilon(n^{s-r-\alpha+\varepsilon}),\ (\varepsilon>0).$$

These prove the statement (i)$\Rightarrow$(v).

Proof of (v)$\Rightarrow$(i). If $p_{M_{K(n)}}$ is an arbitrary element of $P_{M_{K(n)}}$, then we can get the following estimation:

$$E_n(f, X)\leq E_{M_{K(n)}}(f, X)\leq \|f-p_{M_{K(n)}}\|_X\leq$$

$$\leq \|f-p_{M_{K(n)+1},f}\|_X+\|p_{M_{K(n)+1},f}-p_{M_{K(n)}}\|_X.$$

For $p_{M_{K(n)}}$, $p_{M_{K(n)+1},f}$ the above relation has the form

$$E_{M_{K(n)}}(f, X)\leq E_{M_{K(n)+1}}(f, X)+E_{M_{K(n)}}(p_{M_{K(n)+1},f}, X).$$

By induction it is not hard to see that

$$E_{M_{K(n)}}(f, X)\leq E_{M_{K(n)+k}}(f, X)+\sum_{i=1}^{k}E_{M_{K(n)+i-1}}(p_{M_{K(n)+1},f}, X).$$

Since $\lim_k E_{M_{K(n)+k}}(f, X) = 0$, therefore we have

$$E_{M_{K(n)}}(f, X)\leq \sum_{i=1}^{\infty}E_{M_{K(n)+i-1}}(p_{M_{K(n)+i},f}, X).$$

According to Corollary 2 and (3) (iii) we get from (v) for $\varepsilon<r+\alpha$ that

$$E_{M_{K(n)}}(f, X)\leq \sum_{i=1}^{\infty}O_\varepsilon(M_{K(n)+i-1}^{-s+\varepsilon}\|d^{[s]}p_{M_{K(n)+i},f}\|_X) = O_\varepsilon(M_{K(n)}^{-r-\alpha+\varepsilon}) =$$

$$= (M_{K(n)}^{-r-\alpha+\varepsilon}) = O_\varepsilon(n^{-r-\alpha+\varepsilon}),\ (n\in\mathbf{N}),$$

as was stated.

Proof of (iii)$\Rightarrow$(i). By Lemma 1 we have

$$E_n(f, X)\leq E_{M_{K(n)}}(f, X)\leq \omega(f, X, M_{K(n)}^{-1}),\ (n\in\mathbf{N}).$$

Applying Corollary 1 and (3) (iii), (i) follows from (iii).

Proof of (iv)$\Rightarrow$(iii). f satisfies the conditions of Corollary 3, thus applying it we have that $d^{[r]}f\in X(G_m)$ and

$$E_n(d^{[r]}f, X) = O_\varepsilon(n^{-\alpha+\varepsilon}),\ (\varepsilon>0, n\in\mathbf{N}). \tag{8}$$

Since

$$\omega(d^{[r]}f, X, n^{-1})\leq \omega(d^{[r]}f, X, M_{K(n)}^{-1})\leq 2E_{M_{K(n)}}d^{[r]}f, X),$$

(see Lemma 1), thus from (8) and from (3) (iii) we obtain

$$\omega(d^{[r]}f, X, n^{-1}) = O_\varepsilon(n^{-\alpha+\varepsilon}), \quad (\varepsilon > 0, n \in \mathbf{N}).$$

The proof of (iv)$\Rightarrow$(iii) is complete.

The equivalence of (ii) and (iii) is a direct consequence of the definitions of Lip (α, X) and of $\omega(f, X, \delta)$.

The theorem is proved.

We remark that the proof of the theorem in the "bounded" case is similar to the method used in the case $m \in \mathcal{H}$, we need only to apply the appropiate part of the auxiliary lemmas and corollaries and the relation $n = O(M_{K(n)})$, $(n \in \mathbf{P})$.

References

[1] A. V. Efimov, On some approximation properties of periodic multiplicative orthonormal systems, *Mat. Sb.*, **69** (1966), 354–370.

[2] C. W. Onneweer, Differentiability for Rademacher series on groups, *Acta Sci. Math.*, **39** (1977), 121–128.

[3] J. Pál and P. Simon, On a generalization of the concept of derivative, *Acta Math. Acad. Sci. Hungar.*, **29** (1977), 155–164.

[4] T. S. Quek and L. Y. H. Yap, Absolute convergence of Vilenkin-Fourier series, *Journal of Math. Anal. and Appl.*, **74** (1980), 1–14.

[5] P. Simon, Investigations with respect to the Vilenkin systems, *Annales Univ. Sci. Budapest*, Sectio Mathematica, **27** (1984), 87–101.

[6] N. Ja. Vilenkin, On a class of complete orthonormal systems, *Izv. Akad. Nauk SSSR, Ser. Math.*, **11** (1947), 363–400; English transl., *Amer. Math. Soc. Transl.* (2) **28** (1963), 1–35.

[7] H. J. Wagner, *Ein Differential- und Integralkalkül in der Walsh-Fourier-Analysis mit Anwendungen* (Forschungber. des Landes Nordhein-Westfalen Nr.2 334), Westdutscher Verlag (Köln–Opladen), (1973), 71 pp.

A REMARK ON A PROBLEM OF GOEBEL

By

ADAM P. BOSZNAY

Department of Mathematics Faculty of Mechanical Engineering

Technical University of Budapest

(Received December 27, 1982)

Introduction

In [1], B. GOEBEL raised the following question:

Does there exist an open bounded and convex set G in the $\mathcal{H}$ real separable infinite dimensional Hilbert space such that for any $x \in G$ there is exactly one y point satisfying $y \in \partial G$ and

$$\|x-y\| = \text{dist}\ (x, \partial G)$$

where ∂G denotes the norm-boundary of G.

In this paper we show the nonexistence of such G – assuming some additional properties on G.

The result

We call the open bounded convex $\emptyset \neq G \subset \mathcal{H}$ set as a Chebyshev-type set, if for any $x \in G$ there is exactly one point $y \in \partial G$ satisfying

$$\|x-y\| = \text{dist}\ (x, \partial G).$$

THEOREM. *There exists no $0 \in G$ Chebyshev-type set in $\mathcal{H}$ with the following additional property:*

$$\partial G = \bigcup_{n=1}^{\infty} H_n,$$

where H_n is a convex closed subset of

$$\{x \in \mathcal{H};\ \langle x, \xi_n \rangle = 1\}.$$

Here $\xi_n \in \mathcal{H}$ fixed, and H_n has nonempty interior in the relative norm topology of

$$\{x \in \mathcal{H};\ \langle x, \xi_n \rangle = 1\}.$$

(Roughly speaking, we assume that $\overline{G}$ is the intersertion of a countable set of closed half-spaces. It is clear that $0 \in G$ is only a technical assumption.)

PROOF. Without loss of generality, we can assume that $\xi_n \neq \xi_m$ for $n \neq m$. We shall show the Theorem on an indirect way.

Let us denote the nearest element to $x \in G$ in ∂G, $P_G(x)$. By our assumption, for all $x \in G$ $P_G(x) \in H_n$. for some $n \in \mathbf{N}$. First, we prove that $x - P_G(x)$ is orthogonal to H_n for this $n = n(x)$, in the sense that for all $z \in H_n$

$$x - P_G(x) \perp z - P_G(x). \tag{1}$$

We shall prove (1) indirectly. By the indirect assumption, it is easy to prove that $P_G(x) \in \partial H_n$ where ∂H_n is the boundary of H_n in

$$\{x \in \mathcal{H};\ \langle x, \xi_n \rangle = 1\}.$$

It can be shown also easily that

$$\langle x - P_G(x),\ z - P_G(x) \rangle \leq 0 \tag{2}$$

for all $z \in H_n$. (In the other case we would have a nearer element in H_n to x than $P_G(x)$.)

By (2),

$$\langle x - P_G(x), P_G(x) - z \rangle \geq 0$$

for all $z \in H_n$, and using the indirect hypothesis, there exists $z \in H_n$ with

$$\langle x - P_G(x),\ P_G(x) - z \rangle > 0.$$

This implies for sufficiently small $\varepsilon > 0$

$$\|x - (P_G(x) + \varepsilon(P_G(x) - z))\| < \|x - P_G(x)\|, \tag{3}$$

so, $P_G(x) + \varepsilon(P_G(x) - z)$ is nearer to x than $P_G(x)$. Clearly,

$$P_G(x) + \varepsilon(P_G(x) - z) \notin G \tag{4}$$

because of the fact that $P_G(x) - \varepsilon(P_G(x) - z) \in \overline{G}$ and $P_G(x) \in \partial G$ (here we have used that G is open and convex)

Because of (4), on the interval

$$[x,\ P_G(x) + \varepsilon(P_G(x) - z)]$$

there exists a point of ∂G, and for this y^* point we have by (3)

$$\|x - y^*\| < \|x - P_G(x)\|,$$

which is a contradiction. So, (1) is true for all $z \in H_n$.

Using (1) and the fact that G is open, clearly we have

$$P_G^{-1}(H_n) \cap P_G^{-1}(H_m) = \emptyset$$

or $n \neq m$. (Here we apply $\xi_n \neq \xi_m$.)

We prove now that $P_G^{-1}(H_n)$ is closed in G.

Let $x_1, \ldots, x_{K1}, \ldots \in P_G^{-1}(H_n)$,

$$\lim_{k \to \infty} x_k = x_0, x_0 \in G \tag{6}$$

then using (1) and the fact that H_n has nonempty interior in the relative topology of

$$\{x \in \mathcal{H};\ \langle x, \xi_n \rangle = 1\},$$

$$\lim_{K \to \infty} P_G(x_k) = x^* \in H_n. \tag{7}$$

Using the well-known closedness of P_G, we have $P_G(x_0) = x^*$ so, $P_G^{-1}(H_n)$ is closed.

We receive that the convex G set is the union of countable many pairwise disjoint closed sets.

Now, we can use the following result due to ASPLUND [2]:

Let G be a convex set is a Banach space, and G is the union of countably many disjoint closed sets. Then all but one of these sets are empty.

This implies that

$$P_G^{-1}(H_{n^*}) = G$$

for some $n^* \in \mathbf{N}$, which is a contradiction.

The Theorem is proved.

References

[1] Open problems, presented at the third seminar (Poland – GDR) on operator ideals and geometry of Banach spaces, Georgenthal, April 2 – 11, 1979., *Math., Nachrichten*, **95** (1980), 299 – 303.

[2] ASPLUND, E.: Sets with unique farthest points, *Israel J. Math.*, **5** (1967), 201 – 209.

REMARKS ON SUPERLINEAR OPERATORS

By

A. BOGMÉR, I. JOÓ and L. L. STACHÓ

Bolyai Institute of the József Attila University, Szeged and
II. Department for Analysis of the L. Eötvös University, Budapest

(Received December 30, 1982)

In [1] E. M. NIKIŠIN introduced the notions of superlinear and positive superlinear operators concerning his investigation on Fourier series with respect to general orthonormal systems. According to [1], a mapping $T: E \to S(0, 1)$ where E is any Banach space is by definition superlinear if for every $e \in E$ there exists a linear mapping $L_e : E \to S(0, 1)$ such that $L_e e = Te$ and $|L_e f| \leq |Tf|$ for each $f \in E$. Furthermore, if $E = \mathcal{L}^p(X, \mu)$ for some $p \geq 1$ and for every $e \in E$, L_e can be chosen to be a positive linear mapping then T is called a positive superlinear operator ($\mathcal{L}^p \to S$).

The aim of this paper is to examine these concepts in a vector lattice theoretical setting.

DEFINITION 1. Let E, F be a vector space and a vector lattice, respectively. A mapping $T: E \to F$ is *superlinear* if for every $e \in E$ there exists $L_e \in \mathcal{L}(E, F)$ such that $L_e e = Te$ and $|L_e| \leq |T|$. (Throughout this work, we deal with real vector spaces. The symbol $|T|$ means the operator $f \to |Tf|$.)

PROPOSITION 1. Suppose the space F is oder complete (for def. see [3]). Then $T: E \to F$ is superlinear if and only if $|T|$ is a vector norm on E i.e. if, $|T(e_1+e_2)| \leq |Te_1| + |Te_2|$ and $|T\lambda e_1| = |\lambda|\, |Te_1|$ $\forall\, e_1, e_2 \in E$, $\lambda \in \mathbf{R}$.

PROOF. Let $T: E \to F$ be superlinear. Then we can write $|T| = \sup\limits_{e \in E} |L_e|$. But L is clearly a vector norm whenever $L: E \to F$ is linear.

Conversely, assume T is a vector norm on E and F is order complete. Given $e \in E$, define L_e^0 on the subspace $\mathbf{R}_e$ by $L_e^0 \lambda e \equiv \lambda Te$ ($\lambda \in \mathbf{R}$). We have $L_e^0 \leq |T|$ on $\mathbf{R}_e$. Thus by the generalized Hahn-Banach theorem [2], L_e^0 admits a linear extension L_e such that $L_e \leq |T|$. To complete the proof, we show $-L_e \leq |T|$. Indeed, $-L_e f = L_e(-f) \leq |T(-f)| = |Tf|$ $\forall\, f \in E$.

DEFINITION 2. Let E, F be vector lattices. A mapping $T: E \to F$ is positive *superlinear* if for every $e \in E$ there exists $L_e \in \mathcal{L}_+(E, F)$ (i. e. $L_e p \geq 0$ whenever $p \in E_+$ (i. e. $p \geq 0$ in E) such that $L_e\, e = Te$ and $|L_e| \leq |T|$.

THEOREM 1. *Let E, F be vector lattices, $T: E \to F$ a superlinear operator. Assume that the ordering of F is complete. Then equivalent are*

(a) *T is positive superlinear.*
(b) *$e_1 \leq e_2$ implies $Te_1 \leq |Te_2|$ and $-Te_2 \leq |Te_1|$ for all $e_1, e_2 \in E$.*
(c) *By setting $P \equiv |T|$ and $Qe \equiv \inf_{p \in E_+} P(e+p)$, $(e \in E)$, we have*
$Qe \geq (Te) \vee ((-T-e))$ for all $e \in E$.

PROOF. (a)$\Rightarrow$(b): Suppose T is superlinear and $e_1 \leq e_2$ in E. Then choosing L_{e_1}, L_{e_2} in accordance with Definition 2, we obtain

$$Te_1 = L_{e_1} e_1 \leq L_{e_1} e_1 \leq |L_{e_1} e_2| \leq |Te_2|$$

and

$$-Te_2 = -L_{e_2} e_2 = L_{e_2}(-e_2) \leq L_{e_2} e_1 \leq |L_{e_2} e_1| \leq |Te_1|.$$

(b)$\Rightarrow$(c): Let p be any element of E_+ and $e \in E$. An application of (b) to $e_1 \equiv e$ and $e_2 \equiv e+p$ yields $P(e+p) = |T(+e+p)| \geq Te$. Similarly, if $e_1 \equiv \equiv -e-p$, $e_1 \equiv -e$ we have $P(e+p) = |T(e+p)| = |T(-e-p)| \geq -T(-e)$.

(c)$\Rightarrow$(a): By assumption, P is a vector norm on E. Hence for any α_1, $\alpha_2 \in \mathbf{R}_+$, e_1, $e_2 \in E$ and p_1, $p_2 \in E_+$,

$$\sum_{1,2} \alpha_j P(e_j + p_j) = \sum_{1,2} P(\alpha_j e_j + \alpha_j p_j) \geq P\Big(\sum_{1,2} \alpha_j e_j + \sum_{1,2} \alpha_j p_j\Big).$$

Since $\sum_{1,2} \alpha_j p_j \in E_+$, too, it follows that Q is also a vector norm on E Let now $e \in E \setminus \{0\}$ be arbitrarily given and define $L_e^0: \mathbf{R}e \to F$ by $L_e^0(\lambda e) \equiv \lambda\, Te$ $(\lambda \in \mathbf{R})$. Observe that $Q(\lambda e) = \lambda Qe \geq$ by (c) $\geq \lambda((Te) \vee (-T(-e))) \geq \lambda Te = = L_e^0(\lambda e)$ if $\lambda \geq 0$ and $Q(\lambda e) = |\lambda| Q(-e) \geq$ by(c) $\geq |\lambda|\ (T(-e) \vee (-Te)) = = (|\lambda| T(-e)) \vee (\lambda Te) \geq Te = L_e^0(\lambda e)$ if $\lambda \leq 0$. Thus $L_e^0 \leq Q$ on $\mathbf{R}e$. By the generalized Hahn–Banach theorem [2], L_e^0 admits a linear extension L_e to E such that $L_e \leq Q$.

Clearly $L_e \leq P$ since $Q \leq P$. On the other hand, $L_e \in \mathcal{L}_+(E, F)$ since $L_e(-p) \leq Q(-p) \leq P(p-p) = 0$ for all $p \in E_+$.

Next we turn our attention to the continuity properties of positive superlinear operators. It seems that those ranging in $S(0, 1)$ (as in Nikisin's original definition) are of particular importance among them because, as we see, positive superlinear operators between locally convex topological vector lattices are very rarely continuous unless being linear.

LEMMA 1. If E, F are vector lattices and $T: E \to F$ is positive superlinear then T is convex, positive homogeneous and positive valued when restricted to E_+, furthermore $T(-p) = -T_{(p)}$ for all $p \in E_+$.

PROOF. If $p \in E_+$ then $L_p\, p$, $L_{-p}\, p \geq 0$. But $L_p\, p = T_p$ and $L_{-p}(-p) = = T_{(-p)}$ whence $T_p = |T_p|$ and $T(-p) = -|T(-p)| = -|T_p| = -T_p$. Thus T coincides with $|T|$ on E_+. This implies its convexity and homogeneity on E_+ since T is a vector norm.

LEMMA 2. Suppose $T : \mathbf{R}^2 \to \mathbf{R}$ is a continuous positive superlinear map ping. Then T is necessarily linear.

PROOF. We may assume also $T \neq 0$. Then we have $T(\lambda_1, \lambda_2) \neq 0$ for all $\lambda_1, \lambda_2 < 0$. Indeed, $T(\lambda_1, \lambda_2) = 0 < \lambda_1, \lambda_2$ would imply $|L_e(\lambda_1, \lambda_2)| \leq |T(\lambda_2, \lambda_2)| = 0$ i.e. $0 = L_e(\lambda_1, \lambda_2) = \lambda_1 L_e(1, 0) + \lambda_2 L_e(0, 1)$ and hence $L_e = 0$ (for $L_e \in \mathcal{L}_+(\mathbf{R}^2, \mathbf{R})$ by Definition 2) for all $e \in \mathbf{R}$. Thus, by Lemma 1, range T contains both positive and negative numbers. Since T is a vector norm, this means that the set $\mathcal{N} \equiv \{e : Te = 0\}$ is a 1 dimensional subspace of $\mathbf{R}^2$, disjoint from $(0, \infty) \times (0, \infty)$. Therefore we can find a linear functional $\varphi \in \mathcal{L}_+(\mathbf{R}^2, \mathbf{R})$ such that $\varphi(1, 1) = T(1, 1) > 0$ and $\mathcal{N} = \{e : \varphi e = 0\}$. Let $f \in \mathbf{R}^2$ be arbitrarily fixed and consider the linear functional L_f. Since $|L_f| \leq |T|$, we have $\{e : L_f e = 0\} \supset \mathcal{N}$. Hence for some $\lambda_f \in \mathbf{R}_+$, $L_f = \lambda_f \varphi$. To conclude, we prove $\lambda_f = \lambda_g$ for all $f\, g \notin \mathcal{N}$. We may assume $0 < \lambda_f \leq \lambda_g$. Then $\lambda_g |\varphi f| = |L_g f| \leq |Tf| = |L_f f| = \lambda_f |\varphi f|$ whence $\lambda_f = \lambda_g$ completing the proof.

THEOREM 2. *Let E, F be topological vector lattices and let $F_0^* \equiv \{\varphi \in F_+^*) \varphi|f| = |\varphi f| \ \forall f \in F\}$. If F_0^* sparates the points of F then each continuous positive superlinear map $T : E \to F$ is linear.*

PROOF. Let us fix any $\varphi \in F_0^*$. Observe that the functional $\varphi \circ T$ is also positive superlinear ($E \to \mathbf{R}$). In fact, given $e_1 \leq e_2$, from Theorem 1. (b) we obtain $Te_1 \leq |Te_2|$, $-Te_1 \leq |Te_2|$ whence $\varphi Te_1 \leq \varphi |Te_2| = |\varphi Te_2|$ and $-\varphi Te_2 \leq \varphi |Te_1| = |\varphi Te_1|$. Now from Lemma 2. we see that for any p_1, $p_2 \in E_+$, the functional $\mathbf{R}^2 \ni (\varrho_1, \varrho_2) \to \varphi T(\varrho_1 p_1 + \varrho_2 p_2)$ is linear. Thus for all $e_1, e_2 \in \mathbf{R}p_1 + \mathbf{R}p_2$ and $\lambda \in [0, 1]$,

$$\Phi\left(T\left(\frac{1}{2}e_1 + \frac{1}{2}e_2\right) - \frac{1}{2}Te_1 - \frac{1}{2}Te_2\right) = 0.$$

Since F_0^* separates F, it follows

$$T\left(\frac{1}{2}e_1 + \frac{1}{2}e_2\right) = \frac{1}{2}Te_1 + \frac{1}{2}Te_2$$

i.e. the mapping T is linear when restricted to any 2 dimensional subspace of E spanned by positive elements. Thus if $e, f\, E$ then

$$T\left(\frac{1}{2}e + \frac{1}{2}f\right) = T\left(\frac{1}{2}(e_+ + f_+) + \frac{1}{2}(-e_- - f_-)\right) = \frac{1}{2}T(e_+ + f_+) +$$

$$+ \frac{1}{2}T(-e_- - f_-) = \frac{1}{2}T(e_+ + f_+) - \frac{1}{2}T(e_- + f_-) = \left(\frac{1}{2}Te_+ + \frac{1}{2}Tf_+\right) -$$

$$-\left(\frac{1}{2}Te_- + \frac{1}{2}Tf_-\right) = \frac{1}{2}(Te_+ - Te_-) + \frac{1}{2}(Tf_+ - Tf_-) = \frac{1}{2}Te + \frac{1}{2}Tf$$

establishing the linearity of T.

COROLLARY 1. If E is a topological vector lattice and Ω is a compact topological space then each continuous positive superlinear map $E \to C(\Omega)$ is linear.

PROOF. The functionals $\delta_x \equiv [C(\Omega) \in f \to f(x)]$ $(x \in \Omega)$ form a separating family in $C(\Omega)$ and satisfy $\delta_x |f| = |f(x)| = |\delta_x f|$.

COROLLARY 2. If E is a topological vector lattice and μ is an arbitrary measure then each continuous positive superlinear map $E \to L^\infty(\mu)$ is linear.

PROOF. By Kakutani's representation theorem on M-lattices [3], each L^∞ – space is isometrically order isomorphic to some $C(\Omega)$ space for suitable compact topological space.

COROLLARY 3. If E is a topological vector lattice and $1 \leq p \leq \infty$ then every continuous positive superlinear map $E \to l^p$ is linear.

PROOF. Every continuous superlinear operator $T : E \to l^p$ can be viewed as a continuous positive superlinear $E \to l^\infty$. mapping.

The following question arises from the above corollaries: Is there any non-linear continuous positive superlinear operator L^p (0, 1) $\to L^q$(0. 1) if $q < \infty$? The answer is always affirmative in this case.

EXAMPLE. Let $1 \leq p \leq \infty$ and $1 \leq q < \infty$. The mapping $T : L^p(0, 1) \to \to L^q(0, 1)$ defined by

$$Tf \equiv [(0,1) \ni t \to \begin{cases} t \int_0^{1/2} f \quad if \quad \left| t \int_0^{1/2} f \right| \geq \left| (1-t) \int_{1/2}^1 f \right| \\ (1-t) \int_{1/2}^1 f \quad \text{else} \end{cases}$$

is positive superlinear and continuous but non-linear.

PROOF. The non-linear character of T is obvious.

Continuity: Suppose $f_n \to f$ in $L^p(0, 1)$ $(n \to \infty)$.

Now

$$\int_0^{1/2} f_n \to \int_0^{1/2} f \quad \text{and} \quad \int_{1/2}^1 f_n \to \int_{1/2}^1 f, \ (n \to \infty).$$

Hence $Tf_n(t) \to Tf(t)$ whenever

$$\left| t \int_0^{1/2} f \right| \neq \left| (1-t) \int_{1/2}^1 f \right| \quad \text{or} \quad \left| t \int_0^{1/2} f \right| = \left| (1-t) \int_{1/2}^1 f \right| = 0.$$

i. e. almost everywhere. Since the sequence $\{|Tf_n|\}_1^\infty$ consists of functions majorized by the constant $\sup_n \int_0^1 |f_n|$, it follows

$$\|Tf_n - Tf\|_{L^q} = \left(\int_0^1 |Tf_n(t) - Tf(t)|^q \, dt \right)^{1/p} \to 0, \ (n \to \infty).$$

Positive superlinearity: Given $e \in L^p(0, 1)$, it is immediate that the linear mapping $L_e : L^p(0, 1) \to L^q(0, 1)$ defined by

$$L_e f \equiv [(0,1) \ni t \to \begin{cases} t \int_0^{1/2} f \text{ if } \left| t \int_0^{1/2} e \right| \geq \left| (1-t) \int_{1/2}^{1} e \right| \\ (1-t) \int_{1/2}^{1} f \text{ else} \end{cases}$$

is positive and fulfills the requirements of Definition 2.

References

[1] E. M. Никишин, Резонансные теоремы и надлинейные операторы, *Успехи Матем. Наук*, **25** (1970), 129–191.

[2] B. L. Pinto, Banach Extension Theorem for Ordred-Complete Linear Spaces, *Bolletino U. M. I.*, **6** (1972), 181–184.

[3] H. H. Schaeffer, *Banach Lattices and Positive Operators*, Springer Verlag, Berlin–New York, 1976.

[4] A. Sövegjártó, Remark to a paper of E. M. Nikišin, *Annales Univ. Sci. Budapest*, Sect. Math., **22–23** (1979–1980), 135–138.

SOME RESULTS ON LUCAS PSEUDOPRIMES

By

PÉTER KISS

Department of Mathematics, Teacher's Training College, Eger

(Received March 18, 1983)

A Lucas sequence of integers $R = \{R_n\}_{n=1}^{\infty}$ is defined by the recursion

$$R_n = A \cdot R_{n-1} - B \cdot R_{n-2}$$

for $n > 1$, where A and B are fixed integers and the initial terms are $R_0 = 0$ and $R_1 = 1$. Let α and β be the roots of the polynomial

$$f(x) = x^2 - Ax + B$$

and we denote the discriminant of $f(x)$ by D. Thus

$$D = A^2 - 4B = (\alpha - \beta)^2.$$

Throughout this paper we suppose that $AB \neq 0$, $(A, B) = 1$ and α/β is not a root of unity. In this case we say the Lucas sequence is non degenerate.

It is well-known that for an odd prime number n with $(n, B) = 1$ we have

$$n \mid R_{n-(D/n)}, \tag{1}$$

where (D/n) is the Jacobi symbol. If (1) holds for a composite integer n then it is called a Lucas pseudoprime number with respect to the sequence R. The Lucas pseudoprimes are the generalizations of the pseudoprimes with respect to an integer $b(\geqq 2)$. Namely a composite integer n is called pseudoprime number with respect to the integer b if $n \mid (b^{n-1} - 1)$ and one can easily see, using the explicit form

$$R_n = \frac{\alpha^n - \beta^n}{\alpha - \beta} \tag{2}$$

for the terms of Lucas sequences, that a pseudoprime number n with respect to b is a pseudoprime with respect to the Lucas sequence defined by the constants $A = b+1$, $B = b$, if $(n, b-1) = 1$. The pseudoprime numbers with respect to 2 are called pseudoprimes. Furthermore if n is a pseudoprime and every divisor of n is a prime or a pseudoprime then we say n is a super pseudoprime.

The pseudoprimes and the Lucas pseudoprimes have been studied intensively, since they can be well used for tests for primality (see e.g. R. BAILLIE and S. S. WAGSTAFF JR [1] and the references there). A lot of results on them, up to 1971, were collected by E. LIEUWENS [8] and A. ROTKIEWICZ [13].

In this paper we give some new results on Lucas pseudoprimes and shall use these results to study the distribution of them.

Let $\Theta_b(x)$, $\Theta(x)$ and $R(x)$ denote the number of pseudoprimes with respect to b, with respect to 2 and with respect to the Lucas sequence R, respectively, not exceeding the real number x. C. POMERANCE showed that for all large x

$$\Theta_b(x) \geqq \exp\{(\log x)^{5/14}\}$$

for any base b (in [11]) and

$$\Theta(x) \leqq x \cdot [\exp\{\log x \log\log\log x / \log\log x\}]^{-1/2}$$

(in [12]), which were generalizations and improvements of the results of P. ERDŐS [2] and D. H. LEHMER [7] respectively. R. BAILLIE and S. S. WAGSTAFF, JR [1] proved that there are positive constants c_1 and c_2 such that for all large x

$$R(x) < x \cdot \exp\{-c_1 (\log x \log\log x)^{1/2}\}$$

for any non degenerate Lucas sequence R, and

$$R(x) > c_2 \log x$$

for sequences R for which $D > 0$ but D is not a perfect square. Now we extend this lower bound for all non degenerate Lucas sequence.

THEOREM 1. *Let R be a non degenerate Lucas sequence. Then there exists a positive constant c_3, depending only on the parameters of the sequence R, such that for all large x*

$$R(x) > c_3 \cdot \log x.$$

It is known that the number of the super pseudoprime numbers is infinite. For example K. SZYMICZEK [17] showed that $F_n \cdot F_{n+1}$ is a super pseudoprime for $n > 1$, where $F_n = 2^{2^n} + 1$ is the n-th Fermat number. In a joint paper [3] written by J. FEJÉR we proved: If b is an integer with $b > 1$ and $4 \nmid b$, then there exist infinitely many super pseudoprime numbers with respect to b which are products of exactly three distinct primes.

The definition of super pseudoprimes can be extended for Lucas pseudoprimes, too. A composite integer n is called a super Lucas pseudoprime number with respect to a sequence R if every divisor of n (greater than one) is a prime or a Lucas pseudoprime with respect to the sequence R. We note, if n is a super Lucas pseudoprime, then it is a Lucas pseudoprime as well, since n divides itself.

Since $|R_{2p}/A|$ is a composite integer for infinitely many prime p, the next theorem shows that there are infinitely many super Lucas pseudoperimes.

THEOREM 2. *Let R be a non degenerate Lucas sequence defined by the constants A and B. Then $|R_{2p}/A|$ is a super Lucas pseudoprime with respect to the sequence R for any prime p with $p>c_4$, where c_4 is a positive constant depending only on A and B.*

In order to formulate an other result, we introduce some notations. Let R be a non degenerate Lucas sequence defined by constant integers A and B. If n is an integer with $(n, B)=1$ then there are terms in R divisible by n. The least positive integer r, for which $n|R_r$, is called the rank of apparition of n in the sequence R and we shall denote it by $r(n)$. Thus $n|R_{r(n)}$ but $n\nmid R_m$ for $0<m<r(n)$. If $n=p$ is a prime then p is called a primitive prime divisor of $R_{r(p)}$, or more exactly, p is a primitive prime divisor of the term R_n if $p|R_n$ but $p\nmid BDR_m$ for $0<m<n$.

It is known that there is an absolute constant n_0 (>4) such that R_n has at least one primitive prime divisor for any $n\geqq n_0$ (see. A. SCHINZEL [15] or C. L. STEWART [16]). In the followings we denote the product of the primitive prime power divisors of R_n by $\mathcal{R}_n$, where a primitive prime power divisor of R_n means a prime power p^t for which p is a primitive prime divisor of R_n and $p^t\|R_n$. For $n\geqq n_0$ we have $\mathcal{R}_n>1$.

We shall prove:

THEOREM 3. *Let R be a non degenerate Lucas sequence defined by constants A and B let $p>\max(|B|, |D|, n_0)$ be a prime for which $r(p)\neq p-(D/p)$. Then $p\mathcal{R}_{p-(D/p)}$ is a super Lucas pseudoprime with respect to the sequence R. Furthermore the consition $r(p)\neq p-(D/p)$ holds for infinitely many primes p.*

Our theorems imply some results on the distribution of super Lucas pseudoprimes. It is known that the sum of the reciprocals of all Lucas pseudoprimes with respect to a Lucas sequence is convergent (see in [1]). But A. MAKOWSKI [9] showed that if $P_1<P_2<P_3<\dots$ is the sequence of the pseudoprime numbers then the sum $\sum_{i=1}^{\infty} 1/\log P_i$ is divergent. Furthermore A. ROTKIEWICZ [14] proved that for given $\varepsilon>0$ there is a pseudoprime between x and $x^{1+\varepsilon}$ provided $x>x_0(\varepsilon)$. These results can be extended for super Lucas pseudoprimes and so for Lucas pseudoprimes, too.

COROLLARY 1. Let R be a non degenerate Lucas sequence and let $P_1<P_2<P_3<\dots$ be the sequence of the super Lucas pseudoprime numbers with respect to R. Then the sum $\sum_{i=1}^{\infty} 1/\log P_i$ is divergent.

COROLLARY 2. Let R be a non degenerate Lucas sequence defined by constants A and B. Then, for any $\varepsilon>0$, there exists a real number $x_0=x_0(\varepsilon, A, B)$ such that the interval $(x, x^{1+\varepsilon})$ contains a super Lucas pseudoprime with respect to R provided $x>x_0$.

We list some properties of non degenerate Lucas sequences, defined by constants A and B, which will be used in the proofs of the theorems. Let n, m, k, e be positive integers and let q be a prime suc h that $(q, B) = 1$. Using the notations defined above, we have

(i) $r(q)|(q-(D/q))$, supposing that $(D/q) = 0$ if $q|D$.
(ii) $q|R_n$ if and only if $r(q)|n$.
(iii) If $q^e \| R_{r(q)}$ then $r(q^k) = q^{k-e}\, r(q)$ for $k \geqq e$.
(iv) $R_n | R_{nm}$.
(v) $(R_n, R_m) = R_{(n,m)}$.

(For these properties of Lucas sequences we refer to D. H. LEHMER [6]).

For the proof of Theorem 1 we need a lemma.

LEMMA. Let p and q be distinct odd primes. $n = pq$ is a Lucas pseudoprime with respect to a Lucas sequence R if and only if $r(p)|(q-(D/q))$ and $r(q)|(p-(D/p))$.

PROOF of the LEMMA. The integer $n = pq$ is a Lucas pseudoprime if and only if $n|R_{n-(D/n)}$. But

$$n-(D/n) = pq-(D/pq) =$$
$$= (p-(D/p))(q-(D/q))+(D/p)(q-(D/q))+(D/q)(p-(D/p))$$

and by (i), $r(p)|(p-(D/p))$ and $r(q)|(q-(D/q))$, therefore (ii) implies the statement.

PROOF of THEOREM 1. Let $n > 2n_0$ be and integer with condition $n \equiv 2 \pmod 4$. Thus $n = 2k$, where k is an odd integer. Let p and q be a primitive prime divisor of R_{2k} and R_k, respectively. By the definition of $r(n)$ we have $r(p) = 2k$ and $r(q) = k$ and so, by (i),

$$2k|(p-(D/p)) \tag{3}$$

and

$$k|(q-(D/q)). \tag{4}$$

But k is odd and $q-(D/q)$ is even, therefore (3) and (4) imply the relations $r(q) = k|(p-(D/p))$ and $r(p) = 2k|(q-(D/q))$. It shows by the Lemma that $n = pq$ is a Lucas pseudoprime with respect to R.

(2) implies that there exists a positive constant c_5 depending only on the constants of the sequence R such that

$$|R_n| < e^{c_5 n}$$

for all $n \geqq 1$. Let x and y be real numbers with conditions $2n_0 < y = \dfrac{1}{2c_5} \log x$. There are at least $\left[\dfrac{y-2n_0}{4}\right]$ integers n for which $n \leqq y$, $n \equiv 2 \pmod 4$ and both

R_n and $R_{n/2}$ have primitive prime divisors. By using the result proved above, hence it follows that for sufficiently large y there are at least $\frac{1}{5}y = \frac{1}{10c_5}\log x$ Lucas pseudoprimes with respect to R such that they do not exceed $|R_{[y/2]}\cdot R_{[y]}|$. However

$$|R_{[y/2]}\cdot R_{[y]}| < e^{c_5(y+y/2)} < e^{2c_5 y} = x$$

which proves the theorem with $c_3 = 1/(10c_5)$.

Proof of Theorem 2. Let R be a non degenerate Lucas sequence and let p be a prime with $p > \max(n_0, A, D)$. The number R_{2p}/A is an integer since by (iv) $A = R_2|R_{2p}$, furthermore R_{2p}/A is odd by the properties (ii), (iii) and (v), since $r(2) = 2$ or $r(2) = 3$ if there exist terms in the sequence R divisible by 2. The number R_{2p}/A is composite namely it is divisible by R_p and both R_p and R_{2p} has primitive prime divisors for $p > n_0$.

We have to prove that $Q|R_{Q-(D/Q)}$ for any divisor Q of R_{2p}/A.

Let $Q > 1$ be an integer which divides the number R_{2p}/A.This divisor can be written in the form $Q = q_1\cdot q_2\ldots q_s$, where the q_i's ($i = 1, 2, \ldots, s$) are primes. $r(q_i)\neq 2$, namely $r(q_i) = 2$ would imply that $q_i^e\|A = R_2$ and $q_i^{e+1}|R_{2p}$, which contradict to (iii) since $q_i|A$ and $q_i|p$ are impossible by the condition $p > A$. Thus $r(q_i) = p$ or $r(q_i) = 2p$ for $i = 1, 2, \ldots, s$. So, by (i), we have

$$Q = \prod_{i=1}^{s} q_i = \prod_{i=1}^{s} \left(k_i p + (D/q_i)\right) = kp + (D/Q)$$

and

$$Q - (D/Q) = kp \tag{5}$$

with some integers $k_1, k_2, \ldots, k_s$ and k. As we have seen R_{2p}/A and so Q are odd integers. Furthermore $(D/q_i)\neq 0$ for $i = 1, 2, \ldots, s$, since otherwise $q_i|D$ and $r(q_i) = q_i$, and so by (ii) we should have $q_i = p$, which contradicts to the condition $p > D$. Therefore, by (5), kp is an even integer. This and (iv) imply the divisibility

$$R_{2p}|R_{Q-(D/Q)},$$

which proves the statement since $Q|R_{2p}$.

Proof of Theorem 3. Let $Q = q_1\cdot q_2\ldots q_s > 1$, where the q_i's are odd primes and let us suppose that $Q|\mathcal{R}_{p-(D/p)}$. We have to show that $n|R_{n-(D/n)}$ for $n = p^e\cdot Q$, where $e = 0$ or 1.

By the conditions and the properties of the sequence R mentioned above we have $r(q_i) = p-(D/p)$ for any $1 \leqq i \leqq s$, therefore (i) implies the equality

$$q_i = k_i(p-(D/p)) + (D/q_i) \tag{6}$$

with some integer k_i. Using (6) and supposing that $(D/p^e) = 1$ in case $e = 0$, we get

$$n-(D/n) = p^e \prod_{i=1}^{s} [k_i(p-(D/p))+(D/q_i)] - (D/p^e)\prod_{i=1}^{s}(D/q_i) =$$

$$= p^e k(p-(D/p)) + p^e \prod_{i=1}^{s}(D/q_i) - (D/p^e)\prod_{i=1}^{s}(D/q_i) =$$

$$= p^e k(p-(D/p)) + (p^e-(D/p^e))\prod_{i=1}^{s}(D/q_i),$$

where k is an integer. From this it follows that $(p-(D/p))|(n-(D/n))$ which, together with (iv), implies the divisibility

$$R_{p-(D/p)}|R_{n-(D/n)}.$$

But in the case $r(p)\neq p-(D/p)$ we have $(p, \mathcal{R}_{p-(D/p)}) = 1$ by the definition of $\mathcal{R}_m$; furthermore $p|R_{p-(D/p)}$ and $\mathcal{R}_{p-(D/p)})|R_{p-(D/p)}$, therefore $n|R_{n-(D/p)}$. Thus n is a Lucas pseudoprime with respect to the sequence R and so $p\mathcal{R}_{p-(D/p)}$ is a super one.

We have to prove yet that there are infinitely many prime p satisfying the conditions of the theorem. In [4], with B. M. PHONG, we have showed that the function $g(p) = (p-(D/p))/r(p)$ is unbounded for primes $p>B$. This result implies that $(p-(D/p))/r(p)>1$ for infinitely many prime p which completes the proof of Theorem 3.

PROOF of COROLLARY 1. We can suppose that $|\alpha|\geqq|\beta|\geqq 1$ in (2) and $|\alpha|>1$ since α/β is not a root of unity. Therefore there is a positive constant c_6 such that

$$|R_n|<c_6|\alpha|^n \tag{7}$$

for $n>0$. Using it, by Theorem 2 we get

$$\sum_{i=1}^{\infty}\frac{1}{\log P_i} > \sum_{p>c_4}\frac{1}{\log|R_{2p}/A|} > \sum_{p>c_4}\frac{1}{2p\cdot\log|\alpha|+c_7} > c_8\cdot\sum_{p>c_4}\frac{1}{p},$$

where c_7 and c_8 are positive constants depending only on the parameters of the sequence R. Hence the assertion follows.

PROOF of COROLLARY 2. Using Theorem 2, it is enough to show that if $x = |R_{2p_i}/A|$, then $R_{2p_{i+1}}/A|<x^{1+\varepsilon}$ for sufficiently large consecutive primes p_i and p_{i+1}. Let $p_i = p$ and $p_{i+1} = q$. It is known that there is a real number $0<\Theta<1$ such that $q<p+p^{\Theta}$ for sufficiently large p and we may take $\Theta = 17/31$ (see J. PINTZ [10]). This and (7) imply the inequalities

$$|R_{2q}/A|<c_6|\alpha|^{2q} = |\alpha|^{2q+c_9}<|\alpha|^{2p+2p^{\Theta}+c_9} \tag{8}$$

with some constant c_9. We showed in [5] (Theorem 1) that there is a positive constant c_{10} such that

$$|R_n|>|\alpha|^{n-c_{10}\log n}$$

and so

$$x = |R_{2p}/A| > |\alpha|^{2p - c_{10} \log 2p} \tag{9}$$

for $p > p_0$. Furthermore $p^{\Theta} < \delta p$ for any $\delta < 0$ if p is sufficiently large, therefore by (8) and (9) we have

$$|R_{2q}/A| < x \cdot |\alpha|^{2\delta p + c_{10} \log 2p + c_9} = x \cdot |\alpha|^{(2p - c_{10} \log 2p) \cdot \gamma} < x \cdot x^{\gamma},$$

where

$$\gamma = \frac{2\delta p + c_{10} \log 2p + c_9}{2p - c_{10} \log 2p} < 2\delta$$

for sufficiently large p. Hence the statement follows with $\delta = \varepsilon/2$.

References

[1] R. Baillie and S. S. Wagstaff, Jr., Lucas pseudoprimes, *Math. Comp.*, **35** (1980), 1391–1417.

[2] P. Erdős, On pseudoprimes and Carmichael numbers, *Publ. Math. Debrecen*, **4** (1956), 201–206.

[3] J. Fehér and P. Kiss, Note on super pseudoprime numbers, *Annales Univ. Sci. Budapest*, Sect. Math., **26** (1983), 157–159.

[4] P. Kiss and B. M. Phong, On a function concerning second order recurrences, *Annales Univ. Sci. Budapest*, Sect. Math., **21** (1978), 119–122.

[5] P. Kiss, Zero terms in second order linear recurrences, *Math. Sem. Notes Kobe Univ.*, **7** (1979), 145–152.

[6] D. H. Lehmer, An extended theory of Lucas' function, *Ann. of Math.*, **31** (1930), 419–448.

[7] D. H. Lehmer, On the converse of Fermat's theorem, *Amer. Math. Monthly*, **43** (1936), 347–354.

[8] E. Lieuwens, Fermat pseudo primes, Doctor thesis, Delft, 1971.

[9] A. Makowski, On a problem of Rotkiewicz on pseudoprime numbers, *Elem. Math.*, **29** (1974), 13.

[10] J. Pintz, On primes in short intervals I and II, *Studia Sci. Math. Hung.*, **16** (1981), 395–414 and **19** (1984), 89–96.

[11] C. Pomerance, A new lower bound for the pseudoprime counting function, *Illinois J. Math.*, **26** (1982), 4–9.

[12] C. Pomerance, On the distribution of pseudoprimes, *Math. Comp.*, **37** (1981), 587–593.

[13] A. Rotkiewicz, *Pseudoprime numbers and their generalizations*, University of Novi Sad, 1972.

[14] A. Rotkiewicz, Les intervales contenants les nombres pseudopremiers, *Rend. Circ. Mat. Palermo*, **14** (1965), 278–280.

[15] A. Schinzel, Primitive divisors of the expression $A^n - B^n$ in algebraic number fields, *J. reine angev. Math.*, **268/269** (1974), 27–33.

[16] C. L. Stewart, *Primitive divisors of Lucas and Lehmer numbers, Transcendence theory: Advances and applications*, Acad. Press, London – New York – San Francisco, 1977, 79–92.

[17] K. Szymiczek, Note on Fermat numbers, *Elem. Math.*, **21** (1966), 59.

ON A PROPERTY OF THE EIGENFUNCTIONS OF THE SCHRÖDINGER OPERATOR

By

A. JUHÁSZ

Department for Atomic Physics of the L. Eötvös University, Budapest

(Received May 9, 1983)

In [2] Joó proved: If $\Lambda \subset R^n$ is an arbitrary domain, $q(x) = C/|x-x_0|^{2+\varepsilon}$, $(x, x_0 \in \Lambda, n \geq 3)$, then for an arbitrary eigenfunction u of the Schrödinger operator $L = -\triangle + q_0$ the equation $u(x_0) = 0$ holds.

The aim of the present note is to generalize this result for more general class of potentials, namely we assume only that for the potential the relation $q(x) \geq C/|x-x_0|^{2+\varepsilon}$, $(C>0, \varepsilon>0, x \in \Lambda)$[1] holds. For the proof we need essentially new ideas comparing with [2].

Joó used essentially in his proof the spherical symmetry of q.

Our result may have interest in quantum mechanics [1].

We consider the eigenfunctions in weak sense: a function $u \in C(\Lambda)$ is said to be an eigenfunction of the operator $L = -\triangle + q_0$. if for every $\varphi \in C_0^\infty(\Lambda)$

$$\int_\Lambda u(L\varphi) = \lambda \int_\Lambda u(\varphi). \tag{1}$$

We prove the

THEOREM. *If u is a generalized eigenfunction of the operator L, further $q(x) \geq C/|x-x_0|^{2+\varepsilon}$ $(C>0, \varepsilon>0)$, then $u(x_0) = 0$.*

PROOF. The main idea of the proof is: we construct a function φ_0 such that supp $\varphi_0 \subset G(0, r)$,

$$\int_{|x| \leq r} u(\triangle \varphi_0 + \lambda \varphi_0) \leq c \cdot r^n \tag{2}$$

and

$$\int_{|x| \leq r} uq\varphi_0 \geq c \cdot r^{n-\varepsilon}, \quad (0<r<r_0, \quad \varepsilon>0) \tag{3}$$

if $u(x_0)>0$. But (2) and (3) contradicts to (1).

[1] Assume $q \in c(\Lambda \setminus \{X_0\})$.

Suppose $u(x_0) = c > 0$ and let r_0 be such that $|u(x)-c| < c/2$ if $|x-x_0| \leq \leq r_0$. Let $r \in (0, r_0)$. Choose a sequence $\varphi_n \in C_0^\infty(R)$ such that

$$\varphi_n(t) = \begin{cases} cr^2 & \text{if } t \leq r/2, \\ 0 & \text{if } t > r, \end{cases}$$

$$\triangle \varphi_n(x) = \varphi_n''(|x|) + \frac{n-1}{|x|} \varphi_n'(|x|) \to \triangle \varphi_0(|x|),$$

where

$$\varphi_0(s) = \int_s^\infty f(\sigma) d\sigma$$

and

$$r \cdot f(t) = \begin{cases} 0, & t \leq \frac{r}{2}, \\ t - \frac{r}{2}, & \frac{r}{2} \leq t \leq \frac{3}{4} r, \\ -t + r, & \frac{3}{4} r \leq t \leq r, \\ 0, & t > r, \end{cases}$$

further

$$|\triangle \varphi_n(|x|)| \leq \text{const} \quad \text{for } \frac{r}{2} \leq |x| \leq r,$$

$$\triangle \varphi_n(|x|) = 0 \quad \text{if } |x| \notin \left(\frac{r}{2}, r\right).$$

By Lebesgue's dominated convergence theorem

$$\int_A u(\triangle \varphi_n + \lambda \varphi_n) \to \int_A u(\triangle \varphi_0 + \lambda \varphi_0)$$

and

$$\int_A u q \varphi_n \to \int_A u \varphi_0 q_0.$$

Hence

$$B = \int_A u(\triangle \varphi_0 + \lambda \varphi_0) = \int_A u q \varphi_0 = J.$$

But

$$|B| \leq \text{const} \int_{\frac{r}{2} \leq |x| \leq r} u(x)dx \leq c \cdot r^n.$$

On the other hand

$$J = \int_{A} uq\varphi_0 \geq cr^2 \int_{|x| \leq r} q(x)dx \approx cr^{n-\varepsilon}.$$

The obtained contradiction proves the Theorem.

References

[1] G. Marx, *Kvantummechanika*, Műszaki Könyvkiadó, Budapest, 1957.
[2] I. Joó, On the summability of eigenfunction expansions, *Acta Math. Acad. Sci. Hung.*, (to appear).

О ЦЕНТРИРОВКЕ РЕШЁТОК

ЗОЛТАН МАЙОР

Кафедра Начертательной и Проективной Геометрии Унуверситета им. Л. Этвеша, Будапешт

(Поступило 4. 7. 1983)

Введение

В работе Б. Н. Делоне [1] было доказано следующее утверждение: Число различных центрировок решетки с данным индексом ограничено. (см. в частности в **2.**) Там же найдено и это число.

В работе С. С. Рышкова [3] были выведены при $n \leqq 7$ все попарно неэквивалентные допустимые центрировки параллелепипеда минимумов (см. в **2**) — для $n \leqq 6$ они ранее найдены Н. Хофрайтером [2], но другим способом — и для $n = 8$ нашла их Н. В. Захарова [4].

В настоящей работе использованы элементы теории Абелевых групп, даётся доказательство лемм, связанных с центрировками (см. в **2.**), и выведены числа неизоморфных центрировок с данным индексом и неизоморфные допустимые центрировки параллелепипеда минимумов при $n \leqq 8$ (см. в **3.**).

1. Центрировки

Пусть в n-мерном евклидовом пространстве $\mathbf{E}^n$ задан n-мерный репер:

$$\varepsilon = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\},$$

т. е. система n произвольных линейно независимых векторов, с общим началом.

Множество

$$\Gamma = \left\{\mathbf{x} \in \mathbf{E}^n \,|\, \mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i,\ x_i \text{ целое число}\right\}$$

называется n-решёткой, и репер ε, базисом Γ. Полуоткрытый параллелепипед

$$\Pi = \left\{\{\mathbf{x} \in \mathbf{E}^n \,|\, \mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i, 0 \leq x_i < 1\right\}$$

называется *основным параллелепипедом* решётки Γ, мы будем обозначать его объем через $\mathbf{V}(\Gamma)$.

Пусть $\Gamma^{ч}$ тоже n-решётка причём:

$$\Gamma \supset \Gamma^{ч}.$$

Множество

$$\Gamma^{ч} \cap \Pi = \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p\}$$

называется *центрировкой*, и векторы этого множества называются *определяющими векторами* центрировки, где сами векторы задаваются координатными строками относительно того базиса ε, на котором построен параллелепипед Π.

Известно [1], что если $\mathbf{V}(\Gamma^{ч})$ и $\mathbf{V}(\Gamma)$-объёмы основых параллелепипедов соответственно решётки $\Gamma^{ч}$ и её подрешётки Γ, то

$$p \cdot \mathbf{V}(\Gamma^{ч}) = \mathbf{V}(\Gamma), \tag{1}$$

где p — натуральное число (число определяющих векторов), называемое *индексом* центрировки.

Далее известно [3], что из p точек — принадлежащиах параллелепипеду Π — одна является одновременно и точкой решётки Γ,

$$\mathbf{a}_p = (0, 0, \ldots, 0),$$

а остальные $p-1$ точек — только точками решётки $\Gamma^{ч}$, т. е.

$$\mathbf{a}_i = \left(\frac{l_i^1}{m_i}, \frac{l_i^2}{m_i}, \ldots, \frac{l_i^n}{m_i}\right), \quad i = 1, 2, \ldots, p-1 \tag{2}$$

где $(l_i^1, l_i^2, \ldots, l_i^n, m_i) = 1$ и $0 \leq l_i^j \leq m_i$, $j = 1, 2, \ldots, n$.

Число m_i — называется *знаменателем определяющего вектора* $\mathbf{a}_i$. А наименьший общий знаменатель

$$m = [m_1, m_2, \ldots, m_{p-1}] \tag{3}$$

называется *знаменателем центрировки*.

Решётки Γ и $\Gamma^{ч}$ относительно сложения векторов образуют бесконечные Абелевы — группы. Рассмотрим факторгруппу

$$\Gamma^{ч}/\Gamma,$$

элементы которой являются смежными классами подгруппы Γ в группе $\Gamma^{ч}$. Эту факторгруппу будем называть *группой центрировки*, она будет играть большую роль в описании центрировок.

Обозначим смежный класс содержающий вектор $\mathbf{a}$, через $\{\mathbf{a}\}$, тогда известные следующие формулы:

$$\{\mathbf{a}\} = \{\mathbf{b}\} \Leftrightarrow \mathbf{a} - \mathbf{b} \in \Gamma,$$

$$\{\mathbf{a}\} + \{\mathbf{b}\} = \{\mathbf{a} + \mathbf{b}\},$$

$$n\{\mathbf{a}\} = \{n\mathbf{a}\},$$

где $\mathbf{a}, \mathbf{b} \in \Gamma^{ч}$ и n натуральное число. Отсюда видно, что факторгруппа $\Gamma/\Gamma^{ч}$ является конечной Абелевой-группой p — того порядка, где p-индекс центрировки. Заметим, что указанная выше формулировка соответствует формулировке, приведенной в [4] для дробной части вектора, поскольку там $\{\mathbf{a}\}$ обозначает дробную часть вектора $\mathbf{a}$, т. е. $\{\mathbf{a}\}$ определяющий вектор, для которого имеет место

$$\mathbf{a} - \{\mathbf{a}\} \in \Gamma.$$

Далее из (2) следует, что если знаменатель вектора $\mathbf{a}_i$ равен m_i, то порядок смежного класса $\{\mathbf{a}_i\}$ также равен m_i.

Посмотрим конечную Абелеву — группу $\Gamma^{ч}/\Gamma$. Всегда существует у неё базис, т. е. элементы $\{\mathbf{c}_1\}, \{\mathbf{c}_2\}, \ldots, \{\mathbf{c}_r\}$ из $\Gamma^{ч}/\Gamma$, для которых элемент группы $\{\mathbf{c}\}$ получается однозначно в виде $\{\mathbf{c}\} = n_1\{\mathbf{c}_1\} + n_2\{\mathbf{c}_2\} + \ldots + n_r\{\mathbf{c}_r\}$, где $n_1, n_2, \ldots, n_r$ целые, положительные числа, причём число n_i не больше порядка элемента $\{\mathbf{c}_i\}$ $(i = 1, 2, \ldots, r)$. Известно, что число элементов базиса вообще не однозначно. А всегда можно выбрать базис, состоящий из минимального числа элементов $\{\mathbf{c}_1\}, \{\mathbf{c}_2\}, \ldots, \{\mathbf{c}_q\}$ где $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_q$ суть определяющие векторы этих смежных классов. Множество определяющих векторов:

$$\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_q\}$$

будем называть минимальной базой центрировки. Очевидно, что центрировка будет польностью задана, если задать векторы минимальной базы. (Выбор минимальной базы, вообще говоря, неоднозначен.)

Известно далее, что обозначим порядки элементов базиса

$$\{\mathbf{c}_1\}, \{\mathbf{c}_2\}, \ldots, \{\mathbf{c}_q\}$$

через $r_1, r_2, \ldots, r_q$, то существует такой минимальной базис в $\Gamma^{ч}/\Gamma$, для которого

$$r_1 | r_2 | \ldots | r_q \tag{4}$$

и причём числа $r_1, r_2, \ldots, r_q$ и q однозначные. (Это следует из основной теоремы конечных Абелевых групп.) Множество таких определяющих векторов:

$$\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_q\}$$

будем называть канонической базой центрировки.

2. Различные и эквивалентные центрировки

Рассмотрим базис решётки $\Gamma^{ч}$, элементами которого являются векторы

$$\mathbf{b}_1 = (x_{11}, 0, \ldots, 0)$$

$$\mathbf{b}_2 = (x_{21}, x_{22}, \ldots, 0)$$

$$\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\mathbf{b}_n = (x_{n1}, x_{n2}, \ldots, x_{nn}),$$

где координаты векторов рациональные относительно базиса ε решётки Γ, и $0 \leq x_{ij} \leq 1$, $i = 1, 2, \ldots, n$. Из (1) очевидно, что обозначив числа $\frac{1}{x_i}$ через p_i $i = 1, 2, \ldots, p$, то получаем:

$$p = p_1 \cdot p_2 \cdot \ldots \cdot p_n.$$

Делоне доказал [1], что всякой центрировке с индексом p соответствует вполне определенное (принимается во внимание и порядок) разложение p на целые положительные множители, причём некоторые из них могут быть равны и 1. Так он получил, что число $\mathrm{I}_{n,p}$, всех возможных *различны* центрировок n-мерной решетки с индексом p равно

$$\Sigma p_2 \cdot p_2^3 \cdot p_4^3 \ldots p_n^{n-1},$$

где Σ распространнена на все возможные различные разложения числа p на n множителей $p_1, p_2, \ldots, p_n$, причём порядок множителей принимается во внимание.

Две центрировки соответственно n-мерных параллелепипедов Π_1 — в решётке Γ_1, и Π_2 — в решётке Γ_2 будем называть *эквивалентными*, если можно выбрать базисы в решётках Γ_1 и Γ_2, которые определяют параллелепипеды Π_1^* и Π_2^*, где если $\mathbf{e}$ есть один из тех векторов, на которых построен параллелепипед Π_i^* $(i = 1, 2)$, то $\mathbf{e}$ или $-\mathbf{e}$ — есть один из тех векторов на которых построен параллелепипед Π_i $(i = 1, 2)$, и относительно которых обе ентрировки задаются одним и тем же множеством определяющих векторов.

Замечание. Определение эквивалентных центрировок в [3] и в [4] следующее: Две центрировки соответственно n-мерных параллелепипедов Π_1 — в решётке Γ_1 и Π_2 — в решётке Γ_2 будем называть эквивалентными, если можно выбрать базисы в решётках Γ_1 и Γ_2, относительно которых обе центрировки задаются одним и тем же множеством определяющих векторов. В настоящей работе докажем, что это определение эквивалентно с определением неизоморфных центрировок (см. Теорема 2.).

Пусть Γ решётка, в которой существует минимальный базис ε, т. е. векторы $\mathbf{e}_i$ базиса, являются минимальными векторами решётки, причём

$$|\mathbf{e}_i| = 1, \quad i = 1, 2, \ldots, n.$$

Центрировка параллелепипеда Π, который построен на базис ε, называется [3] *допустимой*, если длины векторов $\Gamma^{ч}$ не меньше 1, т. е. в решётке $\Gamma^{ч}$ длина минимальных векторов остаётся той же, что и решётке Γ.

Числа неэквивалентных допустимых центрировок параллелепипеда минимумов — при $n \leqq 8$ — находятся в таблице I. (смотри и введение).

А теперь докажаем две леммы, которые помогут нам искать неэквивалентные центрировки.

Лемма 1. Векторы минимальной базы центрировки линейно независимые.

Доказательство. Пусть минимальная база $\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_q\}$ и предположим, что векторы линейно зависимые, т. е. существуют целые числа для которых имеем:

$$n_1\mathbf{c}_1 + n_2\mathbf{c}_2 + \ldots + n_q\mathbf{c}_q = \mathbf{0}$$

поскольку $\mathbf{c}_i \in \Gamma^{ч}$, $i = 1, 2, \ldots, p$, где можно предполагать что

$$(n_1, n_2, \ldots, n_q) = 1.$$

Обозначим порядки элементов $\{\mathbf{c}_i\}$, т. е. знаменатели векторов $\mathbf{c}_i$ через q_i, $i = 1, 2, \ldots, q$. Известно, что поскольку $\{\mathbf{c}_1\}, \{\mathbf{c}_2\}, \ldots, \{\mathbf{c}_q\}$ является минимальным базисом в факторгруппе $\Gamma^{ч}/\Gamma$, поэтому

$$(q_1, q_2, \ldots, q_n) > 1.$$

А по (5) имеем, что

$$n_1\{\mathbf{c}_1\} + n_2\{\mathbf{c}_2\} + \ldots + n_q\{\mathbf{c}_q\} = \{\mathbf{0}\},$$

т. е. $q_i | n_i$, $i = 1, 2, \ldots, q$, и таким образом получили

$$(q_1, q_2, \ldots, q_q) | (n_1, n_2, \ldots, n_q) = 1$$

очевидное противоречие. Лемма доказана.

Следствие. Число элементов минимальной базы не превосходит n.

Лемма 2. Знаменатель центрировки является наименьшим общим знаменателем векторов минимальной базы, т. е.:

$$m = [q_1, q_2, \ldots, q_q].$$

Доказательство. С одной стороны

$$[q_1, q_2, \ldots, q_q] | [m_1, m_2, \ldots, m_{p-1}] = m,$$

поскольку числа q_i находятся и среди m_i.

С другой стороны из равенства:

$$\{\mathbf{a}_i\} = n_1\{\mathbf{c}_1\} + n_2\{\mathbf{c}_2\} + \ldots + n_q\{\mathbf{c}_q\} = \{n_1\mathbf{c}_1 + n_2\mathbf{c}_2 + \ldots + n_q\mathbf{c}_q\}$$

следует, что

$$m_i | [q_1, q_2, \ldots, q_q] \quad i = 1, 2, \ldots, p-1,$$

и так получаем, что

$$m | [q_1, q_2, \ldots, q_q].$$

Лемма доказана.

Из Леммы 2. и из равенства $p = q_1 \cdot q_2 \cdot \ldots \cdot q_q$ непосредственно слеют Лемма 3. и Лемма 4., которые в [4] Лемма 5, и Лемма 6., но другим способом доказаны.

Лемма 3. Знаменатель m центрировки делит её индекс p, т. е.

$$m | p.$$

Лемма 4. Пусть знаменатель центрировки m — простое число, а q — число векторов её минимальной базы. Тогда для индекса центрировки имеет место:

$$p = m^q.$$

3. Изоморфные центрировки

Две центрировки называются *изоморфными*, если их группы являются изоморфными.

Теорема 1. *Число неизоморфных центрировок с данным индексом p, равно числа разложения p на целые положительные множители:*

$$p = p_1 \cdot p_2 \cdot \ldots \cdot p_t,$$

где p_i степени простих чисел, а порядок сомножителей не принимается во внимание, и заметим, что числа p_i не обязательно различные.

Доказательство. Поскольку индекс p центрировки равен порядку группы центрировки, число неизоморфных центрировок с данном индексом очевидно равно числа неизоморфных конечных Абелевых групп с данном порядком. И так по определению неизоморфных центрировок и по основной теоремы конечных Абелевых групп уже следует утверждение нашей теоремы.

Лемма 5. Пусть $\Gamma \subset \Gamma^{ч}$ n-решётки. Тогда существует базис, $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$ в решётке Γ и сушествуют ненулевые целые числа $r_1, r_2, \ldots, r_n$ такие, что:

1. векторы $r_1\mathbf{a}_1, r_2\mathbf{a}_2, \ldots, r_n\mathbf{a}_n$ образуют базис в решётке Γ,
2. $r_1 | r_2 | \ldots | r_n$
3. числа $r_1, r_2, \ldots, r_n$ суть однозначные.

Доказательство. Пусть $\mathbf{a}_1 \in \Gamma^{ч}$, $\mathbf{b}_1 \in \Gamma$, $r_1 \cdot \mathbf{a}_1 = \mathbf{b}_1$ и $\Gamma^{ч}_{n-1}$, Γ_{n-1} $n-1$ мерные подрешётки в $\Gamma^{ч}$ и в Γ, такие что:

$$\Gamma^{ч} = \{\mathbf{x} \in \mathbf{E}^n | \mathbf{x} = t \cdot \mathbf{a}_1 + \mathbf{a},\ t \in \mathbf{Z},\ \mathbf{a} \in \Gamma^{ч}_{n-1}\}$$

$$\Gamma = \{\mathbf{x} \in \mathbf{E}^n | \mathbf{x} = t \cdot \mathbf{b}_1 + \mathbf{b},\ t \in \mathbf{Z},\ \mathbf{b} \in \Gamma_{n-1}\}.$$

Пусть далее r_1 минимальное такое число.

Тоже самая получим подрешётки $\Gamma_{n-2} \subset \Gamma^{ч}_{n-2}$ и $\Gamma_{n-2} \subset \Gamma_{n-1}$, векторы $\mathbf{a}_2 \in \Gamma^{ч}_{n-1}$, $\mathbf{b}_2 \in \Gamma_{n-1}$ и минимальное число $r_2 (r_2\mathbf{a}_2 = \mathbf{b}_2)$, где нетрудно видеть что

$$r_1 | r_2.$$

Так получим базис $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$ в $\Gamma^{ч}$ и базис $\{\mathbf{b}_1 = r_1\mathbf{a}_1, \mathbf{b}_2 = r_2\mathbf{a}_2, \ldots, \mathbf{b}_n = r_n\mathbf{a}_n\}$ в Γ где $r_1 | r_2 | \ldots | r_n$. Среди чисел $r_1, r_2, \ldots, r_n$ могут быть некоторые равнымы единице, т. е. например $r_1 = r_2 = \ldots = r_{n-q} = 1$. Но тогда смежные классы $\{\mathbf{a}_{n-q+1}\}, \{\mathbf{a}_{n-q+2}\}, \ldots, \{\mathbf{a}_n\}$ образуют канонический базис в факторгруппе $\Gamma^{ч}/\Gamma$, и поэтому числа $r_1, r_2, \ldots, r_n$ однозначные.

Таблица I.

Размерность n	Число неэквивалентных допустимых центрировок параллелепипеда минимумов	Число неизоморфных допустимых центрировок параллелепипеда минимумов
2	1	1
3	1	1
4	2	2
5	3	2
6	6	4
7	14	6
8	42*	11

* Эти данные мы получили по сообщению Н. В. ЗАХАРОВА в *Реферативном журнале* (1982/1.; 1А828ДЕП), где находится дополнение результатов [4] (т. е. ещё три допустимых центрировки 8-мерных решёток).

Таблица II.

Индекс p	Минимальная база (m-кратная)	Знаменатель m	Размерность $n \geqq$	Номера изоморфных центрировок в таблице I. в [4]
1	(00000000)	–	2	1
2	(11110000)	2	4	2–4, 6, 10
3	(11111100)	3	6	19–21
$4 = 2 \cdot 2$	(11110000) (00111100)	2	6	5,7–9, 11–14
4	(11111120)	4	7	23–30
5	(11111122)	5	8	31–33
6	(11122223)	6	8	34–39
$8 = 2 \cdot 2 \cdot 2$	(11110000) (11001100) (01011010)	2	7	15–17
$8 = 4 \cdot 2$	(11111120) (00022222)	4	8	40–42*
$9 = 3 \cdot 3$	(11111100) (00221111)	3	8	22
$16 = 2 \cdot 2 \cdot 2 \cdot 2$	(11110000) (11001100) (10101010) (01101001)	2	8	18

Таким образом лемма доказана.

Замечание. Поскольку решётка $\Gamma^{\text{ч}}$ свободный модуль над кольцом целых чисел и Γ его подмодуль, так наша лемма переизложение основной теоремы свободных модулей (см. например в книге Ленга [5], в главе XV., на стр. 440.).

Из Леммы 5. непосредственно следует следующая важная теорема:

Теорема 2. *Две центрировки соответственно n-мерных параллелепипедов Π_1 – в решетке Γ_1, и Π_2 – в решётке Γ_2 изоморфные тогда и только тогда, если можно выбрать базисы в решетках Γ_1 и Γ_2, относительно которых обе центрировки задаются одним и тем же множеством определяющих векторов.*

Замечание. По утверждению Теоремы 2. можно говорить о центрировках решёток, поскольку изоморфные центрировки точно те, которые получаются относительно различных базисов одной решётки.

Теорема 3. *Число неизомофных допустимых центрировок параллелепипеда минимумов при $n = 2, 3, 4, 5, 6, 7, 8$ равно соотвественно* 1, 1, 2, 2, 4, 6, 11.

Доказательство. Утверждение теоремы видно из таблици II., в которой даётся неизоморфные допустимые центрировки параллелепипеда минимумов при $n \leqq 8$. Таблицу II. составили по определению неизоморфных центрировок и по таблице I. из [4].

Литература

[1] Делоне Б. Н., Фаддеев Д. К.: Теория иррациональностей третей степены, *Труды МИАН СССР*, **11** (1940), 63–69.
[2] Hoffreiter N.: Zur Geometrie der Zahlen, *Monatsh. Math. Phys.*, **40** (1933), 181–192.
[3] Рыщков С. С.: К проблеме отыскания совершенных квадратичных форм от многих переменных, *Труды МИАН СССР*, (1976), **142** 215–239.
[4] Захарова Н. В.: Центрировки 8-мерных решёток, сохраняющие репер последовательных минимчмов, *Труды МИАН СССР*, **152** (1980), 97–123.
[5] Ленг С.: *Алгебра*, Издательство «Мир» Москва 1968, (Addison – Wesley Publishing Company; Reading, Mass., 1965.)

AUSFÜLLUNG UND ÜBERDECKUNG DER EBENE DURCH KREISE

Von

A. BEZDEK

Mathematisches Institut der Ungarischen Akademie der Wissenschaften, Budapest

(Eingegangen am 28. Juli 1983)

Betrachten wir eine unendliche Folge von Kreisen $k_1, k_2, \ldots$, deren Halbmesser von oben beschränkt sind, und deren Gesamtflächeninhalt unendlich ist. L. FEJES TÓTH [1] hat die Packungswirtschaftlichkeit w einer Kreisfolge mit $\overline{\lim} \frac{\sum_{i=1}^{n} k_i}{T_n}$ – wo T_n das kleinstmögliche einem vorgegebenen konvexen Gebiet änliche Gebiet ist, in dem die ersten n Kreise ohne gegen seitige Überdeckung eingelagert werden können – eingeführt. Der Wesen des Unterschiedes zwischen der Dichte d der Dichtesten Packung der Kreise [2] die zu der Folge von Kreisen gehören, und w besteht darin, daß w von der Reihenfolge der Kreise abhängt, aber d nicht. Es läßt sich zeigen, daß besteht die Kreisfolge aus kongruenten Kreisen, so gilt $w = d$. In ähnlicher Weise können wir die Deckungswirtschaftlichkeit der Kreisfolge definieren.

L. FEJES TÓTH hat gezeigt [3], daß

$$\frac{1}{w} \geqq 1 + \frac{\sqrt{12} - \pi}{\pi} \frac{M_{1/3}(k_1, \ldots, k_n)}{M_1(k_1, \ldots, k_n)}.$$

Hier bedeutet $M_\alpha(k_1, \ldots, k_n)$ das α-te Potenzmittel der Flächeninhalte der ersten n Kreise. Hieraus folgt, ist

$$\lim_{n \to \infty} \frac{M_{1/3}(k_1, \ldots, k_n)}{M_1(k_1, \ldots : k_n)} \neq 0$$

dann $w \neq 1$. Das ist der Fall, wenn z.B. die Halbmesser der Kreise der Folge $1^{-\alpha}, 2^{-\alpha}, \ldots, n^{-\alpha}, \ldots$ sind wo $0 < \alpha < 1$ ist. Trotzdem – wie wir es jetzt zeigen werden – läßt sich die Ebene mit der Dichte 1 ausfüllen.

SATZ 1. *Wenn die aus den Halbmessern der Kreisfolge $k_1, \ldots, k_n, \ldots$ bestehende Zahlenfolge $r_1, r_2, \ldots$ den Voraussetzungen $\sum_1^\infty r_i^2 = \infty$ und $\lim_{i\to\infty} r_i = 0$ genügt, so läßt sich die Ebene mit den Kreisen $k_1, \ldots, k_n, \ldots$ mit der Dichte 1 ausfüllen.*

BEWEIS. Im folgenden werden wir die Kreise mit ihren Halbmessern bezeichnen. Wir beweisen erst folgenden später zu verwendenden Hilfsatz.

HILFSATZ 1. Wir geben ein Verfahren, wie wir ein beliebig gewähltes Quadrat A_1B_1CD mit einer endlichen Anzahl der Kreise der Kreisfolge mit der Dichte größer als $\frac{1}{8}$ ausfüllen können.

Zuerst numerieren wir die Kreise so, daß die Halbmesser eine abnehmende Zahlenfolge bilden. Es gibt ein natürliches i_1 so, daß $r_{i_1} < \frac{a}{8}$ ist. Die Kreise werden nach folgender Vorschrift in den Zonen eingebettet (Fig. 1): r_{i_1} wird der erste Kreis der ersten Zone, A_1B_1 wird die obere Randstrecke der ersten Zone sein.

Wenn im allgemeinen der erste Kreis r_{i_n} und die obere Randstrecke A_nB_n der n-ten Zone vorgegeben ist, wählen wir und füllen wir die n-te Zone folgenderweise aus: Der Kreis r_{i_n} soll die Gerade A_1D und A_nB_n berühren. Es schneide die Stützgerade von r_{i_n}, welche mit A_nB_n parallel und von ihr verschieden ist, die Gerade A_1D und B_1C in den Punkte A_{n+1} bzw. B_{n+1}. Die Strecke $A_{n+1}\,B_{n+1}$ wird die obere Randstrecke der $n+1$-ten Zone sein. Das Rechteck $A_nB_nB_{n+1}A_{n+1}$ nennen wir die n-te Zone und wir bezeichnen ihren Flächeninhalt mit S_n. Die nach dem Kreis r_{i_n} folgenden Kreise der Kreisfolge legen wir sukzessiv so, daß jeder Kreis den vorhergehenden berühre, und die Mittelpunkte auf der mit A_nB_n parallelen Symmetrieachse der n-ten Zone liegen. Der erste Kreis $r_{i_{n+1}}$, der sich so in der n-ten Zone nicht einladen läßt, wird der erste Kreis der $n+1$-ten Zone sein.

Der Gesamtflächeninhalt der Kreise, die in der n-ten Zone eingelagert werden, ist größer als die Hälfte vom Flächeninhalt der $n+1$-ten Zone, da

$$\sum_{j=i_n}^{i_{n+1}-1} r_j^2\pi > r_{i_{n+1}} \sum_{j=i_n}^{i_{n+1}-1} 2r_j > r_{i_{n+1}}\left(a - \frac{a}{4}\right) > \frac{S_{n+1}}{4}$$

ist. Da $\sum_{i=i_1}^\infty r_i^2 = \infty$, gilt $\sum_{n=1}^\infty S_n = \infty$. Es sei N die kleinstmögliche ganze Zahl, für die $\sum_{n=1}^N S_n \geq \frac{3}{4}a^2$ ist. Da die Breite der Zonen abnimmt und $2r_{i_1} < \frac{1}{4}a$ ist, sind die ersten N Zonen im Quadrat enthalten. Das bedeutet, daß der Gesamtflächeninhalt der Kreise, die in den ersten $N-1$ Zonen eingelagert werden, das Achtel von Flächeninhalt des Quadrates überschreitet. Wir wenden uns nun dem konstruktiven Beweis des Satzes zu.

Betrachten wir eine um einen festen Ursprungspunkt O geschlagene Kreisfolge, deren Halbmesser $(2n-1)\,r_1$; $n = 1, 2, \ldots$ sind (Fig. 2). Den Kreis r_1 selbst nennen wir den O-ten Kreisring, und den von den Kreisen $(2n-1)r_1$, $(2n+1)r_1$ begrenzten Kreisring nennen wir den n-ten Kreisring. Die Kreise werden nach folgendem Verfahren eingelagert werden.

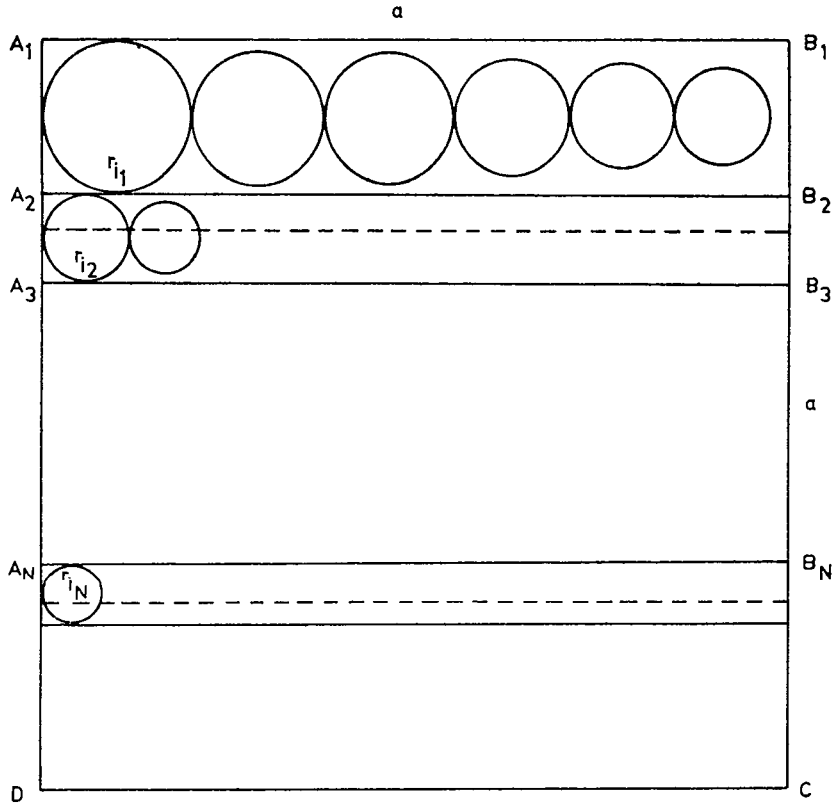

Fig. 1.

Legen wir den Kreis r_1 in den O-ten Kreisring.

Nach der Ausfüllung der ersten i Kreisringe füllen wir den $i+1$-ten Kreisring mit einer endlichen Anzahl der Kreisen aus so, daß die Dichte im $i+1$-ten Kreisring größer als $\left(1-\frac{1}{i+1}\right)$ ist: Wir legen als erstes den Kreis r_{i+1} in den Kreisring, wenn er früher nicht gebraucht wurde.

Wählen wir jetzt ein Quadratnetz. Wir betrachten jetzt den Gesamtflächeninhalt aller solcher Quadrate, die ganz im $i+1$-ten Kreisring liegen und die eingelagerten Kreise nicht schneiden. Wir wählen die Abstände des Geraden des Quadratnetzes so klein, daß dieser Gesamtflächeninhalt größer

als der halbe Flächeninhalt der Restfläche des Kreisringes nach dem Herausnehmen der Kreise ist. (Da nur eine endliche Anzahl von Kreise im $i+1$-ten Kreisring liegt, ist das möglich.) Unter Berücksichtigung des Hilfsatzes 1 ist es möglich, diese Quadrate nacheinander mit der Dichte $\frac{1}{8}$ auszufüllen.

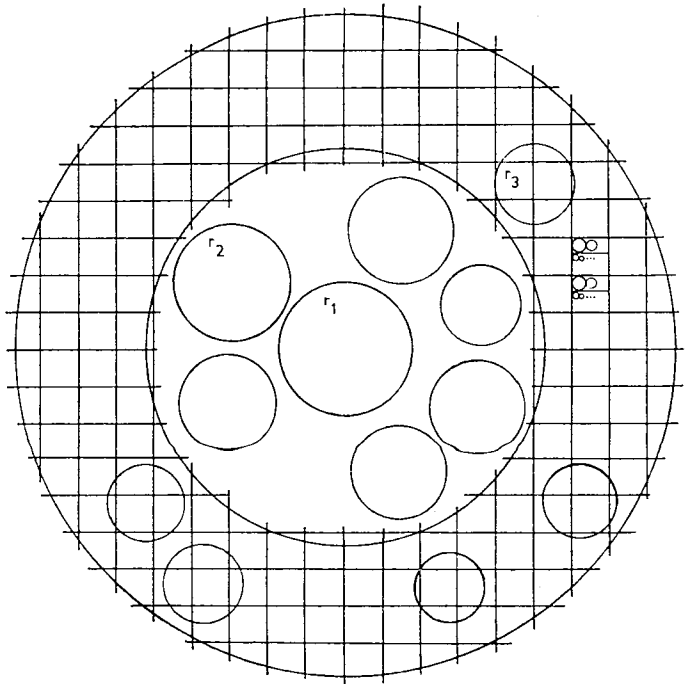

Fig. 2.

Dies bedeutet, daß das von den Kreisen freigelassene Teilgebiet auf das $\frac{7}{8}$-fache zurückgeht. Hieraus folgt unmittelbar, daß nach einer endlichen Anzahl von Wiederholung unseres Verfahren die Dichte im Kreisring größer als $\left(1-\frac{1}{i+1}\right)$ ist.

Jetzt werden wir einen analogen Satz für die Überdeckung beweisen.

SATZ 2. *Wenn die aus den Halbmessern der Kreisfolge $k_1, \ldots$ bestehende Zahlenfolge $r_1, r_2, \ldots$ den Voraussetzungen $\sum_{i=1}^{\infty} r_i^2 = \infty$ und $\lim_{i\to\infty} r_i = 0$ genügt so läßt sich die Ebene mit den Kreisen $k_1 \ldots$, mit der Dichte 1 überdecken.*

Der Beweis des Satzes 2 beruht ähnlich wie der von Satz 1 auf einem Hilfsatz.

HILFSATZ 2. Wir geben ein Verfahren, wie wir ein beliebig gewähltes Quadrat durch eine endliche Anzahl Kreise der Kreisfolge mit einer Dichte kleiner als 8 überdecken können.

Es sei die Kreisfolge wieder so numeriert, daß die Halbmesser eine abnehmende Zahlenfolge bilden. Wir bezeichnen die Seite des zu überdeckenden Quadrates mit b. Füllen wir das Quadrat, dessen Seite $a = 2\sqrt{2}b$ ist, nach dem Verfahren des Hilfsatzes 1 aus. Es seien die Breiten der Zonen $2r_{i_1}, \ldots, 2r_{i_{N-1}}$. Ein Rechteck, das $\frac{2}{\sqrt{2}} r_{i_{n+1}}$ breit und b lang ist, läßt sich durch die in der n-ten Zone liegenden Kreise überdecken. Dies folgt unmittelbar daraus, daß jeder Kreis r_i ein in ihn einbeschreibenes Quadrat mit der Seite $\frac{2}{\sqrt{2}} r_i$ überdeckt und jeder Kreis größer als $r_{i_{n+1}}$ ist. Da N so gewählt war, daß $\sum_{n=2}^{N} 2r_{i_n} > \frac{a}{2} = \sqrt{2}b$ ist, ist es möglich das Quadrat mit der Seite b durch die in den ersten $N-1$-ten Zonen liegenden Kreise zu überdecken. Natürlich ist der Gesamtflächeninhalt dieser Kreise $< a^2 = 8b^2$, und die Dichte ist so < 8. Mit Hilfe der Konstruktion des Satzes 1 ist der Beweis des Satzes 2 sehr einfach. Natürlich ist es genug zu zeigen, daß der n-te Kreisring sich durch Kreise endlicher Anzahl und mit einer Dichte kleiner als $\left(1+\frac{1}{n}\right)$ überdecken läßt.

Erst füllen wir den n-ten Kreisring mit einer Dichte größer als $\left(1-\frac{1}{15n}\right)$ aus.

Wählen wir einen Quadratnetz so, daß der doppelte Flächeninhalt des ausserhalb der Kreise liegenden Teilgebietes von n-ten Kreisring größer ist als der Gesamtflächeninhalt der Quadrate des Quadratnetzes, die mit dem ausserhalb der Kreise liegenden Teilgebiet gemeinsame Punkte haben.

Wir überdecken die Quadrate, mit Hilfe des Hilfsatzes 2, nacheinander. Die Dichte im n-ten Kreisring ist $< \left(1-\frac{1}{15n}\right)1+\frac{2}{15n}\cdot 8 = 1+\frac{1}{n}$.

Damit ist unser Beweis beendet.

Literatur

[1] L. FEJES TÓTH: *Lagerungen in der Ebene auf der Kugel und im Raum*, Springer–Verlag, New York, 1972.
[2] L. FEJES TÓTH: *Reguläre Figuren*, Akadémiai Kiadó, 1965.
[3] L. FEJES TÓTH: Packing and coverings with convex discs., *Studia Sci. Math. Hung.*, **15** (1980), 93–100.

ЦЕНТРИРОВКИ РЕШЁТОК СО ЗНАМЕНАТЕЛЕМ 2, ПРИ $n \leqq 10$

ЗОЛТАН МАЙОР

Кафедра Начертательной и Проективной Геометрии Университета им. Л. Этвеша, Будапешт

(Поступило 9. 9. 1983)

Рассмотрим в n-мерном евкливодом пространстве $\mathbf{E}^n$ n-решётку $\Gamma^ч$, которая содержить кубическую n-решётки Γ_ε, т. е.

$$\Gamma_\varepsilon \subset \Gamma^ч,$$

причём пусть

$$\min \Gamma_\varepsilon = \min \Gamma^ч = 1.$$

Говорят, что решётка $\Gamma^ч$ получается допустимой центрировкой решётки Γ_ε.

Пусть $\varepsilon = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ n-мерный репер, т. е. система n произвольных линейно независимых векторов с общим началом, на котором построен основный куб Π_ε т. е. один из основных параллелепипедов — кубической решётки Γ_ε. Система ε является базисом решётки Γ_ε. Известно [3], что все координаты векторов решётки $\Gamma^ч$ относительно базиса ε, суть рациональные числа. Если наименьший обший знаменатель этих рациональных чисел не превосходит 2 — и решётки Γ_ε и $\Gamma^ч$ различные —, то мы говорим о центрировке со знаменателем 2.

В настоящей заметке выведены при $n \leqq 10$ все неэквивалентные и неизоморфные допустимые центрировки основного куба кубичной решётки со знаменателем 2.

В работе Н. В. Захаровой [4] доказано, что если допустимая центрировка параллелепипеда минимумов со знаменателем 2 существует, то она имеет место по крайней мере для кубической решётки. В работе С. С. Рышкова [3] доказано, что каждая допустимая центрировка произвольной решётки, т. е. центрировка сохранящая репер последовательных минимумов, эквивалентна допустимой центрировке параллелепипеда минимумов. Отсюда следует, что в настоящей заметке выведены при $n \leqq 10$ и все неизоморфные и неэквивалентные допустимые центрировки n-решёток, сохраняющие репер последовательных минимумов со знаменателем 2 (в частности см. в **3**).

В работах С. С. Рышкова [3] и Н. В. Захаровой [4] уже были выведены при $n \leqq 7$ и при $n = 8$ все неэквивалентные допустимые центрировки сохраняющие репер последовательных минимумов.

В настоящей заметке мы докажем несколько новых лемм и наши результаты получаем — при $n \leqq 1$— — с помощчю персонального компютера.

1. Определения и обозначения

Пусть в n-мерном евклидовом пространстве $\mathbf{E}^n$ задан система n произвольных линейно независимых векторов, с общим, началом:

$$\varepsilon = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}.$$

Посмотрим два множества:

$$\Gamma_\varepsilon = \{\mathbf{x} \in \mathbf{E}^n | \mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i,\ x_i \text{ целое число}\},$$

$$\Pi_\varepsilon = \{\mathbf{x} \in \mathbf{E}^n | \mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i,\ 0 \leqq x_i < 1\}.$$

Γ_ε называется *n-решёткой*, репер ε *базисом* Γ_ε, и полуоткрытый параллелепипед Π_ε, построенный на базисе ε, называется *основным параллелепипедом* решётки Γ_ε.

Множество

$$\mathbf{S}_\varepsilon = \Gamma^{\text{ч}} \cap \Pi_\varepsilon = \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p\}$$

называется *центрировкой*, если $\Gamma^{\text{ч}}$ также n-решётка, и причём

$$\Gamma_\varepsilon \subset \Gamma^{\text{ч}}.$$

Векторы множества $\mathbf{S}_\varepsilon$ называются *определяющими векторами* центрировки, где сами векторы $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_p$ задаются координатными строками относительно базиса ε:

$$\mathbf{a}_i = \left(\frac{l_i^1}{m_i}, \frac{l_i^2}{m_i}, \ldots, \frac{l_i^n}{m_i}\right), \quad i = 1, 2, \ldots, p-1,$$

где

$$(l_i^1, l_i^2, \ldots, l_i^n, m_i) = 1 \text{ и } 0 \leqq l_i^j < m_i, \quad j = 1, 2, \ldots, m,$$

далее,

$$\mathbf{a}_p = (0, 0, \ldots, 0).$$

Число

$$m = [m_1, m_2, \ldots, m_{p-1}]$$

называется *знаменателем* центрировки, где m_i знаменатель определяющего вектора $\mathbf{a}_i$ и m является наименьшим общим знаменателем знаменателей m_i.

Число p называется *индексом* центрировки, p натуральное число, число определяющих векторов. Целочисленная матрица:

$$\mathbf{S} = [l_i^j] = \begin{bmatrix} l_1^1 & l_1^2 & \cdots & l_1^n \\ l_2^1 & l_2^2 & \cdots & l_2^n \\ \cdot & \cdot & \cdots & \cdot \\ l_{p-1}^1 & l_{p-1}^2 & \cdots & l_{p-1}^n \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

называется *матрицой* центрировки.

Можно рассматривать факторгруппу:

$$\Gamma^{ч}/\Gamma_{\varepsilon}$$

поскольку решётки $\Gamma^{ч}$ и Γ_{ε} относительно сложению векторов образуют бесконечные Абелевы-группы. Эту факторгруппу будем называть *группой* центрировки. Эта факторгруппа является конечной Абелевой-группой p-того порядка, где p индекс центрировки. Обозначим смежный класс содержающий вектор $\mathbf{a}$ решётки $\Gamma^{ч}$, через $\{\mathbf{a}\}$. Известно, что всегда существует базис в группе $\Gamma^{ч}/\Gamma_{\varepsilon}$, состоящий из минимального числа злементов:

$$\{\mathbf{c}_1\}, \{\mathbf{c}_2\}, \ldots, \{\mathbf{c}_q\},$$

где $\mathbf{c}_1$, $\mathbf{c}_2$,..., $\mathbf{c}_q$ суть определяющие векторы этих смежных классов. Множество определяющих векторов:

$$\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_q\}$$

будем называть минимальной базой центрировки.

Очевидно, что центрировка будет польностью задана, если задать векторы минимальной базы.

Две центрировки соотвественно n-мерных параллелепипедов Π_{ε_1} — в решётке Γ_{ε_1} и Π_{ε_2} — в решётке Γ_{ε_2} будем называть *эквивалентнымы*, если сужествуют такие базисы ε_1^* и ε_2^* в решётках, для которых имеет место:

1. $\varepsilon_1^* \subset \pm\varepsilon_1$ $\varepsilon_2^* \subset \pm\varepsilon_2$,
2. $\mathbf{S}_{\varepsilon_1^*} = \mathbf{S}_{\varepsilon_2^*}$,

т. е. если $\mathbf{e} \in \varepsilon_i^*$, то $\mathbf{e} \in \varepsilon_i$, или $-\mathbf{e} \in \varepsilon_i$, $i = 1, 2$, и относительно которых обе центрировки задаются одним и тем же множеством определяющих векторов.

Две центрировки называются *изоморфными*, если их группы являются изоморфными.

Пусть далее Γ такая решётка, в которой существует *минимальный* базис ε, т. е. векторы $\mathbf{e}_i$ ($j = 1, 2, \ldots, n$) базиса ε являются минимальными векторами решётки, причём:

$$|\mathbf{e}_i| = 1 \quad i = 1, 2, \ldots, n.$$

Центрировка $\mathbf{S}_\varepsilon$ параллелепипеда Π_ε называется *допустимой* если

$$\min \Gamma^{ч} = \min \Gamma_\varepsilon = 1$$

т. е. длины векторов решётки $\Gamma^{ч}$ не меньше 1. Допустимая центрировка $\mathbf{S}_\varepsilon$ называется *свободной*, если параметры решётки Γ_ε можно выбрать такими, что среди векторов решётки $\Gamma^{ч}$ не будет других минимальных векторов, кроме минимальных векторов решётки Γ_ε и сам решётка Γ_ε будет обладать лишь одним параллелепипедом минимумов. В противном случае допустимая центрировка называется *несвободной*.

2. Несколько лемм

Всюду нуже мы будем рассматривать только кубическую решётки, т. е.

$$\mathbf{e}_i \cdot \mathbf{e}_j = 0 \quad \text{при} \quad i \neq j,$$

$$\mathbf{e}_i \cdot \mathbf{e}_j = 1 \quad \text{при} \quad i, j = 1, 2, \ldots, n,$$

и пусть знаменатель центрировки $m = 2$.

Лемма 1. Пусть p-индекс центрировки, и q-число векторов минимальной базы. Тогда имеет место равенство

$$p = 2^q.$$

Лемма 2. Центрировка $\mathbf{S}_\varepsilon$ допустимая тогда и только тогда, если сумма элементов в первых строках матрицы $\mathbf{S}$ не меньше 4:

$$\sum_{j=1}^{n} l_i^j \geqq 4, \quad i = 1, 2, \ldots, p-1. \tag{1}$$

Лемма 3. Если в (1) имеет место знак неравенства при $i = 1, 2, \ldots, p-1$, то центрировка несвободная.

Эти леммы следуют из Леммах 6. и 9. в [4].

Лемма 4. Две центрировки являются эквивалентными, если их матрицы можно перевести друг в друга перестановкой их строк или столбцов между собой.

Доказательство. Поскольку $m=2$, утверждение леммы непосредственно следует из определения эквивалентности.

Лемма 5. Если число элементов минимальной базы центрировки $q = 1$ или $q = 2$, то две центрировки эквивалентны тогда и только тогда когда множество чисел

$$\left\{ \sum_{i=1}^{n} l_i^j \mid \quad i = 1, 2, \ldots, p-1 \right\}$$

один и тот же в матрицах центрировок.

Доказательство. Если $q = 1$, то лемма тривиальна. Если $q = 2$, то лемма следует из того, что координаты векторов минимальной базы польностью определяют число ненулевых координат остальных определяющих векторов.

Лемма 6. Две центрировки являются изоморфными тогда и только тогда, когда их индексы равны.

Доказательство. Эта лемма следует из Леммы 1. и из определения изоморфных центрировок.

Лемма 7. Обозначим через $\mathbf{S}_{n,p}$ n-мерную допустимую центрировку, где $n > 1$, $p > 2$ и

$$p = 2^q$$

(знаменатель $m = 2$, число векторов минимальной базы q). Тогда существует центрировка $\mathbf{S}_{n^*,p^*}$, где

$$n^* \leqq n-1,\ p^* = 2^{q-1}$$

$$\mathbf{S}_{n^*,p^*} \subset \mathbf{S}_{n,p}.$$

Доказательство. Рассмотрим матрицу $\mathbf{S}$ центрировки $\mathbf{S}_{n,p}$. Элементы её равны 0 и 1, а в последной строке стоят 0. Докажем, что во всех строках этой матрици число элементов 1, так и число элементов 0, равно $p/2$. Отметим, поскольку это будет важно в дальнейщем, что строки матрици $\mathbf{S}$ образуют конечную Абелеву-группу $p = 2^q$-того порядка, относительно сложению мод 2 строков. Докажем утверждение индукцией по q. При $q = 1$ утверждение тривиально. Предположим утверждение справдливым при некотором натуральном числе $q-1$, и покажем что имеет место и при q. Действительно, среди строк матрицы $\mathbf{S}$ находятся q, которые образуют минимальную базу выше указанной Абелевой-группы, поэтому они однозначно определяют остальные строки матрици. Обозначим через $\mathbf{S}^*$ матрицу, которую мы получим из строк $\mathbf{S}$, которые определяют первые $q-1$ векторы минимальной базы. Матрицу $\mathbf{S}$ мы можем получить из строк матрицы $\mathbf{S}^*$ со сложением мод 2 строк $\mathbf{S}^*$ и q-того вектора минимальной базы. Таким образом в каждой столбце матрицы $\mathbf{S}$ и число единиц будет равно числу нулей, именно $2(p/4) = p/2$. Утверждение доказано.

Посмотрим далее например первый столбец матрици $\mathbf{S}$, и посмотрим матрицу $\mathbf{S}'$, который получается из тех строк матрици $\mathbf{S}$ в которых в первым столбце стоит 0. Отсюда уже легко видеть, что матрица $\mathbf{S}'$ является матрицей центрировки $\mathbf{S}_{n^*,p^*}$, в которой определяющие строки являются $n^* \leqq n-1$ мерными, и число их равно $p^* = p/2 = 2^{q-1}$ и они образуют бесконечную Абелеву-группу p^*-того порядка.

Таким образом лемма доказана.

Замечание. Аналогичную лемму можно доказать при $m > 2$, если знаменатель центрировки m, простое число, только в этом случае в

лемме выпольняется равенство $p^* = m^{q-1}$, и доказательство тоже аналогично.

Лемма 8. Пусть

$$q = \max\{q \mid \mathbf{S}_{n,2^q} \text{ допустимая центрировка}\}.$$

Тогда имеет место:

$$q_{n-1} \leqq q_n \leqq q_{n-1} + 1.$$

Доказательство. Левая часть очевидно. Правая часть непосредственно следует из Леммы 7.

3. Доказательство теорем

В этом параграфе будем доказать следующую теорему:

Теорема 1. *Числа допустимых нееквивалентных центрировок со знаменателем 2 основного куба кубической решетки при $n \leqq 10$ находятся в 2-ом строке Таблици I., (считая и тривиальную) из них число свободных и число несвободных находятся в 3-ем и 4-ом строке Таблици I., и числа допустимых неизоморфных центрировок со знаменателем 2 основного куба кубической решетки находятся в 5-ом строке Таблици I. Самы нееквивалентные и неизоморфные центрировки находятся в таблицах* II – IV. *и в таблице* V.

Замечание. Н. В. Захарова доказала следующую лемму (это следует из Леммы 9 в [4]):

Лемма. Если допустимая центрировка параллелепипеда минимумов со знаменателем 2 существует, то она имеет место по крайней мере для кубической решётки.

С. С. Рышков в [3] доказал следующую теорему (в [3] теорема 10.):

Теорема. *Каждая допустимая центрировка репера последовательных минимумов эквивалентна допустимой центрировке параллелепипеда минимумов, т. е. токого репера последовательных линейно независимых минимумов, в котором все векторы равны по длине.*

Таким образом из нашей Теоремы 1. непосредственно следует и следующая теорема:

Теорема 2. *На множестве n-решёток, при $n \leqq 10$, числа допустимых и неизоморфных, далее числа свободных и несвободных центрировок репера последовательных минимумов со знаменатели 2, и сами эти центрировки находятся в таблицах* I – V. *(считая и тривиальную).*

Доказательство теоремы 1. Лемма 2 даёт необходимое и достаточное условие для того, чтобы центрировка суть допустимой. С помощью Леммы 3 можно легко узнать, что центрировка свобода или несвобода.

По Лемме 1, индекс центрировки равен $p = 2^q$. С помощью Лемм 4–5 и 7–8 можно выбрать неэквивалентные центрировки из допустиных центрировок следующим образом:

3.1. $p = 2, 4$. Центрировки с индексоами $p = 2$ и $p = 4$ мы сразу получим по Лемме 2 и по Лемме 5. Так мы получим номера 2–13. в таблице II. номера 19–25. в таблице III. и номера 38–46. в таблице IV. Очевидно, что больше таких центрировок нет.

3.2. $p = 8$. Поскольку центрировка с индексом $p = 4$ при $n<6$ не существует поэтому по Леммам 1. и 8. при $n \leqq 6$ центрировки и индексом $p = 8$ не существуют. Для того чтобы получить существенно 7-мерные центрировки с индексом $p = 8$, по Лемме 7 можно исходить из минимальной базы центрировки под номером 7 в таблице II. Можно добавлять к этой базе третий вектор

$$\left(\frac{1}{2}, 0, \frac{1}{2}, 0, \frac{1}{2}, 0, \frac{1}{2}\right)$$

чтобы получить неэквивалентную и допустимую центрировку. Так мы получим центрировку под номером 14. в таблице II. Для того чтобы получить существенно 8-мерные центрировки с индексом $p = 8$, опять по Лемме 7 можно исходить из минимальнх баз, содержащих два вектора под номерами 7. 8. 9 в таблице II. Из этих мы получим номера 15. 16. и 17 в таблице II. Далее, для того чтобы получить существенно 9-мерных и существенно 10-мерные центрировки можно исходить из центрировок под номерами 7–13. в таблице II. и 20–25 в таблице III. Так мы получим номера 26–33 в таблице III. и номера 47–67. в таблице IV.

3.3. $p = 16$. Поскольку центрировки с индексом $p = 8$ при $n<7$ не существует, поэтому по Леммам 1 и 8 при $n \leqq 7$ центрировки с индексом $p = 16$ не существуют. По Лемме 7. существенно 8, 9 и 10-мерные неэквивалентные центрировки с индексом $p = 16$, мы получим из центрировок под номерами 14–17. в таблице II. и 26–33. в таблице III. Они находятся под номерами 18. в таблице II., 34–37. в таблице III. и 68–86 в таблице IV.

Таблица I.

Размерность (n)	2	3	4	5	6	7	8	9	10
Неэквивалентные	1	1	2	3	5	9	18	37	90
Свободные	1	1	1	2	3	4	6	10	18
Не свободные	0	0	1	1	2	5	12	27	72
Неизоморфные	1	1	2	2	3	4	5	5	6

Таблица II.

Неэквивалентные доп. центрировки; $n \leqq 8$, $m = 2$

Номер	Индекс	Число 1 в опред. строк.	Минимальная база	
1	1	0	(00000000)	своб.
2	2	4	(11110000)	несвоб.
3	2	5	(11111000)	своб.
4	2	6	(11111100)	своб.
5	2	7	(11111110)	своб.
6	2	8	(11111111)	своб.
7	4	4,4,4	(11110000) (11001100)	несвоб.
8	4	4,4,6	(11110000) (10001110)	несвоб.
9	4	4,5,5	(11110000) (11001110)	несвоб.
10	4	4,4,8	(11110000) (00001111)	несвоб.
11	4	4,5,7	(11110000) (00011111)	несвоб.
12	4	4,6,6	(11110000) (00111111)	несвоб.
13	4	5,5,6	(11111000) (00011111)	своб.
14	8	4,4,4,4,4,4,4	(11110000) (11001100) (10101010)	несвоб.
15	8	4,4,4,4,4,4,8	(11110000) (11001100) (11000011)	несвоб.
16	8	4,4,4,4,4,6,6	(11110000) (11001100) (10100011)	несвоб.
17	8	4,4,4,5,5,5,5	(11110000) (11001100) (10101011)	несвоб.
18	16	4,4,4,4,4,4,4,4 4,4,4,4,4,4,8	(11110000) (11001100) (10101010) (01101001)	несвоб.

Таблица III.

Неэквивалентные доп. центрировки, $n = 9$, $m = 2$

Номер	Индекс	Число 1 в опред. строк.	Минимальная база	
19	2	9	(111111111)	своб.
20	4	4,5,9	(111100000) (000011111)	несвоб.
21	4	4,6,8	(111100000) (000111111)	несвоб.
22	4	4,7,7	(111100000) (001111111)	несвоб.
23	4	5,5,8	(111110000) (000011111)	своб.
24	4	5,6,7	(111110000) (000111111)	своб.
25	4	6,6,6	(111111000) (000111111)	своб.
26	8	4,4,4,4,6,6,8	(111100000) (110011000) (100000111)	несвоб.
27	8	4,4,4,5,5,5,9	(111100000) (110011000) (110000111)	несвоб.
28	8	4,4,4,5,5,7,7	(111100000) (110011000) (101000111)	несвоб.
29	8	4,4,4,6,6,6,6	(111100000) (110011000) (101010111)	несвоб.
30	8	4,4,4,6,6,6,6	(111100000) (100011100) (010010011)	несвоб.
31	8	4,4,5,5,5,6,7	(111100000) (100011100) (110010011)	несвоб.
32	8	4,5,5,5,5,6,6	(111100000) (110011100) (101010011)	несвоб.
33	8	4,4,5,5,5,5,8	(111100000) (110011100) (110010011)	несвоб.
34	16	4,4,4,4,4,4,4,4 4,4,6,6,6,6,8	(111100000) (110011000) (101010100) (110000011)	несвоб.

Таблица III.

Продложение

Номер	Индекс	Число 1 в опред. строк.	Минимальная база	
35	16	4,4,4,4,4,4,4 5,5,5,5,5,5,5,9	(111100000) (110011000) (101010100) (011010011)	несвоб.
36	16	4,4,4,4,4,4 5,5,5,5,5,5,5,5,8	(111100000) (110011000) (110000110) (101010101)	несвоб.
37	16	4,4,4,4,4,4,4,4,4 6,6,6,6,6,6	(111100000) (110011000) (101000110) (100010101)	несвоб.

3.4. $p = 32$. По Леммами 1 и 8 такие центрировки могут существовать только в 9 и 10-мерных пространствах. Но центрировку под номером 18. в таблице II. нельзя добавить так, чтобы мы получили центрировку с индексом $p = 32$, в 9-мерном пространстве, поэтому такие центрировки существуют только при $n = 10$. Они находятся в таблице IV. под номерами 87–90.

В таблице VI. видно один из них. Там находятся все её определяющие строки.

3.5. $p > 32$. По Леммами 1 и 8 такой центрировки не существуют.

3.6. По Лемме 6. и по таблицам II–IV. уже легко получим неизоморфные центрировки (см. в таблице V.).

3.7. В конце концов, по таблицам II–V. уже не трудно получить таблицу I., и так теорема доказана.

Таблица IV

Неэквивалентные доп. центрировки, $n = 10$, $m = 2$

Номер	Индекс	Число 1 в опред. строк.	Минимальная база	
38	2	10	(1111111111)	своб.
39	4	4,6,10	(1111000000) (0000111111)	несвоб.
40	4	4,7,9	(1111000000) (0001111111)	несвоб.
41	4	4,8,8	(1111000000) (0011111111)	несвоб.
42	4	5,5,10	(1111100000) (0000011111)	своб.
43	4	5,6,9	(1111100000) (0000111111)	своб.
44	4	5,7,8	(1111100000) (0001111111)	своб.
45	4	6,6,8	(1111110000) (0000111111)	своб.
46	4	6,7,7	(1111110000) (0001111111)	своб.
47	8	4,4,4,4,8,8,8	(1111000000) (1100110000) (0000001111)	несвоб.
48	8	4,4,4,5,7,7,9	(1111000000) (1100110000) (1000001111)	несвоб.
49	8	4,4,4,6,6,6,10	(1111000000) (1100110000) (1100001111)	несвоб.
50	8	4,4,4,6,6,8,8	(1111000000) (1100110000) (1010001111)	несвоб.
51	8	4,4,4,7,7,7,7	(1111000000) (1100110000) (1010101111)	несвоб.
52	8	4,4,4,6,6,6,10	(1111000000) (1000111000) (1000000111)	несвоб.
53	8	4,4,4,6,6,8,8	(1111000000) (1000111000) (0100000111)	несвоб.

Таблица IV.

Продолжение I.

Номер	Индекс	Число 1 в опред. строк.	Минимальная база	
54	8	4,4,5,5,6,7,9	(1111000000) (1000111000) (1100000111)	несвоб.
55	8	4,4,5,6,7,7,7	(1111000000) (1000111000) (0100100111)	несвоб.
56	8	4,4,6,6,6,6,8	(1111000000) (1000111000) (1100100111)	несвоб.
57	8	4,4,5,5,7,7,8	(1111000000) (1100111000) (0000100111)	несвоб.
58	8	4,5,5,5,5,6,10	(1111000000) (1100111000) (1100000111)	несвоб.
59	8	4,5,5,5,5,8,8	(1111000000) (1100111000) (1010000111)	несвоб.
60	8	4,5,5,5,6,7,8	(1111000000) (1100111000) (1000100111)	несвоб.
61	8	4,5,5,5,6,6,9	(1111000000) (1100111000) (0000110111)	несвоб.
62	8	4,5,5,6,6,7,7	(1111000000) (1100111000) (1010100111)	несвоб.
63	8	4,4,6,6,6,6,8	(1111000000) (0000111100) (1100110011)	несвоб.
64	8	4,5,5,6,6,7,7	(1111000000) (0001111100) (1000110011)	несвоб.
65	8	4,6,6,6,6,6,6	(1111000000) (0011111100) (1010110011)	несвоб.
66	8	5,5,5,5,6,6,8	(1111100000) (0001111100) (1100010011)	своб.
67	8	5,5,5,6,6,6,7	(1111100000) (0001111100) (1001010011)	своб.

Таблица IV.

Продолжение II.

Номер	Индекс	Число 1 в опред. строк.	Минимальная база	
68	16	4,4,4,4,4,4,4,4 6,6,6,6,8,8,8	(1111000000) (1100110000) (1010101000) (1000000111)	несвоб.
69	16	4,4,4,4,4,4,4 5,5,5,7,7,7,7,9	(1111000000) (1100110000) (1010101000) (1100000111)	несвоб.
70	16	4,4,4,4,4,4,4 6,6,6,6,6,6,6,10	(1111000000) (1100110000) (1010101000) (0110100111)	несвоб.
71	16	4,4,4,4,4,4,4,4 4,4,8,8,8,8,8	(1111000000) (1100110000) (1100001100) (1100000011)	несвоб.
72	16	4,4,4,4,4,4,4,4 6,6,6,6,8,8,8	(1111000000) (1100110000) (1100001100) (1010000011)	несвоб.
73	16	4,4,4,4,4,4,5,5 5,5,7,7,7,7,8	(1111000000) (1100110000) (1100001100) (1010100011)	несвоб.
74	16	4,4,4,4,4,4,6,6 6,6,6,6,6,6,8	(1111000000) (1100110000) (1100001100) (1010101011)	несвоб.
75	16	4,4,4,4,4,4,4 6,6,6,6,6,6,6,10	(1111000000) (1100110000) (1010001100) (0110000011)	несвоб.
76	16	4,4,4,4,4,4,4, 6,6,6,6,6,6,8,8	(1111000000) (1100110000) (1010001100) (1000100011)	несвоб.
77	16	4,4,4,4,4,4,6,6, 6,6,6,6,6,6,8	(1111000000) (1100110000) (1010001100) (0000101011)	несвоб.
78	16	4,4,4,4,4,5,5,5, 5,5,6,6,7,7,9	(1111000000) (1100110000) (1010001100) (1010100011)	несвоб.

Таблица IV.

Продолжение III.

Номер	Индекс	Число 1 в опред. строк.	Минимальная база	
79	16	4,4,4,4,4,5,5, 5,5,6,6,7,7,7,7	(1111000000) (1100110000) (1010001100) (1000101011)	несвоб.
80	16	4,4,4,4,5,5,5,5, 5,5,6,6,7,7,8	(1111000000) (1100110000) (1010101100) (1000001011)	несвоб.
81	16	4,4,4,4,5,5,5,5, 5,5,5,5,8,8,8	(1111000000) (1100110000) (1010101100) (0000001111)	несвоб.
82	16	4,4,4,5,5,5,5, 5,5,5,5,6,6,6,10	(1111000000) (1100110000) (1010101100) (0110100011)	несвоб.
83	16	4,4,4,5,5,5,5,5, 5,5,5,6,6,6,6,9	(1111000000) (1100110000) (1010101100) (1100001011)	несвоб.
84	16	4,4,4,4,4,6,6,6, 6,6,6,6,6,6,6	(1111000000) (1000111000) (0100100110) (0010010101)	несвоб.
85	16	4,4,4,5,5,5,5, 5,5,6,6,6,6,7,7	(1111000000) (1000111000) (0100100110) (1010010101)	несвоб.
86	16	4,4,5,5,5,5,5, 5,5,6,6,6,6,8	(1111000000) (1100111000) (1010100110) (0000110101)	несвоб.
87	32	4,4,4,4,4,4,4,4, 4,4,4,4,4,4,4,4, 4,4,6,6,6,6,6,6, 6,6,8,8,8,8,8	(1111000000) (1100110000) (1010101000) (0110100100) (1100000011)	несвоб.
88	32	4,4,4,4,4,4,4,4, 4,4,4,4,4,4,4,4, 6,6,6,6,6,6,6,6, 6,6,6,6,8,8,8	(1111000000) (1100110000) (1010101000) (1100000110) (1010000101)	несвоб.

Таблица IV.

Продолжение IV.

Номер	Индекс	Число 1 в опред. строк	Минимальная база	
89	32	4,4,4,4,4,4,4,4, 4,4,5,5,5,5,5,5, 5,5,5,5,5,5,5,5, 5,5,8,8,8,8,8	(1111000000) (1100110000) (1100001100) (1010101010) (1100000011)	несвоб.
90	32	4,4,4,4,4,4,4,4, 4,4,4,4,4,4,4,4, 6,6,6,6,6,6,6,6, 6,6,6,6,6,6,6,10	(1111000000) (1100110000) (1010001100) (1000101010) (0001101001)	несвоб.

Таблица V.

Неизоморфные доп. центрировки, $m = 2$

Номер	Индекс p	Размерность $n \geqq$	Минимальная база	Номера изоморфных центрировок в таблицах II–IV.
1	1		(0000000000)	II. 1.
2	2	4	(1111000000)	II. 2–6., III. 19., IV. 38
3	$4 = 2 \cdot 2$	6	(1111000000) (0011110000)	II. 7–13., IV. 20–25., IУ. 39–46.
4	$8 = 2 \cdot 2 \cdot 2$	7	(1111000000) (1100110000) (0101101000)	II. 14–17., III. 26–33. IV. 47–67.
5	$16 =$ $= 2 \cdot 2 \cdot 2 \cdot 2$	8	(1111000000) (1100110000) (1010101000) (0110100100)	II. 18., III. 34–37. IV. 68–86.
6	$32 =$ $= 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2$	10	(1111000000) (1100110000) (1010101000) (0110100100) (1100000011)	IV. 87–90.

Таблица VI.

Пример

Минимальная база $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4, \mathbf{a}_5$	$q=5$	Размерность $n=10$	Знаменатель $m=2$	Индекс $p=32$

$\mathbf{a}_1 = (1111000000)$
$\mathbf{a}_2 = (1100110000)$
$\mathbf{a}_3 = (1010101000)$
$\mathbf{a}_4 = (0110100100)$
$\mathbf{a}_5 = (1100000011)$
$\mathbf{a}_6 = (1100001100)$
$\mathbf{a}_7 = (0000001111)$
$\mathbf{a}_8 = (0110011000)$
$\mathbf{a}_9 = (1010010100)$
$\mathbf{a}_{10} = (0000110011)$
$\mathbf{a}_{11} = (0000111100)$
$\mathbf{a}_{12} = (0011110000)$
$\mathbf{a}_{13} = (0101101000)$
$\mathbf{a}_{14} = (1001100100)$
$\mathbf{a}_{15} = (0011000011)$
$\mathbf{a}_{16} = (0011001100)$
$\mathbf{a}_{17} = (1001011000)$
$\mathbf{a}_{18} = (0101010100)$
$\mathbf{a}_{19} = (1010100111)$
$\mathbf{a}_{20} = (0110101011)$
$\mathbf{a}_{21} = (0110010111)$
$\mathbf{a}_{22} = (1010011011)$
$\mathbf{a}_{23} = (0101100111)$
$\mathbf{a}_{24} = (1001101011)$
$\mathbf{a}_{25} = (1001010111)$
$\mathbf{a}_{26} = (0101011011)$
$\mathbf{a}_{27} = (0011111111)$
$\mathbf{a}_{28} = (1100111111)$
$\mathbf{a}_{29} = (1111001111)$
$\mathbf{a}_{30} = (1111110011)$
$\mathbf{a}_{31} = (1111111100)$
$\mathbf{a}_{32} = (0000000000)$

Литература

[1] Делоне В. Н., Фаддеев Д. К.: Теория иррациональностей третей степены, *Труды МИАН СССР*, **11** (1940), 63–69.

[2] Hoffreiter N.: Zur Geometrie der Zahlen, *Monatsh. Math. Phys.*, **40** (1933), 181–192.

[3] Рыщков С. С.: К проблеме отыскания совершенных квадратичных форм от многих переменных, *Труды МИАН СССР*, **142** (1976), 215–239.

[4] Захарова Н. В.: Центрировки 8-мерных решёток, сохраняющие репер последовательных минимумов, *Труды МИАН СССР*, **152** (1980), 97–123.

[5] Ленг С.: *Алгебра*, Издательство «Мир», Москва, 1968. (Addison–Wesley Publishing Company; Reading, MASS, 1965.)

[6] Майор З.: О центрировке решёток, *Annales Univ. Sci. Budapest*, Sect. Math., **28** (1985), 165–172.

A REMARK ON SIGNUM TYPE ORTHONORMAL SYSTEMS

By

N. H. LOI and M. HORVÁTH

II. Department for Analysis of the L. Eötvös University, Budapest

(Received October 26, 1983)

Consider the orthogonal series

$$\sum_{n=1}^{\infty} a_n \varphi_n(x), \tag{1}$$

where $\{a_n\}$ is a sequence of real numbers and $\{\varphi_n\}$ is an orthonormal system on the interval (0, 1) (briefly: ONS). The series (1) is said to be convergent almost everywhere (a. e.) unconditionally if it is convergent in every rearrangement

$$\sum_{l=1}^{\infty} a_{n(l)} \varphi_{n(l)}(x) \tag{2}$$

a. e. (The set of divergence points may depend on the rearrangement). The orthonormal system $\{\varphi_n\}$ is said to be signum type, if $|\varphi_n(x)| = 1$ a. e. on (0, 1) for $n = 1, 2, \ldots$.

For any $\{a_n\} \in l_2$ denote by $\{a_n^*\}$ such a rearrangement of $\{a_n\}$ for which $|a_1^*| \geq |a_2^*| \geq \ldots$ and define

$$S = \sum_{k=0}^{\infty} \left(\sum_{n=\nu(k)+1}^{\nu(k+1)} a_n^{*2} \log^2 n \right)^{1/2} \quad \left(\nu(k) \overset{\text{def}}{=} 2^{2^k}\right).$$

I. Joó proved in [2]: If $S = \infty$, then there exists a signum type orthonormal system and a rearrangement $\{n(l)\}$ such that the series (2) is divergent a. e. on (0, 1). His theorem results from the following.

LEMMA. Let $c_0 \leq k_1 < k_2$ be integers (where c_0 and later on $c_1, c_2, \ldots$ are absolute constants) and $\{a_n\}$ $(n = \nu(k_1)+1, \ldots, \nu(k_2)$ a sequence monotone decreasing in absolute value. Then there is a constant c_1 such that if

$$\sum_{k=k_1}^{k_2} \left(a_{\nu(k)+1}^2 + \sum_{l=2}^{\nu(k+1)-\nu(k)} a_{\nu(k)+l}^2 \log^2 l \right)^{1/2} \geq c_1 \left(\sum_{n=\nu(k_1)+1}^{\nu(k_2)} a_n^2 \right)^{1/2}. \tag{3}$$

then there exists a signum type orthonormal system $\{\varphi_n\}$ $(n = \nu(k_1)+1, \ldots, \ldots, \nu(k_2))$ of stepfunctions and a simple set $E \subset (0, 1)$ (i. e. E is the union of finitely many intervals) with mes $E \geq c_2$ such that the series

$$\sum_{n=\nu(k_1)+1}^{\nu(k_2)} a_n \varphi_n(x)$$

has a rearrangement

$$\sum_{l=1}^{\nu(k_2)-\nu(k_1)} a_{m(l)} \varphi_{m(l)}(x)$$

for which

$$\max_{1 \leq \lambda \leq \nu(k_2)-\nu(k_1)} \sum_{l=1}^{\lambda} a_{m(l)} \varphi_{m(l)}(x) \geq c_3 \sum_{k=k_1}^{k_2} \left(a_{\nu(k)+1}^2 + \sum_{l=1}^{\nu(k+1)-\nu(k)} a_{\nu(k)+l}^2 \log^2 l \right)^{1/2} \tag{4}$$

$$(x \in E)$$

is fulfilled.

For the proof of this Lemma in [2] are used independent functions and the well known Kolmogorov's inequality for them (see also in [3]).

The aim of the present note is to show that this Lemma follows by an elementary construction. Our method may be useful also for other related problems.

Indeed. Use the notation

$$A_k \overset{\text{def}}{=} \left(a_{\nu(k)+1}^2 + \sum_{l=2}^{\nu(k+1)-\nu(k)} a_{\nu(k)+l}^2 \log^2 l \right)^{1/2}.$$

It is well known [2], that there exists a 1-bounded orthogonal system $\{\psi_n\}$ $(n = \nu(k_1)+1, \ldots, \nu(k_2))$ from stepfunctions (i. e. $|\psi_n(x)| \leq 1$ a. e. on $(0, 1)$) and a simple set $E^* \subset (0, 1)$ with mes $E^* \geq c_2^*$ such that for an appropriate rearrangement $\{m(l)\}$ the relation (4) is fulfilled for $\{\psi_n\}$ in every $x \in E^*$, with some c_3^*. Here c_2^* and c_3^* are absolute constants. Now we construct from the system $\{\psi_n\}$ a signum type orthonormal system $\{\varphi_n\}$ satisfying the requirements of the Lemma.

To this let $I_1, \ldots, I_N$ be a partition of $(0, 1)$ into disjoint intervals so that every function ψ_n $(n = \nu(k_1)+1, \ldots, \nu(k_2))$ is constant on every I_i $(i = 1, \ldots, N)$, further

$$E^* = \bigcup_{\substack{i \\ E^* \cap I_i \neq 0}} I_i. \tag{5}$$

Denote $\vartheta_{i,j} = \psi_{j+\nu(k_1)}(x)$ for $x \in I_i$, $i = 1, \ldots, N$ and $j = 1, \ldots, \nu(k_2)-\nu(k_1)$; we know that $|\vartheta_{i,j}| \leq 1$.

We construct a series $\{f_{i,j}\}$, $i = 1, \ldots, N$ and $j = 1, \ldots, \nu(k_2)-\nu(k_1)$ of stepfunctions such that

$$\operatorname{supp} f_{i,j} \subset I_i, \tag{6}$$

$$f_{i,j}(x) = 1-\vartheta_{i,j} \text{ or } -1-\vartheta_{i,j} \text{ on } I_i, \tag{7}$$

$$\int_{I_i} f_{i,j} = 0, \tag{8}$$

$$\int_{I_i} f_{i,j_1} f_{i,j_2} = 0 \text{ if } j_1 \neq j_2 \tag{9}$$

holds for all possible choices of the indices i and j.

Denote $I_i = (a_i, b_i)$. Let

$$f_{i,1}(x) = \begin{cases} 1-\vartheta_{i,1}, & \text{if } x \in \left(a_i, \dfrac{(1-\vartheta_{i,1})a_i+(1+\vartheta_{i,1})b_i}{2}\right), \\ -1-\vartheta_{i,1}, & \text{if } x \in \left(\dfrac{(1-\vartheta_{i,1})a_i+(1+\vartheta_{i,1})b_i}{2}, b_i\right), \\ 0 & \text{otherwise.} \end{cases}$$

Suppose $f_{i,j}$ is given. Let $f_{i,j+1}(x) = 0$ if $x \notin I_j$. The interval I_i can be divided into (maximal) subintervals on which $f_{i,j}$ is constant. On each of these subintervals we give $f_{i,j+1}$ as follows:

Denote by (a, b) a subinterval described above.

Define

$$c = \frac{1}{2}[a(1-\vartheta_{i,j+1})+b(1+\vartheta_{i,j+1})]$$

and let

$$f_{i,j+1}(x) = \begin{cases} 1-\vartheta_{i,j+1} & \text{if } x \in (a, c), \\ -1-\vartheta_{i,j+1} & \text{if } x \in (c, b). \end{cases}$$

Since $\int_a^b f_{i,j+1} = 0$, it is easy to verify (6) (7) (8) and (9). Having constructed the $f_{i,j}'$ s, define for $x \in I$

$$\varphi_{j+\nu(k_1)}(x) = \psi_{j+\nu(k_1)}(x) + \sum_{i+1}^{N} f_{i,j+\nu(k_1)}(x).$$

By (8) and (9) the system $\{\varphi_n\}$ $n = \nu(k_1)+1, \ldots, \nu(k_2)$ is a signum type ONS. To prove our Lemma we have to verify (4).

Define $1 \leq \lambda_i \leq \nu(k_2)-\nu(k_1)$ such that for $x \in I_i$

$$\max_{1 \leq \lambda \leq \nu(k_2)-\nu(k_1)} \sum_{l=1}^{\lambda} a_{m(l)}\psi_{m(l)}(x) = \sum_{l=1}^{\lambda_i} a_{m(l)}\psi_{m(l)}(x).$$

Applying the inequality

$$\operatorname{mes}\{x\in I: |f(x)|>\varepsilon\}\leq\frac{1}{\varepsilon^2}\int_I f^2$$

we get from (3), (9):

$$\operatorname{mes}\left\{x\in I_i: \left|\sum_{l=1}^{\lambda_i} a_{m(l)}f_{i,\,m(l)}(x)\right|>\frac{c_3^*}{2}\sum_{k=k_1}^{k_2}A_k\right\}\leq$$

$$\leq\frac{4\sum\limits_{n=\nu(k_1)+1}^{\nu(k_2)} a_n^2}{\left(\frac{c_3^*}{2}\sum\limits_{k=k_1}^{k_2}A_k\right)^2}\cdot\operatorname{mes} I_i\leq c_4\operatorname{mes} I_i.$$

Here $c_4=\dfrac{16}{c_3^{*2}c_1^2}$ we can suppose that the constant c_1 is great, namely that $c_4<\dfrac{1}{2}$. Now for $I_i\subset E^*$ we have

$$\operatorname{mes}\left\{x\in I_i: \max_{1\leq\lambda\leq\nu(k_2)-\nu(k_1)}\sum_{l=1}^{\lambda} a_{m(l)}\varphi_{m(l)}(x)\geq\frac{c_3^*}{2}\sum_{k=k_1}^{k_2}A_k\right\}\geq$$

$$\geq\operatorname{mes}\left\{x\in I_i: \sum_{l=1}^{\lambda_i} a_{m(l)}\varphi_{m(l)}(x)\geq\frac{c_3^*}{2}\sum_{k=k_1}^{k_2}A_k\right\}\geq$$

$$\geq\operatorname{mes}\left\{x\in I_i: \sum_{l=1}^{\lambda_i} a_{m(l)}\psi_{m(l)}(x)>c_3^*\sum_{k=k_2}^{k_2}A_k\right\}-$$

$$-\operatorname{mes}\left\{x\in I_i: \left|\sum_{l=1}^{\lambda_i} a_{(ml)}f_{l,\,m(l)}(x)\right|>\frac{c_3^*}{2}\sum_{k=k_1}^{k_2}A_k\right\}>\frac{1}{2}\operatorname{mes} I_i.$$

We obtained the statement of the Lemma with $c_3=\dfrac{c_3^*}{2}$ and $c_2=\dfrac{c_2^*}{2}$.

References

[1] G. Alexits, *Convergence problems of orthogonal series*, Pergamon Press, (Oxford, 1961).
[2] I. Joó, On signum type orthonormal systems, *Analysis Mathematica*, 4 (1978), 17–26.
[3] K. Tandori, Über die unbedingte Konvergenz der Orthogonalreihen, *Acta Sci. Math.* (Szeged), 32 (1971), 11–40.

ON THE USE OF ESPIONAGE IN A CLASS OF POSITIONAL GAMES

By

L. A. SZÉKELY

Bolyai Institute of the József Attila University, Szeged

(Received September 22, 1983)

In this paper we consider positional games from an unusual point of view. We are going to calculate the use of espionage in the simplest situation where it can be asked for. As far as we know this problem is new at all, so we have no references.

The simplest situation is the following one. Two players, Red and Blue are playing on the set of natural numbers $N = \{1, 2, \ldots, n\}$. They colour alternately one of the uncolured numbers of N with Red colouring first. Before the game Blue fixed a set $A \subset N$ ($|A| = k \leq n/2$) and his goal is to colour the numbers of A till the end of the game. Red does not know A, so, it is wise for him to act randomly.

Red has two main possibilities to make use of random. On the one hand, Red may colour randomly and independently with the same probability one number of the uncoloured rest by move. Now there is nothing to know for Blue. He has two choices when he is on move: to colour a number of the rest of A or to colour a number of the rest of $N-A$. We define a program for Blue to be a function that says what to do of the above mentioned two choices. Let $q_{n,k}$ denote the probability of the event that Blue can colour A. (It depends on Blue's program, of course.) Let $Q_{n,k}$ denote $q_{n,k}$ if the program is to colour an element of A as long as it is possible.

On the other hand, Red may choose a strategy randomly, Strategy means a function whose domain is the set of positions in which Red is to move. Two positions are different iff they evolved on different ways. Its range is the set of possible moves. All the strategies have the same probability. The notion of espionage means that Blue knows the randomly chosen strategy. Let $p_{n,k}$ denote the probability of the following event: Blue has possibility to colour A.

There are absolutely different limit distributions of $Q_{n,k}$ and $p_{n,k}$. This difference is that, what we call as "use of espionage".

THEOREM 1. *Every program of Blue satisfies* $q_{n,k} \leq Q_{n,k}$.

If $k = o(\sqrt{n})$, *then* $Q_{n,k} \to 1$;

$k = c\sqrt{n} + o(\sqrt{n})$, *then* $Q_{n,k} \to e^{-c^2/2}$;

$k/\sqrt{n} \to \infty$, *then* $Q_{n,k} \to 0$.

PROOF. Because of the independent choices $q_{n,k}$ is a product of probabilities. So

$$Q_{n,k} = \frac{n-k}{n} \cdot \frac{n-k-1}{n-2} \cdots \frac{n-2k}{n-2k},$$

and it is easy to see that $q_{n,k} \leq Q_{n,k}$ factor by factor. Using the asymptotic formula

$$e^{-\frac{k}{n}} \sim 1 - \frac{k}{n} \quad \text{if} \quad \frac{k}{n} \to 0$$

and the inequalities

$$\frac{k(k-1)}{2} \cdot \frac{1}{n} \leq \frac{k}{n} + \frac{k-1}{n-2} + \cdots + \frac{0}{n-2k} \leq \frac{k(k-1)}{2} \cdot \frac{1}{n-2k}$$

we have our statements. ■

THEOREM 2. *If* $k \leq n/2$, *then* $p_{n,k} \sim 1 - \frac{k}{n}$.

PROOF. Let $S(n)$ denote the number of Red strategies, $F(n, k)$ denote the number of those Red strategies which give possibility for Blue to colour A if he spied.

It is obvious, that $p_{n,k} = F(n, k)/S(n)$. We make recursive formulae for $S(n)$ and $F(n, k)$.

There are n possible choices for Red in the first move, all these choices allow $n-1$ choices for Blue in the first move. Having done the first pair of moves the number of the Red strategies on the set of the remaining $n-2$ points is $S(n-2)$. After any of the $n-1$ possible Blue choices in the first move Red can continue his play by all of the $S(n-2)$ then possible strategies.

So we have $S(n) = n \cdot S(n-2)^{n-1}$.

Suppose $p \in N-A$. Let us define $B_p(X)$ to be the set of those Red strategies, in which $X \subset N-A$, (i) and (ii).

(i) Red's first move is p,

(ii) Blue can win if he choose in the first move whichever element of X.

Using the sieve we have:

$$F(n,k) = \sum_{P \in N-A} \sum_{\emptyset \neq x < N-\{P\}} (-1)^{|x|-1} |B_p(x)|.$$

Using the notations $|X| = x$, $|X \cap A| = m$, we have

$$|B_p(x)| = F(n-2, k-1)^m F(n-2, k)^{x-m} S(n-2)^{n-x-1}$$

since an element of $X \cap A$ can be followed by any Red strategy that allows Blue's win on $k-1$ of $n-2$ points, an element of $X-A$ can be followed by any Red strategy that allows Blue's win on k of $n-2$ points, and an element of $N-\{p\}-X$ can be followed by any Red strategy on $n-2$ points. We have

$$\sum_{|x|=x} |B_p(x)| =$$

$$= \sum_{m=0}^{x} \binom{k}{m}\binom{n-k-1}{x-m} F(n-2, k-1)^m F(n-2, k)^{x-m} S(n-2)^{n-x-1}.$$

Changing the sums and using the Binomial Theorem twice

$$F(n, k) = (n-k) \sum_{a=1}^{n-1} (-1)^{x-1} \sum_{|X|=x} |B_p(x)| = (n-k)S(n-2)^{n-1} +$$

$$+ (n-k) \sum_{m=0}^{n-1} (-1)^{m-1} \binom{k}{m} F(n-2, k-1)^m \sum_{x=m}^{n-1} (-1)^{x-m} \cdot$$

$$\cdot \binom{n-k-1}{x-m} F(n-2, k)^{x-m} S(n-2)^{n-x-1} = (n-k)S(n-2)^{n-1} -$$

$$-(n-k)(S(n-2) - F(n-2, k-1))^k (S(n-2) - F(n-2, k))^{n-k-1}.$$

Dividing to $S(n)$ we have

$$p_{n,k} = \left(1 - \frac{k}{n}\right)[1 - (1 - p_{n-2,k})^{n-k-1}(1 - p_{n-2,k-1})^k].$$

It is easy to see by induction that $p_{n,\left[\frac{n}{2}\right]} \to \frac{1}{2}$, since $p_{n-2,\left[\frac{n}{2}\right]} = 0$. Hence $p_{n,\left[\frac{n}{2}\right]}$ is separated from zero. Obviously $i < j$ implies $p_{n,i} > P_{n,j}$, so, $p_{n,k}$ is separated from zero.

We have $\dfrac{p_{n,k}}{1 - \dfrac{k}{n}} = 1 + o(\varepsilon^n)$ with an $\varepsilon > 1$. ■

ASYMPTOTIC FORMULA FOR THE NUMBER OF SOLUTIONS OF A DIOPHANTIC SYSTEM

By

A. BOGMÉR and L. A. SZÉKELY

Department of Numerical Methods and Computer Science of the Eötvös University, Budapest

(Received October 31, 1983)
(Revised October 2, 1984)

1. Diophantic systems and partitions of numbers

In the present paper we are investigating the number of solutions of the following diophantic system:

(i) $1x_1+2x_2+\ldots+nx_n = N_1$,

(ii) $x_1+x_2+\ldots+x_n = N_2$,

(iii) $0 \le x_j \le p \quad (j = 1, 2, \ldots, n)$.

Although the above system has been out of regard, all of its subsystems have been intensively investigated.

It is well-known, that the number of solutions of (i), $n = N_1$ is $P(n)$, the number of partitions of the number n,

$$P(n) = \left(1 + O\left(\frac{1}{\sqrt{n}}\right)\right) \frac{1}{4\sqrt{3}\,n} e^{\frac{2\pi}{\sqrt{6}}\sqrt{n}}$$

as it was proved by HARDY and RAMANUJAN in 1918 [4]. Here the term of error $n^{-1/2}$ was changed to $n^{-1/4+\varepsilon}$ by an easier proof of FREYMAN in 1955 (see in POSTNIKOFF's book [9]). In 1943 PAUL ERDŐS found an elementary proof for [2]

$$P(n) \sim \frac{c}{n} \exp\left(\frac{2\pi}{\sqrt{6}}\sqrt{n}\right).$$

Every introduction in combinatorics contains, that the number of solutions of (ii) is $\binom{N_2+n-1}{n-1}$.

HARDY and RAMANUJAN found for the number of solutions of (i) and (iii) in case of $n = N_1$, $p = 1$ [4]

$$\frac{1+o(1)}{4\cdot 3^{1/4} n^{3/4}} \exp\left(\frac{\pi}{\sqrt{3}}\sqrt{n}\right).$$

In case of fixed p rather than $p = 1$, by the private communication of M. SZALAY is

$$(1+o(1))\frac{\left(1-\frac{1}{p+1}\right)^{3/4}}{2\cdot 6^{1/4}\,p^{1/2}\,n^{3/4}}\exp\left(\frac{2\pi}{\sqrt{6}}\sqrt{1-\frac{1}{p+1}}\,n^{1/2}\right).$$

I. Joó [6] proved the number of solutions of (i), (iii) to be

$$(p+1)^n\frac{1}{\sqrt{2\pi}\,D_n}\left[e^{-\frac{(N-A_n)}{2D_n^2}}+\frac{\Theta}{n}\right],$$

where $|\Theta|$ is bounded by an universal constant,

$$A_n=\frac{n(n+1)p}{4},\quad D_n=\frac{n(n+1)(2n+1)p(p+2)}{72}.$$

The system (ii), (iii) was investigated by CASTELNOUVO [1], and in case of more generality by POSTNIKOFF. POSTNIKOFF proved that the number of solutions is ([9], [10]):

$$(p+1)^n\frac{1}{\sqrt{2\pi n}\,\delta}\left[e^{-\frac{(N-a_n)^2}{2n\delta^2}}+O\left(\frac{(p+1)^{n-1}}{n}\right)\right],$$

where

$$a=\frac{p+1}{2},\quad \delta^2=\frac{1}{p+1}\sum_{x=0}^{P}x^{2p}-\left(\frac{p+1}{2}\right)^2.$$

We have found the following theorem.

THEOREM 1. *Suppose* $3\le n$, $1\le p\le(1{,}0014)^{n/4}n^{-3/2}$, N_1, N_2 *are natural numbers. Let us define*

$$A_{n,1}=\frac{np}{2},\qquad A_{n,2}=\frac{n(n+1)p}{4},$$

$$D_{n,1}^2=\frac{np(p+2)}{12},\qquad D_{n,2}^2=\frac{n(n+1)(2n+1)p(p+2)}{72},$$

$$h_n=\sqrt{\frac{3n+3}{4n+2}},\qquad Q(u_1,u_2)=\frac{1}{1-h_n^2}(u_1^2-2h_nu_1u_2+u_2^2).$$

Now we have $R_n(N_1,N_2)$ *the number of solutions of* (i), (ii), (iii)

$$R_n(N_1,N_2)=\frac{(p+1)^n}{D_{n,1}D_{n,2}}\left[\frac{e^{-\frac{1}{2}Q\left(\frac{N_1-A_{n,1}}{D_{n,1}},\ \frac{N_2-A_{n2}}{D_{n,2}}\right)}}{2\pi\sqrt{1-h_n^2}}+\frac{\Theta}{n}\right],$$

where $|\Theta|\le 10^5$.

The proof of the theorem is contained in the following two sections of the paper. We apply some techniques due to Joó and to the book of SIRASHDINOFF, AZLAROFF and ZUPAROFF [12]. The final section is devoted to an-application to Wilcoxon statistics.

2. The proof of the theorem

Let us consider for $k = 1, 2, \ldots, n$ the independent random vectors $\boldsymbol{\xi}_k$ as follows ($x = 0, 1, \ldots, p$):

$$\boldsymbol{\xi}_k = (\xi_{k,1}, \xi_{k,2}), \quad P(\boldsymbol{\xi}_k = (x,kx)) = \frac{1}{p+1}.$$

Set

$$S_{n,j} = \sum_{k=1}^{n} \xi_{k,j} \quad (j = 1, 2),$$

$$P_n(N_1, N_2) = P((S_{n,1}, S_{n,2}) = (N_1, N_2)).$$

It is easy to check, that

$$E(S_{n,1}) = A_{n,1}, \quad E(S_{n,2}) = A_{n,2},$$
$$D^2(S_{n,}) = D_{n,}{}^2{}_1, \quad D^2(S_{n,2}) = D^2_{n,2},$$

where the quantities $A_{n,i}$, $D_{n,i}$ are defined in the theorem. Set

$$D_{n,1,2} = E\left\{\sum_{k=1}^{n} (\xi_{k,1} - E(\xi_{k,1}))(\xi_{k,2} - E(\xi_{k,2}))\right\},$$

$h_n = \dfrac{D_{n,1,2}}{D_{n,1}D_{n,2}}$, Easy computation gives $h_n = \sqrt{\dfrac{3n+3}{4n+2}}$. Let us consider the characteristic functon $f_n(t_1, t_2)$ of the random vector variable $(S_{n,1}, S_{n,2})$:

$$f_n(t_1, t_2) = E\{e^{i(S_{n,1}t_1 + S_{n,2}t_2)}\} = \prod_{k=1}^{n} \frac{1}{p+1} \sum_{x=0}^{p} e^{i(t_1 x + t_2 kx)} =$$

$$(1) \qquad = \frac{1}{(p+1)^n} e^{i(A_{n,1}t_1 + A_{n,2}t_2)} \prod_{k=1}^{n} \frac{\sin (p+1) \dfrac{t_1 + kt_2}{2}}{\sin \dfrac{t_1 + kt_2}{2}},$$

and the characteristic function $f_n^*(t_1, t_2)$ of the normed random vector variable $\left(\dfrac{S_{n,1} - A_{n,1}}{D_{n,1}}, \dfrac{S_{n,2} - A_{n,2}}{D_{n,2}}\right)$:

$$(2) \qquad f_n^*(t_1, t_2) = E \exp i\left(\frac{S_{n,1} - A_{n,1}}{D_{n,1}} t_1 + \frac{S_{n,2} - A_{n,2}}{D_{n,2}} t_2\right) =$$

$$= f_n\left(\frac{t_1}{D_{n,1}}, \frac{t_2}{D_{n,2}}\right) \exp -i\left(\frac{A_{n,1}}{D_{n,1}} t_1 + \frac{A_{n2}}{D_{n2}} t_2\right).$$

It is important, that f_n^* is real-valued. The following identity is easy to check:

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f_n(t_1, t_2)e^{-i(t_1N_1+t_2N_2)}\,dt_1,dt_2 = \tag{3}$$

$$= \frac{1}{(p+1)^n}\int_{\pi}^{\pi}\int_{\pi}^{\pi}\sum_{x_1, x_2, \ldots, x_n=0}^{p} e^{i[t_1(x_1+x_2+\ldots+x_n-N_1)+t_2(1\cdot x_1+2x_2+\ldots+nx_n-N_2)]}\,dt_1\,dt_2 =$$

$$= \frac{4\pi^2}{(p+1)^2}R_n(N_1, N_2) = 4\pi^2P_n(N_1, N_2).$$

Let us consider the quadratic forms

$$q(t_1, t_2) = t_1^2+2h_nt_1t_2+t_2^2,$$

$$Q(u_1, u_2) = \frac{1}{1-h_n^2}(u_1^2-2h_nu_1u_2+u_2^2).$$

The density of the two-dimensional standard normal distribution is

$$\varphi(u_1, u_2) = \frac{1}{2\pi\sqrt{1-h_n^2}}\,e^{-\frac{1}{2}Q(u_1u_2)}$$

[3]. We recall the following identity: if ψ is a positive definite quadratic form with the symmetric matrix A and ψ^* is the quadratic form of the matrix A^{-1}, then

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\pi\psi(x_1, x_2)}e^{2\pi i(u_1x_1+u_2x_2)}\,dx_1\,dx_2 = \frac{1}{\sqrt{\det A}}e^{-\pi\psi^*(u_1, u_2)} \tag{4}$$

(see [7] Chapter XI, section 2). Let us apply (4) to the matrices

$$A = \frac{1}{2\pi}\begin{pmatrix}1 & h_n\\ h_n & 1\end{pmatrix}, \quad A^{-1} = \frac{2\pi}{1-h_n^2}\begin{pmatrix}1 & -h_n\\ -h_n & 1\end{pmatrix}.$$

We have

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i(u_1t_1+u_2t_2)-\frac{1}{2}q(t_2, t_2)}\,dt_1dt_2 = 4\pi^2\varphi(u_1, u_2). \tag{5}$$

It is easy to get from (2), (3)

$$4\pi^2P_n(N_1, N_2) = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f_n(t_1, t_2)e^{-i(N_1t_1+N^2t_2)}\,dt_1, dt_2 = \tag{6}$$

$$= \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f_n^*(D_{n,1}t_1, D_{n2}t_2)e^{i[(A_{n,1}-N_1)t_1+(A_{n,2}-N_2)t_2]}dt_1\,dt_2 =$$

$$= \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\cos[(N_1-A_{n,1})t_1+(N_2-A_{n,2})t_2]\prod_{k=1}^{n}\frac{\sin(p+1)\dfrac{t_1+kt_2}{2}}{(p+1)\sin\dfrac{t_1+kt_2}{2}}dt_1dt_2.$$

By slight modification of (5) we have

$$4\pi^2\varphi\left(\frac{N_1-A_{n,1}}{D_{n,1}},\ \frac{N_2-A_{n,2}}{D_{n,2}}\right)= \tag{7}$$

$$= D_{n,1}D_{n,2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\cos\left[(N_1-A_{n,1})t_1+(N_2-A_{n,2})t_2\right]\cdot$$

$$\cdot\, e^{-\frac{1}{2}q(D_{n,1}t_1,\,D_{n,2}t_2)}dt_1\,dt_2.$$

Let us consider

$$O_n=\left\{(t_1,t_2):|t_1|<\frac{\sqrt{n}}{4D_{n,1}},\quad |t_2|<\frac{\sqrt{n}}{2\sqrt{12}\,D_{n,2}}\right\}. \tag{8}$$

Now multiplying (6) to $D_{n,1}$, $D_{n,2}$ and subtracting (7) we have

$$4\pi^2\left[D_{n,1}D_{n2}P_n(N_1,N_2)-\varphi\left(\frac{N_1-A_{n,1}}{D_{n,1}},\ \frac{N_2-A_{n,2}}{D_{n,2}}\right)\right]= \tag{9}$$

$$= D_{n,1}D_{n,2}[I_1+I_2-I_3],$$

where

$$I_1=\iint_{O_n}\cos\left[(N_1-A_{n,1})t_1+(N_2-A_{n,2})t_2\right]\cdot$$

$$\cdot\left\{\prod_{k=1}^{n}\cdot\frac{\sin(p+1)\dfrac{t_1+kt_2}{2}}{(p+1)\sin\dfrac{t_1+kt_2}{2}}-e^{-\frac{1}{2}q(D_{n,1}t_1,\,D_{n,2}t_2)}\right\}dt_1\,dt_2,$$

$$I_2=\iint_{[-\pi,\pi]^2-O_n}\cos\left[(N_1-A_{n,1})t_1+(N_2-A_{n,2})t_2\right]\prod_{k=1}^{n}\frac{\sin(p+1)\dfrac{t_1+kt_2}{2}}{(p+1)\sin\dfrac{t_1+kt_2}{2}}dt_1\,dt_2,$$

$$I_3=\iint_{\mathbf{R}^2-O_n}\cos\left[(N_1-A_{n,1})t_1+(N_2-A_{n,2})t_2\right]e^{-\frac{1}{2}q(D_{n,1}t_1,\,D_{n,2}t_2)}\,dt_1\,dt_2.$$

In the following we have to give good estimates for I_1, I_2, I_3.

3. The estimation of the integral I_1

This estimation is the hardest of the three ones. Our base is

$$(10)\qquad |I_1| \leq \iint\limits_{O_n} |f_n^*(D_{n,1}t_1, D_{n,2}t_2) - e^{-\frac{1}{2}q(D_{n,1}t_1,\, D_{n,2}t_2)}|\,dt_1dt_2 =$$

$$= \frac{1}{D_{n,1}D_{n,2}} \iint\limits_{\tilde{O}_n} |f_n^*(t_1, t_2) - e^{-\frac{1}{2}q(t_1,\, t_2)}|\,dt_1dt_2,$$

where

$$\tilde{O}_n = \left\{(t_1, t_2): |t_1| \leq \frac{\sqrt{n}}{4},\quad |t_2| \leq \frac{\sqrt{n}}{2\sqrt{12}}\right\}.$$

Using higher moments we expand the function and estimate the term of error. There is no restriction for our method but the convergency to the function, and this condition is given by the definition (8). The new quantities are:

$$\delta_{k,i}^2 = E\{(\xi_{k,i} - E(\xi_{k,i}))^2\} \quad i = 1, 2,$$

$$\beta_{k,i}^4 = E\{(\xi_{k,i} - E(\xi_{ki,}))^4\} \quad i = 1, 2,$$

$$E(\xi_{k,1}) = \frac{1}{p+1}\sum_{x=0}^{P} x = \frac{p}{2},$$

$$E(\xi_{k,2}) = \frac{1}{p+1}\sum_{x=0}^{P} kx = \frac{p}{2}k,$$

$$E(\xi_{k,1}^2) = \frac{1}{p+1}\sum_{x=0}^{P} x^2 = \frac{p(2d+1)}{6},$$

$$E(\xi_{k,2}^2) = \frac{1}{p+1}\sum_{x=0}^{P} k^2x^2 = \frac{p(2pH)}{6}k^2,$$

$$E(\xi_{k,1}^3) = \frac{1}{p+1}\sum_{x=0}^{P} x^3 = \frac{p^2(p+1)}{4},$$

$$E(\xi_{k,2}^3) = \frac{1}{p+1}\sum_{x=0}^{P} k^3x^3 = \frac{p^2(p+1)}{4}k^3,$$

$$E(\xi_{k,1}^4) = \frac{1}{p+1}\sum_{x=0}^{P} x^4 = \frac{p(2p+1)(3p^2+3p-1)}{30},$$

$$E(\xi_{k,2}^4) = \frac{1}{p+1}\sum_{x=0}^{P} k^4x^4 = \frac{p(2p+1)(3p^2+3p-1)}{30}k^4,$$

$$\delta_{k,1}^2 = E(\xi_{k,1}^2) - E^2(\xi_{k,1}) = \frac{p(p+2)}{12},$$

$$\delta_{k,2}^2 = E(\xi_{k,2}^2) - E^2(\xi_{k,2}) = \frac{p(p+2)}{12}k^2,$$

$$\beta_{4,1}^k = \frac{p(p+2)(3p^2+6p-4)}{240},$$

$$\beta_{k,2}^4 = \frac{p(p+2)(3p^2+6p-4)}{240}k^4,$$

$$\frac{\beta_{k,i}^4}{\delta_{k,i}^4} = \frac{3}{5}\frac{3p^2+6p-4}{p(p+2)} = \frac{9}{5}\left(1-\frac{4}{3}\frac{1}{p(p+2)}\right),$$

and by $p \geq 1$,

$$\delta_{k,i}^4 \leq \beta_{k,i}^4 \leq \frac{9}{5}\delta_{k,i}^4.$$

We have also:

$$\delta_{k,1,2,} = E\{(\xi_{k,1}-E(\xi_{k,1}))(\xi_{k,2}-E(\xi_{k,2}))\} =$$

$$= \frac{1}{p+1}\sum_{x=0}^{P}\left(x-\frac{p}{2}\right)\left(kx-k\frac{p}{2}\right) = \frac{p(p+2)}{12}k,$$

$$A_{n,1} = \sum_{k=1}^{n} E(\xi_{k,1}) = \frac{np}{2},$$

$$A_{n,2} = \sum_{k=1}^{n} E(\xi_{k,2}) = \frac{n(n+1)p}{4},$$

$$D_{n,2}^2 = \sum_{k=1}^{n} \delta_{k,1}^2 = \frac{np(p+2)}{12},$$

$$D_{n,1}^2 = \sum_{k=1}^{n} \delta_{k,2}^2 = \frac{n(n+1)(2n+1)p(p+2)}{72},$$

$$D_{n,1,2} = \sum_{k=1}^{n} \delta_{k,1,2} = \frac{n(n+1)p(p+2)}{24},$$

$$h_n = \frac{D_{n,1,2}}{D_{n,1}D_{n,2}} = \sqrt{\frac{3n+3}{4n+2}},$$

$$\sum_{k=1}^{n} \beta_{k,1}^4 = \frac{np(P+2)(3p^2+6p-4)}{240},$$

$$\sum_{k=1}^{n} \beta_{k,2}^4 = \frac{n(n+1)(2n+1)(3n^2+3n-1)p(p+2)(3p^2+6p-4)}{7200}.$$

Now we have

$$\sum_{k=1}^{n} \frac{\beta_{k,1}^4}{D_{4,1}^4} \leq \frac{9}{5n}, \quad \sum_{k=1}^{n} \frac{\beta_{k,2}^4}{D_{n,2}^4} \leq \frac{81}{25n}.$$

Let us consider the characteristic function $\varphi_k(t_1, t_2)$ of the random vector ξ_k and expand it:

$$\varphi_k(t_1, t_2) = \frac{1}{p+1} \sum_{x=0}^{P} e^{ix(t_1+kt_2)} = 1 + iE(\xi_{k,1})(t_1+kt_2) -$$

$$-\frac{1}{2} E(\xi_{k,1}^2)(t_1+kt_2)^2 - \frac{i}{6} E(\xi_{k,1}^3)(t_1+kt_2)^3 + \frac{1}{24} E(\xi_{k,1}^4)(t_1+tk_2)^4 + \dots,$$

and let us define the functions

$$\varphi_k^*(t_1, t_2) = e^{-iE(\xi_{k,1})\left(\frac{t_1}{D_{n,1}} + k\frac{t_2}{D_{n,2}}\right)} \varphi_k\left(\frac{t_1}{D_{n,1}}, \frac{t_2}{D_{n,2}}\right). \tag{11}$$

It is essential, that φ_k^* is real-valued and their product is f_n^*.

On the other hand we have

$$e^{-iE(\xi_{k,1})\left(\frac{t_1}{D_{n,1}} + k\frac{t_2}{D_{n,2}}\right)} 1 - iE(\xi_{k,1})\left(\frac{t_1}{D_{n,1}} + k\frac{t_2}{D_{n,2}}\right) -$$

$$-\frac{1}{2} E^2(\xi_{k,1})\left(\frac{t_1}{D_{n,1}} + k\frac{t_2}{D_{n,2}}\right)^2 + \frac{i}{6} E^3(\xi_{k,1})\left(\frac{t_1}{D_{n,1}} + k\frac{t_2}{D_{n,2}}\right)^3 +$$

$$+\frac{1}{24} E^4(\xi_{k,1})\left(\frac{t_1}{D_{n,1}} + k\frac{t_2}{D_{n,2}}\right)^4 + \dots$$

and multiplying the expansion, in (11), we have by Taylor's formula with a $0 < \Theta < 1$

$$\varphi_k^*(t_1, t_2) = 1 - \frac{1}{2}\left(\frac{\delta_{k,1}}{D_{n,1}} t_1 + \frac{\delta_{k,2}}{D_{n,2}} t_2\right)^2 + \frac{1}{24}\left(\frac{\beta_{k,1}}{D_{n,1}} \Theta t, + \frac{\beta_{k,2}}{D_{n,2}} \Theta t_2\right)^4.$$

If $(t_1, t_2) \in \tilde{Q}_n$, then

$$|1 - \varphi_k^*(t_1, t_2)| \leq \frac{1}{2}\left|\frac{\delta_{k,1}}{D_{n,1}} t_1 + \frac{\delta_{k,2}}{D_{n,2}} t_2\right|^2 +$$

$$+\frac{1}{24}\left|\frac{\beta_{k,1}}{D_{n,1}} |t_1| + \frac{\beta_{k,2}}{D_{n,2}} |t_2|\right|^4 \leq \frac{1}{2} + \frac{1}{24} < 1,$$

and the possibility is given to expand $\log(1-\varphi_k^*)$ on $\tilde{Q}_n$. Using $(a+b)^2 \leq \leq 2(a^2+b^2)$ we notice

$$|1-\varphi_k^*)|^2 \leq \frac{1}{2}\left(\frac{\sigma_{k,1}}{D_{n,1}}|t_1| + \frac{\sigma_{k,2}}{D_{n,2}}|t_2|\right)^4 + \frac{1}{288}\left(\frac{\beta_{k,1}}{D_{n,1}}|t_1| + \frac{\beta_{k,2}}{D_{n,2}}|t_2|\right)^8 \leq$$

$$\leq \left(\frac{1}{2}+\frac{1}{288}\right)\left(\frac{\beta_{k,1}}{D_{n,1}}|t_1| + \frac{\beta_{k,2}}{D_{n,2}}|t_2|\right)^4 = \frac{145}{288}\left(\frac{\beta_{k,1}}{D_{n,1}}|t_1| + \frac{\beta_{k,2}}{D_{n,2}}|t_2|\right)^4.$$

This way we have with $0<\Theta,\ \Theta_k<1$

$$\log f_n^*(t_1, t_2) = \sum_{k=1}^{n} \log\left[1-(1-\varphi_k^*(t_1, t_2))\right] =$$

$$= \sum_{k=1}^{n}\left\{(\varphi_k^*(t_1, t_2)-1) - \frac{\Theta_k}{2}(\varphi_k^*(t_1, t_2)-1)^2\right\} =$$

$$= -\frac{1}{2}\sum_{k=1}^{n}\left(\frac{\sigma_{k,1}}{D_{n,1}}t_1 + \frac{\sigma_{k,2}}{D_{n,2}}t_2\right)^2 - \frac{121}{576}\sum_{k=1}^{n}\left(\frac{\beta_{k,1}}{D_{n,1}}\Theta|t_1| + \frac{\beta_{k,2}}{D_{n,2}}\Theta|t_2|\right)^4.$$

Having applied

$$\sum_{k=1}^{n}\left(\frac{\sigma_{k,1}}{D_{n,1}}t_1 + \frac{\sigma_{k,2}}{D_{n,2}}t_2\right)^2 = t_1^2 + 2h_n t_1 t_2 + t_2^2,$$

we have for $(t_1, t_2)\in\tilde{Q}_n$

$$\left|f_n^*(t_1, t_2) - e^{-\frac{1}{2}q(t_1, t_2)}\right| = e^{-\frac{1}{2}q(t_1, t_2)} \cdot \left|e^{-\frac{121}{576}\sum_{k=1}^{n}\left(\frac{\beta_{k,1}}{D_{n,1}}\Theta|t_1| + \frac{\beta_{k,2}}{D_{n,2}}\Theta|t_2|\right)^4} - 1\right| \leq$$

$$\leq e^{-\frac{1}{2}q(t_1, t_2)}\frac{121}{72}\left(\frac{\sum_{k=1}^{n}\beta_{k,1}^4}{D_{n,1}^4}t_1^4 + \frac{\sum_{k=1}^{n}\beta_{k,2}^4}{D_{n,2}^4}t_2^4\right). \tag{12}$$

$$\cdot\exp\left[-\frac{121}{72}\left(\frac{\sum_{k=1}^{n}\beta_{k,1}^4}{D_{n,1}^4}t_1^4 + \frac{\sum_{k=1}^{n}\beta_{k,2}^4}{D_{n,2}^4}t_2^4\right)\right] \leq e^{-\frac{1}{2}q(t_1, t_2)}\frac{121}{40n}\left(t_1^4 + \frac{9}{5}t_2^4\right).$$

Comparing (10) and (12) we have

$$|I_1| \leq \frac{121}{40nD_{n,1}D_{n,2}}\iint\limits_{\tilde{O}_n}\left(t_1^4 + \frac{9}{5}t_2^4\right)e^{-\frac{1}{2}q(t_1, t_2)}dt_1\,dt_2 \leq \frac{6{,}5\cdot 10^5}{nD_{n,1}D_{n,2}}.$$

The latter estimate is given by $h_n \leq \sqrt{\frac{6}{7}}$ for $n\geq 3$, and $t_1^2+2h_n t_1 t_2 + t_2^2 \geq \geq \left(1-\sqrt{\frac{6}{7}}\right)(t_1^2+t_2^2)$. We have $t_1^2 + \frac{9}{5}t_2^2 \leq \frac{9}{5}(t_1^2+t_2^2)^2$, and having an rota-

tion-invariant majorant, we substitute polar coordinates and in the later integral of one variable we integrate partial twice. It is worth noting $I_2 = o\left(\frac{1}{nD_{n,1}D_{n,2}}\right)$, $I_3 = o\left(\frac{1}{nD_{n,1}D_{n,2}}\right)$, but we must not ignore them, since we are interested also in the values of constants.

4. The estimation of I_2 and I_3

We begin it with I_2.

$$|I_2| \leq \iint\limits_{[-\pi,\pi]^2-O_n} \prod_{k=1}^{n} \left| \frac{\sin (p+1) \frac{t_1+kt_2}{2}}{(p+1) \sin \frac{t_1+kt_2}{2}} \right| dt_1 dt_2 =$$

$$= 4 \iint\limits_{\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]^2-\frac{1}{2}O_n} \prod_{k=1}^{n} \left| \frac{\sin (p+1)(t_1+kt_2)}{(p+1) \sin (t_1+kt_2)} \right| dt_1 dt_2.$$

The essence of the proof is $\left| \frac{\sin (p+1)x}{(p+1) \sin x} \right| \leq 1$, and being x from the nearest $m\pi$ at a great distance apart, it is much lower than 1. This way we give an uniform estimate for the integrand not only out of $\frac{1}{2} Q_n$, but out of the less domain

$$H = \left\{(x, y): |x| \leq \frac{\sqrt{3}}{8(p+1)}, |y| \leq \frac{\sqrt{3}}{4(n+1)(p+1)}\right\},$$

Indeed, if $\frac{\sqrt{3}}{8(p+1)} \leq |x| \leq \frac{\pi}{2}$, then $x+ky$ $(k = 1, 2, \ldots, n)$ is at least $n/2$ times at least at distance $\frac{\sqrt{3}}{16(p+1)}$ from the nearest $m\pi$. If $|x| < \frac{\sqrt{3}}{8(p+1)}$, then $\pi/2 \geq |y| \geq \frac{\sqrt{3}}{4(n+1)(p+1)}$, and $x+ky$ is at least $n/2$ times at least at $\frac{\sqrt{3}}{16(p+1)}$ distance apart from the nearest $m\pi$. If $(x, y) \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]^2 - H$, then for all $p \geq 1$

$$\prod_{k=1}^{n} \left| \frac{\sin (p+1)(x+ky)}{(p+1) \sin (x+ky)} \right| \leq \left| \frac{\sin (p+1) \frac{\sqrt{3}}{16(p+1)}}{(p+1) \sin \frac{\sqrt{3}}{16(p+1)}} \right|^{n/2} \leq$$

$$\leq \left(\frac{\sin \sqrt{3}/16}{\frac{\sqrt{3}}{16} - \left(\frac{\sqrt{3}}{16}\right)^3 \frac{1}{6} \frac{1}{(p+1)^2}} \right)^{n/2} \leq 0{,}99855^{n/2}.$$

It follows

$$|I_2| \leq 4\pi^2\, 0{,}99855^{n/2} < \frac{100}{nD_{n,1} D_{n,2}}, \text{ since } p \leq (1{,}0014)^{n/4} n^{-3/2}.$$

We can estimate in a similar way as we made the last estimate in the previous section:

$$e^{-\frac{1}{2} q(D_{n,1} t_1,\, D_{n,2} t_2)} \leq e^{-\frac{1}{2}\left(1-\sqrt{\frac{6}{7}}\right)\left(D_{n,1}^2 t_1^2 + D_{n,2}^2 t_2^2\right)},$$

and

$$|I_3| \leq \frac{1}{D_{n,1} D_{n,2}} \iint\limits_{\mathbf{R}-\tilde{O}_n} e^{-\frac{1}{2}\left(1-\sqrt{\frac{6}{7}}\right)(\tau_1^2+\tau_2^2)}\, d\tau_1 d\tau_2 \leq$$

$$\leq \frac{2\pi}{D_{n,1} D_{n,2}} \int\limits_{\frac{\sqrt{n}}{2\sqrt{12}}}^{\infty} r e^{-\frac{1}{2}\left(1-\sqrt{\frac{6}{7}}\right) r^2} dr \leq \frac{1{,}1 \cdot 10^6}{nD_{n,1} D_{n,2}}.$$

5. Application to Wilcoxon statistics

Suppose f and g are continuous distribution functions, $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_t$ are independent random variables with these distribution functions. The event, that two of the $m+t$ variables have the same value, is of probability zero, so, this event will be out of regard. Let us consider the increasing sequence of $m+t$ numbers of a sample and denote by R_k the number of position (among the $m+t$ numbers) of the increasing subsequence of X-es. Several order statistics try to decide using the positions of X-es whether $f = g$ or not.

In the Wilcoxon statistics $W = \sum_{k=1}^{m} R_k$ is considered, see the book of SCHMETTERER [11]. From this point we suppose $f = g$. MANN and WHITNEY [8] proved the Wilcoxon statistics to be asymptotically normal.

ISMATULLAEFF and JOÓ [5] noticed the following, quite general connection of order statistics and diophantic systems.

Let us denote by $X_{(k)}$ the kth element of the increasing sequence of X-es, $X_{(0)} =: -\infty$, $X_{(m+1)} =: +\infty$, by S_k the number of Y-s in the interval $(X_{(k-1)}, X_{(k)})(k = 1, 2, \ldots, m+1)$. Let us consider the statistics

$$T_{m,t} = \sum_{k=1}^{m+1} a(k, S_k),$$

where $a(i, j)$ is an integer-valued function. The notice of the authors is

$$P(T_{m,t} = K) = \tau_{m,t}(K) \Big/ \binom{m+t}{t},$$

where $\tau_{m,t}(K)$ is the number of solutions of the following diophantic system:

$$\text{(i')} \qquad \sum_{k=1}^{m+1} a(k, x_k) = K,$$

$$\text{(ii')} \qquad \sum_{k=1}^{m+1} x_k = t,$$

$$\text{(iii')} \qquad 0 \leq x_1, \ldots, x_{m+1} \leq t.$$

(It is easy to check by $x_1 = R_1 - 1$, $x_{m+1} = t + m - R_m$, $x_k = R_k - R_{k-1} - 1$, $k = 2, 3, \ldots, m$). The authors applied it to the case of Dixon statistics $a(k, x_k) = x_k^2$. Connecting (i'), (ii'), (iii') to (i), (ii), (iii) of the first section we have information on the distribution of Wilcoxon statistics through Theorem 1. We have

$$W = \sum_{k=1}^{m} R_k = \binom{m+1}{2} + \sum_{k=1}^{m+1} (m-k+1)(R_k - R_{k-1} - 1),$$

where $R_0 = 0$, $R_{m+1} = n + m - 1$ and introducting the function

$$a(k, x_k) = (m-k+1)\, x_k,$$

we have $W - \binom{m+1}{2} = T_{m,t}$.

Denoting by $P(A|B)$ the conditional probability, we have

$$P(W = K \,|\, S_{m+1} = s) = P\left(T_{m,t} = K - \binom{m+1}{2} \,\middle|\, S_{m+1} = s\right) =$$

$$= \frac{\tau_{m,t}\left(K - \binom{m+1}{2}\right)}{\binom{m+t}{t}} : \frac{\binom{m+t-s-1}{m-1}}{\binom{m+t}{t}}$$

where $\tau_{m,t}$ is the number of solutions of the diophantic system

$$\text{(i'')} \qquad \sum_{k=1}^{m} k x_k = K - \binom{m+1}{2},$$

$$\text{(ii'')} \qquad \sum_{k=1}^{m} x_k = t - s,$$

$$\text{(iii'')} \qquad 0 \leq x_1, \ldots, x_m \leq t - s.$$

This way we have the following theorem.

THEOREM 2. *Suppose* $1 \le t-s \le (1{,}0014)^{m/4} m^{-3/2}$, $3 \le m$, *then*

$$P(W = K | S_{m+1} = s) = \frac{(t-s+1)^m}{\binom{m+t-s-1}{m-1} D_{m,1} D_{m,2}} \cdot$$

$$\cdot \left[\frac{e^{-\frac{1}{2} Q\left(\frac{t-s-A_{m,1}}{D_{m,1}}, \frac{K-\binom{m+1}{2}-A_{m,2}}{D_{m,2}} \right)}}{2\pi\sqrt{1-h_m^2}} + \frac{\Theta}{m} \right]$$

where

$$h_m = \sqrt{\frac{3m+3}{4m+2}}, \quad |\Theta| \le 10^5,$$

$$Q(u_1, u_2) = \frac{1}{1-h_m^2} (u_1^2 - 2h_m u_1 u_2 + u_2^2),$$

$$A_{m,1} = \frac{1}{2} m(t-s), \; A_{m,2} = \frac{1}{4} m(m+1)(t-s),$$

$$D_{m,1}^2 = \frac{1}{12} m(t-s)(t-s+2),$$

$$D_{m,2}^2 = \frac{1}{72} m(m+1)(2m+1)(t-s)(t-s+2).$$

References

[1] CASTELNOUVO, G., Sur quelques problemes se rattachant au calcul de probabilitiés, *Ann. l'Inst. H. Poincaré*, **3** (1933), 465–490.

[2] ERDŐS, P., On an elementary proof of some asymptotic formulas in the theory of partitions, *Annals of Math.*, **43** (1942), 437–450.

[3] FELLER, W., *An introduction to probability theory and its applications*, Wiley-Chapman, New York–London, 1950.

[4] HARDY, G. H. and RAMANUJAN, S., Asymptotic formula in combinatorial analysis, *Proc. London Math. Soc.* (2), **17** (1918), 74–115.

[5] ISMATULLAEV, S. A. and Joó, I., Asymptotical analysis of Dixon's statistics, to appear in *Acta Math. Sci. Hung.*, 1985 (in Russian).

[6] Joó, I., On the number of partitions of the number N into terms of 1, 2, ..., n repeating a term at most p times, *Annales Univ. Sci. Budapest*, Sect. Math., **28** (1985), 217–227.

[7] KAHANE, J. P., *Some random series of functions*, D. C. Heath and Co., Lexington, Massachusetts, 1968.

[8] MANN, H. B. and WHITNEY, D. R., ON a test of whether one of two random variables is stochastically larger than the other, *Ann. Math. Stat.*, **18** (1947), 50–60.

[9] POSTNIKOFF, A. G., *Introduction in analytic number theory*, Moscow. 1971 (in Russian).

[10] POSTNIKOFF, A. G., Additive problems in number theory with increasing number of terms, *Izv. AN. SSSR*, Series Math. **20** (1956) (in Russian).

[11] SCHMETTERER, L., *Einführung in die mathematische Statistik* (2. Auflage), Springer–Verlag, Wien–New York, 1966.

[12] SIRASHDINOFF, S. CH., AZLAROFF, T. A. and ZUPAROFF, T. M., Problems in additive theory of increasing number of terms, *Fan*, Tashkent, 1975 (in Russian).

ON THE NUMBER OF PARTITIONS OF THE NUMBER N INTO TERMS OF $1, 2, \ldots, n$ REPEATING A TERM AT MOST p TIMES

By

I. JOÓ

Institute for Mathematics of the L. Eötvös University, Budapest

(Received, November 30, 1983)

Dedicated to Professor Á. CSÁSZÁR on the occasion of his 60th birthday

1. Introduction

A great number of papers investigate asymptotic formulae for the number of solutions of Diophantine equations (Cf. [1]–[3] and references there), even a prominent monography has been written in this topic by S. CH. SIRASHDINOFF, T. A. AZLAROFF and T. M. ZUPAROFF [4]. The number of solutions of Diophantine equations in connection with the number of partitions is of particular interest. Nevertheless, the problem mentioned in the title of our present paper seems to be new in the literature.

Let us consider the following equation

$$1 \cdot x_1 + 2x_2 + \ldots + nx_n = N, \quad 0 \leq x_i \leq p \quad \text{integer.} \tag{1}$$

Let us denote by R the number of solutions of (1).

G. H. HARDY and S. RAMANUJAN in 1918 [2] proved

$$R = \left(1 + o\left(\frac{1}{\sqrt{n}}\right)\right) \frac{1}{4\sqrt{3}n} \exp\left(\frac{2\pi}{\sqrt{6}}\sqrt{n}\right) \quad \text{if } n = N = p$$

and

$$R = \frac{1 + o(1)}{4 \cdot 3^{1/4} \cdot n^{3/4}} \exp\left(\frac{\pi}{\sqrt{3}}\sqrt{n}\right) \quad \text{if } n = N,\ p = 1.$$

The aim of our paper is to prove the following theorem using probabilistic methods.

THEOREM. *Suppose $n \geq 2$, $p \geq 1$ and N are natural numbers. Let us introduce the notations*

$$A_n = \frac{1}{4} n(n+1)p \text{ and } D_n^2 = \frac{1}{72} n(n+1)(2n+1)p(p+2).$$

Then the number of solutions of (1) *is*

$$R = (p+1)^n \frac{1}{\sqrt{2\pi}D_n}\left\{\exp\left[-\frac{(N-A_n)^2}{2D_n^2}\right]+\frac{\Theta}{n}\right\}$$

where $|\Theta| \leq K$ *with an universal constant* K.

REMARK. The ideas of our proof were further developed in [5] for the investigation of Diophantine system of equations. The proof of the Theorem will be given in several sections.

2. Proof of the Theorem

Let us consider the independent random variables ξ_k $(k = 1, \ldots, n)$ with the following distributions:

$$P(\xi_k = k\cdot x) = \frac{1}{p+1} \quad (x = 0, 1, \ldots, p).$$

Let us define S_n to be

$$S_n = \sum_{k=1}^{n} \xi_k.$$

It is easy to check that

$$E(S_n) = A_n \text{ and } D^2(S_n) = D_n^2.$$

Let us consider the $f_n(t)$ characteristic function of S_n and the $f_n^*(t)$ characteristic function of $(S_n - A_n)/D_n$. We have

$$(2) \qquad f_n(t) = E(e^{iS_nt}) = \prod_{k=1}^{n} \frac{1}{p+1}\sum_{x=1}^{p} e^{ixkt} = \frac{e^{iA_nt}}{(p+1)^n}\prod_{k=1}^{n}\frac{\sin(p+1)\frac{kt}{2}}{\sin\frac{kt}{2}}$$

and

$$(3) \qquad f_n^*(t) = E\left(e^{i\frac{S_n-A_n}{D_n}t}\right) = e^{-i\frac{A_n}{D_n}t} f_n\left(\frac{t}{D_n}\right).$$

In fact we are going to prove a local limit distribution theorem, what is a standard tool in this field (See [4]). However, the realization of the proof requires also new ideas in our contribution. We have

$$\int_{-\pi}^{\pi} f_n(t)e^{-itN}dt = \frac{1}{(p+1)^n}\int_{-\pi}^{\pi}\sum_{x_1,\ldots,x_n=0}^{p} e^{it(1\cdot x_1+2x_2+\ldots+nx_n-N)}dt =$$

$$(4) \qquad = \frac{2\pi}{(p+1)^n}R = 2\pi P(S_n = N)$$

and the well known inversion formula

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-iut}e^{-\frac{t^2}{t}}dt = \frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}. \tag{5}$$

We have to prove

$$\left|\sqrt{2\pi}\,D_n\frac{R}{(p+1)^n} - e^{-\frac{(N-A_n)}{2D_n^2}}\right| = O\left(\frac{1}{n}\right), \tag{6}$$

since (6) is equivalent to the Theorem.

Using (2), (3), (4) and (5) we can write

$$D_nP(S_n=N) - \frac{1}{\sqrt{2\pi}}e^{-\frac{(N-A_n)}{2D_n^2}} =$$

$$= \frac{D_n}{2\pi}\int_{-\pi}^{\pi} f_n(t)e^{-itN}dt - \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{t^2}{t}}e^{-i\frac{N-A_n}{D_n}t}dt =$$

$$= \frac{D_n}{2\pi}\left(\int_{-\pi}^{\pi} f_n^*(D_nt)\,e^{-i(N-A_n)}dt - \int_{-\infty}^{+\infty} e^{-\frac{D_n^2t^2}{2}}e^{-i(N-A_n)t}dt\right)$$

and

$$D_nP(S_n=N) - \frac{1}{\sqrt{2\pi}}e^{-\frac{(N-A_n)^2}{2D_n^2}} = \frac{D_n}{2\pi}(I_1+I_2-I_3), \tag{7}$$

where

$$I_1 = \int_{|t|\leq\frac{c}{np}} e^{-i(N-A_n)t}[f_n^*(D_nt) - e^{-D_n^2t^2/2}]\,dt,$$

$$I_2 = \int_{\frac{c}{np}\leq|t|\leq\pi} e^{-i(N-A_n)t}f_n^*(D_nt)dt,$$

$$I_3 = \int_{|t|\geq\frac{c}{np}} e^{-i(N-A_n)t}\cdot e^{-D_n^2t^2/2}\,dt.$$

We shall prove

$$|I_1| \leq \frac{K_1}{nD_n},\quad |I_2| \leq \frac{K_2}{p}e^{-cn},\quad |I_3| \leq \frac{K_3}{p}e^{-cn},$$

what proves

$$D_n(|I_1|+|I_2|+|I_3|) = O\left(\frac{1}{n}\right)$$

and having a look at (7) we get (6).

3. The estimation of I_1

Let us consider the following quantities:

$$\sigma_k^2 = E[(\xi_k - E(\xi_k))^2],$$

$$\beta_k^4 = E[(\xi_k - E(\xi_k))^4],$$

$$E(\xi_k^2 = \frac{1}{p+1}\sum_{x=0}^{p} kx = \frac{p}{2}k,$$

$$E(\xi_k^2) = \frac{1}{p+1}\sum_{x=0}^{p} k^2x^2 = \frac{1}{6}p(2p+1)k^2,$$

$$E(\xi_k^3) = \frac{1}{p+1}\sum_{x=0}^{p} k^3x^3 = \frac{1}{4}p^2(p+1)k^3,$$

$$E(\xi_k^4) = \frac{1}{p+1}\sum_{x=0}^{p} k^4x^4 = \frac{1}{30}p(2p+1)(3p^2+3p-1)k^4,$$

$$\sigma_k^2 = E(\xi_k^2) - E^2(\xi_k) = \frac{1}{12}p(p+2)k^2,$$

$$\beta_k^4 = \frac{1}{240}p(p+2)(3p^2+6p-4)k^4,$$

$$\frac{\beta_k^4}{\sigma_k^4} = \frac{3}{5}\cdot\frac{3p^2+6p-4}{p(p+2)} = \frac{9}{5}\left(1-\frac{4}{3}\cdot\frac{1}{p(p+2)}\right).$$

We have

(8)
$$1 \le \frac{\beta_k^4}{\sigma_k^4} \le \frac{9}{5},$$

$$\sum_{k=1}^{n} E(\xi_k) = A_n = \frac{1}{4}n(n+1)p,$$

$$\sum_{k=1}^{n} \sigma_k^2 = D_n^2 = \frac{1}{72}n(n+1)(2n+1)p(p+2),$$

$$B_n^4 \overset{\text{def}}{=} \sum_{k=1}^{n} \beta_n^4 = \frac{1}{7200}n(n+1)(2n+1)(3n^2-3n-1)p(p+2)(3p^2+6p-4),$$

$$\frac{17}{20n} \le \frac{B_n^4}{D_n^4} \le \frac{81}{25n}.$$

Let us consider the characteristic function $\varphi_k(t)$ of ξ_k and

$$\varphi_k^*(t) = e^{-iE(\xi_k)t/D_n} \cdot \varphi_k\left(\frac{t}{D_n}\right) \text{ of } \frac{\xi_k - E(\xi_k)}{D_n}.$$

Expanding them we have

$$\varphi_k(t) = \frac{1}{p+1} \sum_{x=0}^{p} e^{ixkt} = 1 + iE(\xi_k)t - \frac{1}{2} E(\xi_k^2)t^2 -$$

$$-\frac{i}{6} E(\xi_k^3)t^3 + \frac{1}{24} E(\xi_k^4)t^4 - \ldots,$$

$$e^{-iE(\xi_k)\frac{t}{D_n}} = 1 - iE(\xi_k)\frac{t}{D_n} - \frac{1}{2} E^2(\xi_k)\frac{t^2}{D_n^2} + \frac{i}{6} E^3(\xi_k)\frac{t^3}{D_n^3} +$$

$$+\frac{1}{24} E^4(\xi_k)\frac{t^4}{D_n^4} - \ldots,$$

$$\varphi_k^*(t) = 1 - \frac{1}{2}\frac{\sigma_k^2}{D_n^2} t^2 + \frac{\Theta}{24}\frac{\beta_k^4}{D_n^4} t^4, \quad (0 < \Theta < 1).$$

If $|t| \leq c\sqrt{n}$ with a $c \leq 3^{-3/4}$, then

$$|1 - \varphi_k^*(t)| \leq \frac{1}{2}\frac{\sigma_k^2}{D_n^2} t^2 + \frac{1}{24}\frac{\beta_k^2}{D_n^2} t^4 \leq \frac{1}{2} + \frac{1}{24} < 1. \tag{9}$$

Using $(a+b)^2 \leq 2(a^2+b^2)$ and (9) we get

$$|1 - \varphi_k^*(t)|^2 \leq \frac{1}{2}\frac{\sigma_k^4}{D_n^4} t^4 + \frac{1}{288}\frac{\beta_k^8}{D_n^8} t^8 \leq$$

$$\leq \left(\frac{1}{2} + \frac{1}{288}\right)\frac{\beta_k^4}{D_n^4} t^4 \leq \frac{145}{288}\frac{\beta_k^4}{D_n^4} t^4. \tag{10}$$

Since $f_n^*(t) = \prod_{k=1}^{n} \varphi_k^*(t)$ is real and (10), we have

$$\log f_n^*(t) = \sum_{k=1}^{n} \log\left[1 - (1 - \varphi_k^*(t))\right] =$$

$$= \sum_{k=1}^{n} \left\{(\varphi_k^*(t) - 1) - \frac{\Theta}{2}(\varphi_k^*(t) - 1)^2\right\} = -\frac{1}{2}\sum_{k=1}^{n} \frac{\sigma_k}{D_n^2} t^2 -$$

$$-\frac{\Theta}{2}\frac{157}{288}\sum_{k=1}^{n} \frac{\beta_k^4}{D_n^4} = -\frac{t^2}{2} - \Theta\frac{157}{576}\frac{B_n^4}{D_n^4} t^4. \tag{11}$$

From (11) and (8) follows

$$|f_n^*(t)-e^{-t^2/2}|\le e^{-t^2/2}\left|e^{-\Theta\frac{157}{576}\frac{B_n^2}{D_n^2}t^4}-1\right|\le$$

$$\le e^{-t^2/2}\frac{157}{576}\frac{B_n^4}{D_n^4}t^4e^{-\Theta\frac{157}{576}\frac{B_n^4}{D_n^4}t^4}\le\frac{K_1}{n}t^4e^{-t^2/2}$$

and

$$|I_1|\le\int\limits_{|t|\le\frac{c}{np}}|f_n^*(D_nt)-e^{-D_n^2t^2/2}|dt=$$

$$=\frac{1}{D_n}\int\limits_{|t|\le c\sqrt{n}}|f_n^*(t)-e^{-t^2/2}|dt\le\frac{K_1}{D_nn}\int\limits_{|t|\le c\sqrt{n}}t^4e^{-t^2/2}dt\le K_1/D_nn.$$

4. The estimation of I_2

We need a number of lemmata to estimate the characteristic function $f_n(t)$.

LEMMA 1. We have for $p=0,1,\ldots$

$$\left|\frac{\sin(p+1)x}{\sin x}\right|\le 1+|\cos x|+\ldots+|\cos x|^p, \tag{12}$$

$$\left|\frac{\sin(p+1)x}{(p+1)\sin x}\right|\le 1. \tag{13}$$

PROOF. Use induction on p:

$$\left|\frac{\sin(px+x)}{\sin x}\right|\le\left|\frac{\sin px}{\sin x}\right|\cdot|\cos x|+|\cos px|\le 1+$$

$$+|\cos x|(1+|\cos x|+\ldots+|\cos x|^{p-1}).$$

LEMMA 2. Let us define m and M as follows:

$$m=\inf_{1\le t\le\infty}\frac{\left(\frac{t}{e}\right)^t\sqrt{2\pi t}}{\Gamma(t+1)},\quad M=\sup_{1\le t\le\infty}\frac{\left(\frac{t}{e}\right)^t\sqrt{2\pi t}}{\Gamma(t+1)}.$$

Denote

$$A=\left[\frac{3\pi}{n(p+1)},\frac{\pi}{p(p+1)}\right]\quad\text{and}\quad g(x)=\left|\prod_{k=1}^{n}\frac{\sin(p+1)kx}{(p+1)\sin kx}\right|.$$

Then

$$\int_A g(x)\,dx \le C\frac{M}{m}\frac{1}{n(n-1)(p+1)}\left(\frac{e}{6}\right)^n,\ (n=2,3,\ldots). \tag{14}$$

Here and in below C denotes a constant not depending on n, N, p, and it may be different in different places.

PROOF. Using (13) we have

$$g(x) \ge \prod_{\substack{k:\\ \sin kx \ge \frac{1}{p+1}}} \frac{1}{(p+1)\sin kx} \ge \prod_{\substack{k:\\ \frac{2}{\pi}kx \ge \frac{1}{p+1}}} \frac{1}{(p+1)\frac{2}{\pi}kx}$$

and

$$g(x) \le \prod_{\substack{k:\\ \frac{2(p+1)x}{\pi} \le k \le n}} \frac{\pi}{2(p+1)\,kx} \le \left(\frac{\pi}{p(p+1)x}\right)^{n-\frac{\pi}{2(p+1)x}+1}\cdot \tag{15}$$

$$\cdot\frac{\Gamma\left(\frac{\pi}{2(p+1)x}+1\right)}{n!} \le \frac{\pi}{2(p+1)x}\left(\frac{\pi e}{2n(p+1)x}\right)^n\cdot$$

$$\cdot\frac{\frac{M}{m}\frac{\pi}{\sqrt{(p+1)x}}}{e^{\frac{\pi}{2(p+1)x}}\sqrt{2n}} \le c\left(\frac{\pi e}{2n(p+1)x}\right)^{n+\frac{1}{2}}$$

The right hand side of (14) is the integral of (15).

LEMMA 3. Denote $B=\left[\frac{\pi}{2(p+1)},\frac{\pi}{2n}\right]$. Then

$$\int_B g(x)\,dx \le C\frac{1}{(p+1)(n-1)\,n!},\ (n=2,3,\ldots).$$

PROOF. As in (15),

$$g(x) \le \prod_{\substack{k:\\ \frac{\pi}{2(p+1)x} \le k \le n}} \frac{\pi}{2(p+1)kx}$$

and

$$\int_B g(x)\,dx \le \int_B \left(\frac{\pi}{2(p+1)x}\right)^n \frac{1}{n!}\,dx \le \frac{\pi}{2(p+1)(n-1)\,n!}.$$

LEMMA 4. Assume $n \geq 4$ and $1/2 \leq x \leq 1/8$. Then there exists at least $\max(n/8, 1)$ number of the sequence $x, 2x, \ldots, nx$ being at least at distance $1/4$ apart from the nearest integer.

PROOF. At least $[n\,x]+1$ of the intervals $\left[0, \frac{1}{2}\right], \left[\frac{1}{2}, 1\right], \left[1, 1\frac{1}{2}\right], \ldots$ contains $[1/4x]$ elements with the claimed property of the sequence $x, 2x, \ldots, \ldots nx$. Hence

$$([nx]+1)\cdot[1/4x] \geq nx\left(\frac{1}{4x}-1\right) \geq \frac{n}{4}-\frac{n}{8} \geq \frac{n}{8}.$$

We are ready, since it is evident, that there exists at least one element with the claimed property.

LEMMA 5. Denote $C = \left[\frac{\pi}{2n}, \frac{\pi}{8}\right], D = \left[\frac{\pi}{8}, \frac{\pi}{2}\right]$. If $p \geq 3$, $x \in D \cup C$ then

$$g(x) \leq \frac{C}{(p+1)} e^{-cn}, \ (n = 2, 3, \ldots).$$

PROOF. Denote $\langle t \rangle = \min\limits_{S \text{ integer}} |t - s\pi|$. From (12) and (13) we have

$$g(x) \leq \prod_{k=1}^{n} \frac{\min\left(p+1, \dfrac{1}{1-|\cos x|}\right)}{p+1} \leq$$

$$\leq \left(\frac{2}{2-\sqrt{2}} \cdot \frac{1}{p+1}\right)^{\left|\left\{k \in \{1, \ldots, n\} : k\langle kx \rangle \geq \frac{\pi}{4}\right\}\right|}.$$

Lemma 4 implies for $n \geq 4$ and $x \in C$

$$g(x) \leq \left(\frac{3{,}5}{p+1}\right)^{\max(n/8, 1)},$$

and there exist C_0, n_0, p_0 such that for every $n \geq n_0$, $p \geq p_0$

$$g(x) \leq \frac{C_0}{(p+1)} e^{-cn}, \ (x \in C).$$

We let denote

$$C_1 = \sup_{2 \leq n \leq n_0} \sup_{3 \leq p} \frac{e^{-cn}}{(p+1)} \left(\frac{3{,}5}{p+1}\right)^{\max(n/8,1)},$$

$$C_2 = \sup_{2 \leq n} \sup_{3 \leq p \leq p_0} \frac{e^{-cn}}{(p+1)} \left(\frac{3{,}5}{p+1}\right)^{\max(n/8,1)},$$

these numbers are finite.

Comparing these results for $x \in C$, $n \geq 2$, $p \geq 3$:

$$g(x) \leq \frac{\max(C_0, C_1, C_2)}{(p+1)} e^{-cn}.$$

In the case $x \in D$ we use similar arguments. There exist at least $\max\left(\left[\frac{n}{8}\right], 1\right)$ elements of the sequence $x, 2x, \ldots, nx$ such that $\langle k \cdot x \rangle \geq \pi/4$. Then

$$g(x) \leq \left(\frac{3{,}5}{p+1}\right)^{\max\left(\left[\frac{n}{8}\right], 1\right)}$$

for $x \in D$, $n \geq 2$, $p \geq 3$; what means

$$g(x) \leq C \frac{e^{-cn}}{(p+1)}.$$

LEMMA 6. If $p = 1$, $\frac{3\pi}{2n} \leq x \leq \frac{\pi}{2}$, then

$$g(x) \leq \left(\frac{1}{\sqrt{2}}\right)^{\max\left(\left[\frac{n}{8}\right], 1\right)}.$$

PROOF.

$$\prod_{k=1}^{n} \frac{\sin 2kx}{2 \sin kx} = \prod_{k=1}^{n} \cos kx,$$

applying the methods of Lemma 4 and 5, we have: $\langle kx \rangle \geq \pi/4$ implies $|\cos kx| \leq 1/\sqrt{2}$.

LEMMA 7. If $p = 2$, $\frac{3\pi}{2n} \leq x \leq \frac{\pi}{2}$, we have

$$g(x) \leq \frac{1}{\sqrt{3}} \cdot 3^{-n/4}.$$

PROOF. Use the elementary inequality

$$|1-t| \leq 27^{-t/4} = e^{-\frac{\ln 27}{4} t} \quad \text{for} \quad 0 \leq t \leq \frac{4}{3}.$$

We have

$$g(x) = \left| \prod_{k=1}^{n} \frac{\sin 3kx}{3 \sin kx} \right| = \left| \prod_{k=1}^{n} 1 - \frac{4}{3} \sin^2 kx \right| \leq$$

$$\leq \prod_{k=1}^{n} e^{-\frac{\ln 27}{3} \sin^2 kx} = e^{-\frac{n+1}{2} - \frac{\sin (n+1)x \cos nx}{2 \sin x}} \leq \frac{1}{\sqrt{3}} 3^{-n/4}.$$

Summarizing our results, we obtain: there exists universal constant K with

$$\int_{\frac{3\pi}{n(p+1)}}^{\frac{\pi}{2}} \left| \prod_{k=1}^{n} \frac{\sin (p+1)kx}{(p+1)\sin kx} \right| dx \leq \frac{K}{(p+1)} e^{-cn} \tag{16}$$

for every $n \geq 2$, $p \geq 1$, further

$$|I_2| \leq \int_{\frac{c}{np} \leq |t| \leq \pi} \prod_{k=1}^{n} \left| \frac{\sin p+1)kt/2}{(p+1)\sin kt/2} \right| dt =$$

$$= 2 \int_{\frac{c}{2np} \leq |t| \leq \frac{\pi}{2}} \prod_{k=1}^{n} \left| \frac{\sin (p+1)\, kt}{(p+1)\sin kt} \right| dt \leq \frac{K_2}{p} e^{-cn}.$$

5. The estimation of I_3

This estimation is much more easier than the previous ones. We have

$$e^{-D_n^2 t^2/2} \leq e^{-n^3 p^2 t^2}$$

and hence

$$|I_3| \leq \int_{|t| \geq \frac{c}{np}} e^{-n^3 p^2 t^2} dt = \frac{1}{n^{3/2} p} \int_{|t| \geq c\sqrt{n}} e^{-t^2} dt \leq \frac{K_3}{p} e^{-cn}.$$

6. Remark

Our theorem gives an asymptotic formula if and only if

$$(1-\varepsilon)\, D_n^2 \log n > (N - A_n)^2,$$

since in the opposite case the term of error exceeds the main term. It follows, that our results are incomparable with those of HARDY and RAMANUJAN.

The ideas and the result of the present paper were, further developed by A. BOGMÉR and L. A. SZÉKELY in [5].

References

[1] P. ERDŐS and J. LEHNER, The distribution of the number of summands in the partitions of a positive integer, *Duke Math., Journal*, **8** (1941), 335–345.
[2] G. H. HARDY and S. RAMANUJAN, Asymptotic formulae in combinatory analysis, *Proc. London Math. Soc.*, **17** (1918), 75–115.
[3] M. SZALAY and P. TURÁN, On some problems of the statistical theory of partitions with application to characters of the symmetric group. III, *Acta Math. Acad. Sci. Hung.*, **32** (1978), 129–155.
[4] S. CH. SIRASHDINOFF, T. A. AZLAROFF and T. M. ZUPAROFF, *Problems in additive number theory with increasing number of terms*, FAN, Tashkent, 1975.
[5] A. BOGMÉR and L. A. SZÉKELY, Asymptotic formula for the number of solutions of a Diophantic system, *Annales Univ. Sci. Budapest*, Sectio Mathematica, **28** (1985), 203–215.

ON THE RANGE OF CERTAIN FUNCTIONALS OF THE CALCULUS OF VARIATIONS

By

A. KÓSA and A. SHAMANDY

II. Department for Analysis of the L. Eötvös University, Budapest and Department of Mathematics, Mansoura University, Egypt

(Received August 31, 1983)

In problems of calculus of variations it occurs frequently that after a substantial extension of a functional the infinum and supremum of its range do not change (see e.g. [2], [4]). In this paper a theorem of this type will be proved, namely for such an extension of a functional studied in the classical case, the investigation of which is important from the point of view of applications ([5]).

Let $a, b, c, d \in \mathbf{R}$, $a < b$. Denote by F the class of all functions of first kind defined in $[a, b]$. $\varphi \in F$ means that the domain of φ is $[a, b]$ except for an at most countable set, moreover, φ has a finite right limit at each point of the interval $[a, b[$ and a finite left limit at each point of the interval $(]a, b]$ (see [1], [3]). Denote by D the part of F consisting of functions which are not defined on an at most finite set and for which the set where the left and right limits are different is also at most finite.

Let D_1 and F_1 the set of the integral functions of the functions in D and F respectively. For a $\varphi \in F$ we shall also use the following notation

$$\int_a \varphi : [a, b] \to \mathbf{R}, \quad t \mapsto \int_a^t \varphi.$$

Define the following norm in F: for any $\varphi, \psi \in F$

$$\|\varphi - \psi\| := \sup_{t \in [a, b[} |\varphi(t+0) - \psi(t+0)| + \sup_{t \in]a, b]} |\varphi(t-0) - \psi(t-0)|.$$

Let $j := id_{[a, b]}$ and for any $\alpha \in \mathbf{R}$ define

$$\alpha(\cdot) := [a, b] \to \mathbf{R}, \quad t \mapsto \alpha.$$

Introduce the following sets,

$$L := \{x \in F_1 | \varphi(a) = c, \quad \varphi(b) = d\},$$

$$M := \{x \in D_1 | \varphi(a) = c, \quad \varphi(b) = d\}.$$

Let $f: \mathbf{R}^3 \to \mathbf{R}$ a continuously differentiable function and denote by f_i the partial derivative of f with respect to the i-th variable ($i = 1, 2, 3$). Define the following functionals:

$$I: L \to \mathbf{R}, \quad x \mapsto \int_a^b f \circ (j, x, \dot{x}),$$

$$J: M \to \mathbf{R}, \quad x \mapsto \int_a^b f \circ (j, x, \dot{x}).$$

Clearly

$$I|_M = J.$$

The infimum and supremum of the range of a real-valued function Φ are called shortly the infimum and supremum of Φ and are denoted by $\inf \Phi$ and $\sup \Phi$ respectively.

The following theorem holds.

THEOREM. $\inf J = \inf I$, $\sup J = \sup I$.

PROOF. For the proof of the theorem it is clearly enough to show the following: for any $\varphi \in L$ and $\varepsilon > 0$ there exists a $\psi \in M$ such that

$$|I(\varphi) - I(\psi)| < \varepsilon. \tag{1}$$

Fix the function $\varphi \in L$ arbitrarily.

The function φ' is of first kind, therefore, there exists a finite interval $[\sigma, \tau]$ containing the range of φ'.

Put

$$U := \{(t, s, u) \in \mathbf{R}^3 | t \in [a, b],\ s \in [\varphi(t) - 1, \varphi(t) + 1],\ u \in [\sigma - 1, \tau + 1]\}.$$

U is clearly a bounded set. Define

$$Q := \max\{\sup|f_{\cdot 2}|_U|, \sup|f_{\cdot 3}|_U|\}.$$

Now fix an arbitrary number

$$\delta \in]0, \min\{1, 1/(b-a)\}[. \tag{2}$$

The function φ' is of first kind, so there exists a step function $g: [a, b] \to \mathbf{R}$ such that

$$\|\varphi' - g\| < \frac{\delta}{2}. \tag{3}$$

Hence

$$g - \frac{1}{2}\delta(\cdot) < \varphi' < g + \frac{1}{2}\delta(\cdot)$$

which implies that

$$\int_a^b \left(g - \frac{1}{2}\delta(\cdot)\right) < \int_a^b \varphi' < \int_a^b \left(g + \frac{1}{2}\delta(\cdot)\right).$$

From the above inequalities it is clear that a number

$$\alpha \in \left]-\frac{\delta}{2}, \frac{\delta}{2}\right[\tag{4}$$

can be chosen such that the equality

$$\int_a^b \varphi' = \int_a^b (g+\alpha(\cdot)) \tag{5}$$

holds. Define

$$\psi := c(\cdot)+\int_a (g+\alpha(\cdot)).$$

It is obvious that $\psi \in D_1$. From (3) and (4) and from the definition of ψ it follows that

$$\|\varphi'-\psi'\| < \delta. \tag{6}$$

Taking in account (5), from the definition of ψ we get

$$\psi(a) = c,$$

$$\psi(b) = c+\int_a^b (g+\alpha(\cdot)) = c+\int_a^b \varphi' = c+(d-c) = d.$$

From the first equality by (6) we obtain that

$$\|\varphi-\psi\| \leq \delta(b-a). \tag{7}$$

From the relation (2) and the inequalities (6), (7) by the definition of the set U it follows that at any point t of the intersection of the domains of φ' and ψ' for any $\vartheta \in]0, 1[$ we have

$$(t, \varphi(t)+\vartheta(\psi(t)-\varphi(t)), \varphi'(t)+\vartheta(\psi'(t)-\varphi'(t)) \in U. \tag{8}$$

Now estimate the difference of corresponding values of the functional. Using (8) and the definition of Q a simple calculation shows that

$$|I(\varphi)-I(\psi)| \leqq \int_a^b |f\circ(j,\varphi,\varphi')-f\circ(j,\psi,\psi')| \leqq$$

$$\leq Q\delta(b-a)^2+Q\delta(b-a) = \delta\{Q(b-a)^2+Q(b-a)\}.$$

Hence if $\varepsilon > 0$ is fixed and the number δ is chosen sufficiently small the inequality (1) holds. Theorem is proved.

REMARK 1. Let N be the part of M consisting of all polynomials and let K be the restriction of the function J onto the set N:

$$K := J|_N.$$

From the classical results of calculus of variations it is known (see e.g. [4] chapter 6) that

$$\inf K = \inf J, \quad \sup K = \sup J.$$

So our theorem implies that

$$\inf K = \inf I, \ \sup K = \sup I.$$

REMARK 2. Introduce in L the following norm

$$\|\varphi\|_1 := \|\varphi\| + \|\varphi'\|. \tag{9}$$

In the proof of our theorem, essentially, we have proved first that D_1 is dense in F_1 with respect to the metrics induced by (9), then it has been shown that I is a continuous functional with respect to the same metrics.

References

[1] DIEUDONNEÉ, J.: *Foundations of Modern Analysis*, Academic Press, 1969.
[2] COURANT, R. – HILBERT, D.: *Methods of Mathematical Physics*, Vol. 1, Interscience Publishers, 1953.
[3] KÓSA, A.: *Ismerkedés a matematikai analízissel* (A first look at mathematical analysis), Műszaki Könyvkiadó, Budapest, 1981.
[4] KÓSA, A : *Variációszámítás*, Tankönyvkiadó, 1970. (Russian edition: КОША, А.: Вариационное исчисление, Высшая школа, 1983.)
[5] KÓSA, A. – SHAMANDY, A.: On the smoothness properties of stationary functions arising in calculus of variations, *Annales Univ. Sci. Budapest*, Sectio Computatorica, 5 (1984), 29 – 35.

DIE ZAHL DER OVALE IN DER BOLYAI – LOBATSCHEFSKY EBENE S⟨3, 3⟩

Von

T. HORVÁTH

L. Eötvös Universität, Budapest

(Eingegangen am 19. September 1983)

Es sei **S** eine endliche Punktmenge. Bestimmte Teilmengen von **S** werden Geraden genannt, wenn die folgenden Erforderungen erfüllt sind:

1. *Zwei beliebige verschiedene Elemente von* **S** *werden von genau einer Geraden genannten Teilmenge von* **S** *enthält.*

2. *Für jedes Paar, das aus einem beliebigen Punkt und aus einer nicht durch diesen Punkt verlaufenden Gerade von* **S** *besteht, gilt es, daß es m schneidende bzw. n nicht schneidende Geraden durch den Punkt eines Paares zu dem Geradenelement dieses Paares gibt.*

3. *Es gelten* $m > 2$ *und* $n > 2$.

Die den vorigen drei Erforderungen entsprechende Struktur wird eine BOLYAI – LOBATSCHEFSKY Ebene (kurz **B – L** Ebene) mit Charakter $\langle m, n\rangle$ genannt. Diese Struktur wird mit $\mathbf{S}\langle m, n\rangle$ bezeichnet.

Das Oval ist eine Punktmenge mit höchsten Punktanzahl in einer endlichen Ebene, die höchstens zwei gemeinsame Punkte mit einer Gerade hat und die in jedem von ihren Punkten genau eine Gerade *(Tangente)* besitzt, die keinen weiteren gemeinsamen Punkt mit der obigen Teilmenge hat.

Zwei Geraden, die bezüglich eines Punktes perspektiv sind, und das Zentrum der Perspektivität werden eine *Π-Konfiguration* genannt. Diese Konfiguration besteht aus sieben Punkten und fünf Geraden (Fig. 1.).

Vier Geraden bilden ein vollständiges Vierseit, wenn sich jede zwei Geraden schneiden und nicht drei von ihnen durch einen gemeinsamen Punkt verlaufen.

In dieser Arbeit beschäftigen wir uns nur mit der **B – L** Ebene **S** ⟨3, 3⟩. Diese Ebene besteht aus 13 Punkten und 26 Geraden, weiterhin jedes Oval besteht aus 6 Punkten. Deshalb gilt die zweite Erforderung in der Defintion des Ovals offenbar, wenn die erste Erforderung gilt. Bis jetzt sind zwei, miteinander nicht isomorphe Ebenen **S**⟨3, 3⟩ (siehe in [2]) bekannt.

Zuerst beschäftigen wir uns mit den Kollineationen dieser zwei Ebenen. Dann beweisen wir, daß $o = 3n+p$ gilt, wo o die Zahl der Ovale, n die Zahl der vollständigen Vierseite und p die Zahl der Π-Konfigurationen sind.

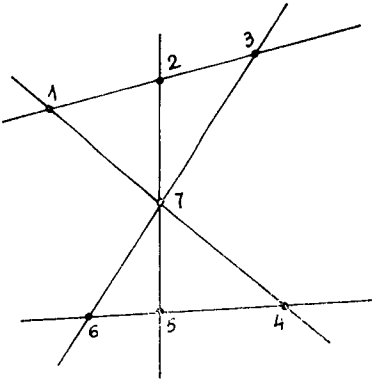

Fig. 1.

1. Wir schreiben die Geraden der zwei bekannten miteinander nicht isomorphen Ebenen $\mathbf{S}\langle 3, 3\rangle$ auf. Mit den Elementen der Menge $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13\}$ werden die Punkte der erwähnten zwei Ebenen $\mathbf{S}_1$ und $\mathbf{S}_2$ bezeichnet. Die Geraden der Ebene $\mathbf{S}_1\langle 3, 3\rangle$ sind die Punktmengen (siehe [1]):

$\{1, 2, 3\}$, $\{1, 4, 5\}$, $\{1, 6, 7\}$, $\{1, 8, 9\}$, $\{1, 10, 11\}$, $\{1, 12, 13\}$, $\{2, 4, 6\}$, $\{2, 5, 7\}$, $\{2, 8, 10\}$, $\{2, 9, 12\}$, $\{2, 11, 13\}$, $\{3, 4, 8\}$, $\{3, 5, 9\}$, $\{3, 6, 11\}$, $\{3, 7, 13\}$, $\{3, 10, 12\}$, $\{4, 7, 10\}$, $\{4, 9, 13\}$, $\{4, 11, 12\}$, $\{5, 6, 12\}$, $\{5, 8, 11\}$, $\{5, 10, 13\}$, $\{6, 8, 13\}$, $\{6, 9, 10\}$, $\{7, 8, 12\}$, $\{7, 9, 11\}$,

und die Geraden der Ebene $\mathbf{S}_2\langle 3, 3\rangle$ sind die folgenden Punktmengen (s. [2]):

$\{1, 2, 3\}$, $\{1, 4, 5\}$, $\{1, 6, 7\}$, $\{1, 8, 9\}$, $\{1, 10, 11\}$, $\{1, 12, 13\}$, $\{2, 4, 6\}$, $\{2, 5, 7\}$, $\{2, 8, 10\}$, $\{2, 9, 12\}$, $\{2, 11, 13\}$, $\{3, 4, 8\}$, $\{3, 5, 9\}$, $\{3, 6, 10\}$, $\{3, 7, 13\}$, $\{3, 11, 12\}$, $\{4, 7, 12\}$, $\{4, 9, 11\}$, $\{4, 10, 13\}$, $\{5, 6, 11\}$, $\{5, 8, 13\}$, $\{5, 10, 12\}$, $\{6, 8, 12\}$, $\{6, 9, 13\}$, $\{7, 8, 11\}$, $\{7, 9, 10\}$.

Die Ebene $\mathbf{S}_1\langle 3, 3\rangle$ ist mit der endlichen hyperbolischen Ebene isomorph, die man im Artikel [1] finden kann. Deshalb sind die Kollineationen der Ebene $\mathbf{S}_1\langle 3, 3\rangle$ die folgenden:

	1	2	3	4	5	6	7	8	9	10	11	12	13
Π_1:	1	2	3	4	5	6	7	8	9	10	11	12	13
Π_2:	1	2	3	5	4	7	6	9	8	12	13	10	11
Π_3:	7	13	3	8	12	6	1	4	10	9	11	5	2
Π_4:	7	13	3	12	8	1	6	10	4	5	2	9	11
Π_5:	6	11	3	9	10	7	1	5	12	8	13	4	2
Π_6:	6	11	3	10	9	1	7	12	5	4	2	8	13

Die Gruppe dieser Kollineationen ist mit der symmetrischen Gruppe $\mathbf{S}_3$ *isomorph.*

Man kann eine beliebige Kollineation von $\mathbf{S}_2\langle 3, 3\rangle$ folgenderweise aufschreiben:

	1	2	3	4	5	6	7	8	9	10	11	12	13
ϱ:	13	5	8	1	12	4	10	9	6	3	7	11	2

Man kann sehen, daß ϱ der Zyklus (1, 13, 2, 5, 12, 11, 7, 10, 3, 8, 9, 6, 4) mit der Länge 13 ist. Die Potenzen von ϱ sind auch Kollineationen, deshalb hat die Gruppe der Kollineationen von $\mathbf{S}_2\langle 3, 3\rangle$ eine zyklische Teilgruppe von der Ordnung 13. Auch aus den Vorhergehengen folgt der in [2] bewiesene Satz: *die Ebene* $\mathbf{S}_1\langle 3, 3\rangle$ *und* $\mathbf{S}_2\langle 3, 3\rangle$ *sind miteinander nicht isomorph.*

2. Im folgenden betrachten wir eine beliebige Ebene $\mathbf{S}\langle 3, 3\rangle$, und zuerst beweisen wir den folgenden

SATZ 1. *In einer beliebigen* **B–L** *Ebene* $\mathbf{S}\langle 3, 3\rangle$ *ist die Zahl der Ovale mindestens dreimal soviel wie die Zahl der vollständigen Vierseite.*

Es sei N ein vollständiges Vierseit. Wir nennen den dritten, zu N nicht gehörigen Punkt einer Diagonale von N *einen Diagonalpunkt von* N. Ein Diagonalpunkt ist *mehrfach*, wenn dieser Punkt ein Diagonalpunkt von mehreren Diagonalen ist. Wenn der Diagonalpunkt nur zu einer einzigen Diagonale von N gehört, dann ist der Diagonalpunkt *einfach.*

Zuerst beweisen wir zwei Hilfssätze.

HILFSSATZ 1. *Einem beliebigen einfachen Diagonalpunkt des vollständigen Vierseits* N *können wir ein Oval zuordnen.*

BEWEIS. Diese Zuordnung ist das folgende. Mit 1 bezeichnen wir einen beliebigen einfachen Diagonalpunkt von N. Die Punkte 2, 3, 4, 5, 6, 7 von N seien so gewählt, daß die Punktmengen {1, 2, 3}, {2, 4, 5}, {2, 6, 7}, {3, 4, 7}, {3, 5, 6} Geraden bilden. Weil 1 ein einfacher Diagonalpunkt ist, deshalb liegen 1, 4 und 6 bzw. 1, 5 und 7 nicht auf einer Gerade, d.h., durch 1 kann man vier weitere Geraden ziehen. Diese Geraden sind {1, 4, }, {1, 5, }, {1, 6, }, {1, 7, }, deren dritte Punkte nacheinander 8, 9, 10, 11 sind. In der Ebene $\mathbf{S}\langle 3, 3\rangle$ können wir durch jeden Punkt sechs Geraden ziehen, deshalb gibt es eine sechste, von den vorigen verschiedene Gerade durch 1. Diese Gerade ist {1, 12, 13}. Wir beweisen, daß *die Punktmenge* Ω = {8, 9, 10, 11, 12, 13} *ein Oval ist* (Fig. 2.).

Nehmen wir indirekt an, daß Ω kein Oval ist. Dann hat Ω drei Punkte, die auf derselben Gerade g liegen. Wir zeigen, daß keine **B–L** Ebene $\mathbf{S}\langle 3, 3\rangle$ existiert, wo die vorige indirekte Bedingung gilt.

Die Gerade g kann zweierlei sein.

(1) *Die Punkte von* g *gehören zu der Menge* {8, 9, 10, 11}.

(2) *Zwei von den Punkten der Gerade* g *sind in der Menge* {8, 9, 10, 11} *und der dritte Punkt von* g *gehört zu der Menge* {12, 13}.

Wir betrachten den Fall (1). Ohne Beschränkung der Allgemeinheit können wir annehmen, daß {8, 9, 10} die oben erwähnte Gerade ist. In diesem Fall sind die Geraden {2, 8, }, {2, 9, }, {2, 10, } verschieden und die dritten Punkte dieser Geraden sind 11, 12, 13. (Auch eine andere Reihenfolge der Punkte kann vorkommen.) Ebendas kann man auch über die Geraden {3, 8, }, {3, 9, }, {3, 10, } sagen.

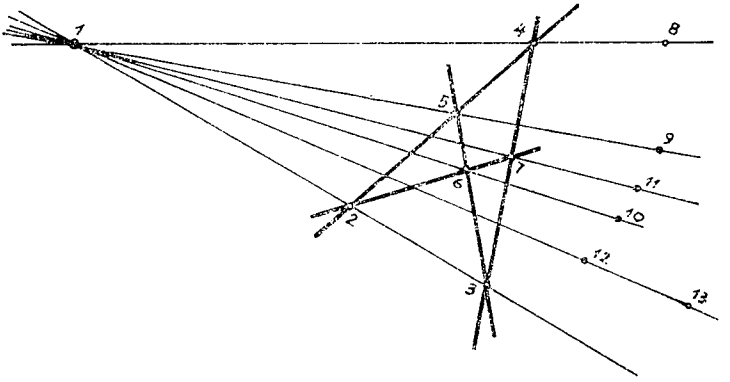

Fig. 2.

Wir betrachten die Geraden {7, 8, }, {7, 9, }, {7, 10, }. Die dritten Punkte dieser Geraden können die Punkte 5, 11, 12, 13 sein. Der dritte Punkt von {7, 9, } kann weder 5 noch 11 sein, weil 5 und 9 bzw. 7 und 11 schon mit einer Gerade ({1, 5, 9}, {1, 7, 11}) verbunden sind. Deshalb ist der dritte Punkt von {7, 9, } 12 oder 13. Wir können annehmen, daß dieser Punkt 12 ist.

Was kann der dritte Punkt der Gerade {4, 9, } sein? 8 und 10 sind nicht möglich, weil die Menge {8, 9, 10} eine Gerade ist. Auch 12 ist nicht möglich, weil auch {7, 9, 12} eine Gerade ist. 11 und 13 können keine dritte Punkte sein, weil 11 und 13 die dritten Punkte der Geraden {2, 9, } und {3, 9, } sind. Die Punkte 1, 2, 3, 5, 7 waren mit dem Punkt 4 verbunden, deshalb sind auch diese Punkte nicht möglich. Endlich bleibt nur eine einzige Möglichkeit für den dritten Punkt, nämlich der Punkt 6 und dann ist die Gerade eben die Gerade {4, 6, 9}. Aus ähnlichen Überlegungen ergibt sich, daß der dritte Punkt der Gerade {5, 12, } nur 11 sein kann.

Wir betrachten die Gerade {6, 12, }. Die Punkte 1, 2, 3, 4, 5, 7, 10 können die dritten Punkte dieser Gerade nicht sein, weil sie mit dem Punkt 6 schon verbunden sind. Auch die Punkte 8, 9, 11, 13 sind nicht möglich, weil sie mit dem Punkt 12 verbunden sind. Folglich hat die Gerade {6, 12, } keinen dritten Punkt. Das ist aber in *der* **B – L** *Ebene* **S**⟨3, 3⟩ *unmöglich.*

Nun betrachten wir den Fall (2). Hier müssen wir zwei Unterfälle untersuchen. {8, 9, 12} und {8, 10, 12} sind diese Fälle, die wir gemeinsam erörtern.

Wir betrachten die dritten Punkte der Geraden {2, 12, } und {3, 12, }. Diese Punkte können nur 10, 11 bzw. 9, 11 sein. Daraus folgt, daß {6, 4, 12} und {7, 5, 12} Geraden sind. Deshalb sind die Geraden {1, 12, 13}, {2, 13, }, {3, 13, }, {4, 13, }, {5, 13, }, {6, 13, }, {7, 13, } verschieden, d.h., durch den Punkt 13 verlaufen sieben verschiedene Geraden. Das ist aber *in der* **B–L** *Ebene* **S**⟨3, 3⟩ *nicht möglich.*

Wir können also ausgehend von einem beliebigen einfachen Diagonalpunkt eines vollständigen Vierseits ein Oval bekommen.

HILFSSATZ 2. *Das vollständige Vierseit N in der* **B–L** *Ebene* **S**⟨3, 3⟩ *hat nur einfache Diagonalpunkte, d. h., alle drei Diagonalpunkte von N sind verscheieden.*

BEWEIS. Wenn die drei Diagonalpunkte von N zusammenfallen, dann bilden N und dieser Diagonalpunkt eine Galois-Ebene **PG** (2). Betrachten wir einen Punkt, der von den vorigen Punkten verschieden ist. Wenn wir diesen Punkt mit den Punkten von **PG** (2) verbinden, bekommen wir sieben verschiedene Geraden. Das ist aber ein Widerspruch.

Wenn N einen zweifachen und einen einfachen Diagonalpunkt hat, dann tritt derselbe Fall auf, wie bei der Untersuchung des Falles (2) im Hilfssatz 1. Der Punkt 1 ist der einfache, 12 ist der zweifache Diagonalpunkt. Deshalb ist auch dieser Fall nicht möglich.

Der BEWEIS des SATZES 1. Nach dem Hilfssatz 2. hat ein vollständiges Vierseit drei verschiedene einfache Diagonalpunkte. Mit dem Verfahren, das wir im Hilfssatz 1. geschrieben haben, bekommen wir ein Oval zu jedem einfachen Diagonalpunkt. Diese Ovale sind verschieden, weil sie in je einem Punkt verschieden sind.

Wir müssen noch einsehen, daß wir aus sieben Punkten zwei vollständige Vierseite N_1 und N_2 derart nicht vorstellen können, daß der siebente Punkt in beiden Fällen der Diagonalpunkt des vollständigen Vierseits sei. Der Punkt 1 sei der Diagonalpunkt von N_1. Weil $N_1 \neq N_2$ ist, muß der Punkt 1 zu N_2 gehören. Deshalb kann man durch den Punkt 1 noch eine weitere Gerade ziehen, die zwei Punkte von N_1 enthält. Diese Gerade kann nur {1, 4, 6} oder {1, 5, 7} sein. Das würde aber bedeuten, daß 1 ein mehrfacher Diagonalpunkt wäre. Das ist aber nach dem Hilfssatz 2. unmöglich.

Also ist die *Zahl der Ovale mindestens dreimal soviel, wie die Zahl der vollständigen Vierseite.*

SATZ 2. *Ist eine aus sieben Punkten bestehende Teilmenge eine Π-Konfiguration in der Ebene* **S**⟨3, 3⟩, *dann bilden die zurückbleibenden sechs Punkte ein Oval.*

BEWEIS. Eine Π-Konfiguration besteht aus fünf Geraden. Es gibt sechs weitere Geraden in **S**⟨3, 3⟩, die je zwei Punkte von Π enthalten. Jede von den zurückbleibenden 15 Geraden enthält mindestens zwei Punkte, die nicht zu Π gehören. Höchstens 15 Geraden gehören zu sechs Punkten derart, daß jede Gerade mindestens zwei Punkte enthält. Wenn eine dieser Geraden schon drei Punkte von diesen sechs Punkten enthält, dann existieren weniger als 15 Geraden durch diese sechs Punkte. Deshalb gibt es keine Gerade, die von

den zurückbleibenden sechs Punkten bestimmt ist und die drei von diesen Punkten enthält. Daraus folgt schon, daß die Definition des Ovals erfüllt ist, d. h., die zurückbleibenden sechs Punkte ein Oval bilden.

Wir bemerken, daß man einen ähnlichen Beweis für den Hilfssatz 1. angeben kann.

SATZ 3. *Bilden sechs Punkte ein Oval Ω in der Ebene* $\mathbf{S}\langle 3, 3\rangle$, *dann bilden die zurückgebliebenen sieben Punkte und fünf Geraden, die keinen Punkt von Ω enthalten, entweder eine Π-Konfiguration oder ein vollständiges Vierseit mit seinem Diagonalpunkt.*

BEWEIS. Betrachten wir das Oval Ω und die Geraden, die durch die Punkte von Ω verlaufen. Aus der Definition des Ovals folgt, daß die Zahl dieser Geraden 21 ist. Die zurückgebliebenen Konfiguration besteht aus sieben Punkten und fünf Geraden. Diese Konfiguration wird eine L-Konfiguration genannt.

Nehmen wir an, daß höchstens zwei Geraden durch jeden Punkt von L durchlaufen. Dann ist die Zahl der Geraden höchstens $\frac{7\cdot 2}{3} > 5$, weil höchstens zwei Geraden durch jeden Punkt verlaufen und jede Gerade drei Punkte enthält. Deshalb gibt es einen Punkt M, durch den mindestens drei Geraden durchgehen. Auf diesen drei Geraden sollen diese sieben Punkte liegen. L hat noch zwei weitere Geraden. Diese zwei zurückgebliebenen Geraden enthalten je einen Punkt der obigen drei Geraden, der von M verschieden ist. So kann man diese zwei Geraden nur zweierlei legen. Die erste Möglichkeit ist, wenn sich diese zwei Geraden schneiden. In diesem Fall bildet L ein vollständiges Vierseit mit seinem Diagonalpunkt. Im anderen Fall, wenn sich die Geraden nicht schneiden, bekommen wir eine Π-Konfiguration.

Damit haben wir den Satz bewiesen.

SATZ 4. *In einer beliebigen Ebene* $\mathbf{S}\langle 3, 3\rangle$ *gilt* $o = 3n+p$, *wo o die Zahl der Ovale, n die Zahl der vollständigen Vierseite und p die Zahl der Π-Konfigurationen sind.*

BEWEIS. Aus den ersten und zweiten Sätzen dieses Abschnittes folgt $o \geq 3n+p$. Aus dem dritten Satz folgt, daß man ein einziges vollständiges Vierseit oder eine einzige Π-Konfiguration einem Oval zuordnen kann. Deshalb besteht die Gleichheit in der Ungleichung, d. h., $o = 3n+p$ gilt.

FOLGERUNG. In der Ebene $\mathbf{S}_1\langle 3, 3\rangle$ sind 8 vollständige Vierseite und 10 Π-Konfigurationen. Diese sind:

$\{1, 2, 4, 5, 6, 7\}$, $\{1, 3, 4, 5, 8, 9\}$, $\{1, 8, 6, 7, 13, 12\}$,
$\{1, 9, 6, 7, 10, 11\}$, $\{2, 4, 9, 12, 13, 11\}$, $\{2, 5, 8, 10, 11, 13\}$,
$\{3, 6, 5, 9, 12, 10\}$, $\{3, 7, 4, 8, 10, 12\}$,
$\{1, 2, 11, 13, 3, 10, 12\}$, $\{1, 4, 11, 12, 5, 10, 13\}$,
$\{2, 1, 8, 9, 3, 10, 12\}$, $\{3, 1, 6, 7, 2, 11, 13\}$,
$\{6, 2, 8, 10, 4, 13, 9\}$, $\{6, 2, 11, 13, 4, 3, 8\}$,
$\{7, 2, 9, 12, 5, 11, 8\}$, $\{7, 2, 11, 13, 5, 9, 3\}$,
$\{11, 3, 4, 8, 6, 12, 5\}$, $\{13, 3, 5, 9, 7, 10, 4\}$.

Deshalb gibt es $3\cdot 8+10 = 34$ Ovale in der Ebene $\mathbf{S}_1\langle 3, 3\rangle$.

In der Ebene $\mathbf{S}_2\langle 3, 3\rangle$ gibt es 13 vollständige Vierseite und gibt es keine Π-Konfiguration. Die vollständigen Vierseiten sind:

$$\{1, 2, 4, 5, 6, 7\}, \{1, 3, 4, 5, 8, 9\}, \{1, 6, 8, 9, 12, 13\},$$
$$\{1, 7, 8, 9, 11, 10\}, \{1, 10, 4, 5, 13, 12\}, \{1, 11, 2, 3, 13, 12\},$$
$$\{2, 3, 4, 6, 8, 10\}, \{2, 8, 5, 7, 13, 11\}, \{2, 9, 4, 6, 11, 13\},$$
$$\{2, 10, 5, 7, 12, 9\}, \{3, 5, 6, 10, 11, 12\}, \{3, 7, 4, 8, 12, 11\},$$
$$\{3, 9, 6, 10, 13, 7\}.$$

Deshalb gibt es $3 \cdot 13 + 0 = 39$ Ovale in der Ebene $\mathbf{S}_2\langle 3, 3\rangle$.

Literatur

[1] KÁRTESZI F., Egy legkisebb véges reguláris hiperbolikus sík, *MTA III. Oszt. Közl.*, **19** (1969), 5–7.

[2] KÁRTESZI, F.–HORVÁTH, T., Einige Bemerkungen bezüglich der Struktur von endlichen Bolyai–Lobatschefsky Ebenen, *Annales Univ. Sci. Budapest*, Sectio Math., **28** (1985), 263–270.

ON STRONGLY NONLINEAR ELLIPTIC EQUATIONS IN UNBOUNDED DOMAINS

By

L. SIMON

II. Department of Analysis of the L. Eötvös University, Budapest

(Received March 2, 1983)

Introduction

In [1] J. R. L. WEBB the following elliptic equalion has considered:

$$\sum_{|\alpha|\leq m}(-1)^{|\alpha|}D^{\alpha}f_{\alpha}(x,u,\ldots,D^{\beta}u,\ldots)+g(x,u)=F,\ x\in\Omega \tag{0.1}$$

where Ω is a possibly unbounded domain in $\mathbf{R}^n$, $|\beta|\leq m$ and the terms $f_\alpha(x,\xi)$ are required to have polynomial growth in ξ, in the term $g(x,u)$, however, no such growth restriction is imposed but it is supposed that g (essentially) satisfies the sign condition $g(x,u)u\geq 0$. He has proved the existence of solutions of the boundary value problems for (0.1).

In the present paper a similar existence theorem is proved for the more general (elliptic) equation

$$\sum_{|\alpha|\leq m}(-1)^{|\alpha|}D^{\alpha}f_{\alpha}(x,u,\ldots,D^{\beta}u,\ldots)+ \\ +\sum_{|\alpha|\leq l}(-1)^{|\alpha|}D^{\alpha}g_{\alpha}(x,u,\ldots,D^{\gamma}u,\ldots)=F \tag{0.2}$$

where $l\leq m-1$, $|\gamma|\leq m-1$ and in the terms $g_\alpha(x,\xi_0,\ldots,\xi_\gamma,\ldots)$ no growth restriction is imposed with respect to ξ_α but it is supposed that g (essentially) satisfies the condition $g_\alpha(x,\xi_0,\ldots,\xi_\gamma,\ldots)\,\xi_\alpha\geq 0$. The assumptions on f_α are similar to that of [1] but in certain sense are less restrictive than the assumptions in [1]. (Such assumptions has been formulated in [2]–[4].)

Another generalization of the equation (0.1) has been considered in [3].

1. The formulation of the main result

Let $\Omega\subset\mathbf{R}^n$ be a (possibly unbounded) domain, $p>1$ and m a nonnegative integer. Denote by $W_p^m(\Omega)$ the Sobolev space of real valued functions u whose distributional derivatives of order $\leq m$ belong to $L^p(\Omega)$. The norm in $W_p^m(\Omega)$ is defined by

$$\|u\|_{W_p^m(\Omega)}=\left\{\sum_{|\alpha|\leq m}\int_\Omega|D^\alpha u|^p\right\}^{1/p}$$

where $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a multiindex, $D^\alpha = D_1^{\alpha_1} \ldots D_n^{\alpha_n}$ $D_j = \dfrac{\partial}{\partial x_j}$. The expression $W_{p,0}^m(\Omega)$ will denote the closure in $\|\cdot\|_{W_p^m(\Omega)}$ of $C_0^\infty(\Omega)$, the infinitely differentiable functions with compact support contained in Ω.

Let N be the number of multiindices α satisfying the condition $|\alpha| \leqq m$. Assume that

a) V is a closed subspace of $W_p^m(\Omega)$ with the property that for any $u \in V$ there exist constants $c > 0$, $c' > 0$ and a sequence of functions $w_j \in V \cap L^\infty(\Omega)$ such that

$$(w_j) \text{ converges to } u \text{ weakly in } V \tag{1.1}$$

and for $|\alpha| \leqq l \leqq m-1$, $D^\alpha w_j \in L^\infty(\Omega)$,

$$|D^\alpha w_j(x)| \leqq c|D^\alpha u(x)| + c' \text{ a.e. in } \Omega. \tag{1.2}$$

b) The functions f_α: $\Omega \times \mathbf{R}^n \to \mathbf{R}$ satisfy the Carathéodory conditions, i.e. they are measurable in x for each fixed $\xi = (\xi_0, \ldots, \xi_\beta, \ldots) \in \mathbf{R}^N$ and continuous in ξ for almost all $x \in \Omega$.

c) For any number $c_1 > 0$ there is a number $c_2 > 0$ such that

$$\|u\|_V^p \leqq c_1 \text{ implies } \int_\Omega |f_\alpha(x, u, \ldots, D^\beta u, \ldots)|^q dx \leqq c_2, \tag{1.3}$$

where $\dfrac{1}{p} + \dfrac{1}{q} = 1$.

d) There exist constants $c_3 > 0$, $c_4 > 0$ such that for all $u \in V$ the inequality

$$\sum_{|\alpha| \leqq m} \int_\Omega f_\alpha(x, u, \ldots, D^\beta u, \ldots) D^\alpha u dx \geqq c_3 \|u\|_V^p - c_4 \tag{1.4}$$

holds.

e) There exist a function $\psi \in C_0^\infty(\mathbf{R}^n)$ and a continuous function $F_1 = F_1(R, \lambda) \geqq 0$ $(R > 0, \lambda > 0)$ with the following properties: for any fixed $R > 0$

$$\lim_{\lambda \to +0} \frac{F_1(R, \lambda)}{\lambda} = 0 \tag{1.5}$$

and $u, v \in V$, $\|u\|_V \leqq R$, $\|v\|_V \leqq R$ imply

$$\sum_{|\alpha| \leqq m} \int_\Omega [f_\alpha(x, u, \ldots, D^\beta u, \ldots) - f_\alpha(x, v, \ldots, D^\beta v, \ldots)] \cdot \tag{1.6}$$
$$\cdot (D^\alpha u - D^\alpha v) dx \geqq -F_1(R, \|(u-v)\psi\|_{W_p^{m-1}(\Omega)}).$$

f) The functions p_α, r_α: $\Omega \times \mathbf{R}^M \to \mathbf{R}$ satisfy the Carathéodory conditions (i.e. p_α, r_α are measurable in x for each fixed $\xi' \in \mathbf{R}^M$ and continuous in ξ' for almost all $x \in \Omega$) and

$$g_\alpha(x, \xi') = p_\alpha(x, \xi') + r_\alpha(x, \xi'), \ |\alpha| \leqq l$$

where M denotes the number of multiindices γ with $|\gamma| \leqq m-1$.

g) For almost all $x \in \Omega$, for all $\xi' \in \mathbf{R}^M$ and $|\alpha| \leqq l$

$$p_\alpha(x, \xi')\xi_\alpha \geqq 0 \tag{1.7}$$

and

$$|r_\alpha(x, \xi')| \leqq h_\alpha(x), \text{ where } h_\alpha \in L^q(\Omega) \cap L^1(\Omega). \tag{1.8}$$

h) Denote by A the set of multiindices $\alpha = (\alpha_1, \ldots, \alpha_n)$ satisfying the condition $|\alpha| \leqq m$ and let

$$A = \bigcup_\nu A_\nu \text{ where } A_\nu \cap A_\mu = \emptyset \text{ if } \nu \neq \mu$$

and

$$\alpha, \alpha^* \in A_\nu \text{ if and only if } g_\alpha = g_{\alpha^*}.$$

Further, for a fixed ν denote by ξ'' the vector consisting of those coordinates of $\xi' = (\xi_0, \ldots, \xi_\gamma, \ldots) \in \mathbf{R}^M$ which satisfy $\gamma \notin A_\nu$. For any $s > 0$, $x \in \Omega$ let

$$g_{\nu,s}(x, \xi'') = \sup\{|g_\alpha(x, \xi')| : |\xi_\alpha| \leqq s \text{ for } \alpha \in A_\nu\}.$$

Suppose that for any number $c > 0$ there is a function $g^*_{\nu,s} \in L^1(\Omega)$ such that

$$\|u\|_V^p \leqq c \text{ implies } g_{\nu,s}(x, u, \ldots, D^\gamma u, \ldots) \leqq g^*_{\nu,s}(x) \text{ a.e. (where } \gamma \notin A_\nu). \tag{1.9}$$

The main result of this paper is the following

THEOREM. *Assume that conditions* a)–h) *are fulfilled. Then for any* $F \in V^*$ *(i. e. for any linear continuous functional on* V*) there exists* $u \in V$ *such that*

$$g_\alpha(\cdot, u, \ldots, D^\alpha u, \ldots) \in L^1(\Omega), \tag{1.10}$$

$$g_\alpha(\cdot, u, \ldots, D^\gamma u, \ldots) D^\alpha u \in L^1(\Omega) \quad (|\alpha| \leqq l) \quad \textit{and}$$

$$\sum_{|\alpha| \leqq m} \int_\Omega f_\alpha(x, u, \ldots, D^\beta u, \ldots) D^\alpha v dx + \tag{1.11}$$

$$+ \sum_{|\alpha| \leqq l} \int_\Omega g_\alpha(x, u, \ldots, D^\gamma u, \ldots) D^\alpha v dx = \langle F, v \rangle$$

for all $v \in V$ *satisfying* $\partial^\alpha v \in L^\infty(\Omega)$ *(if* $|\alpha| \leqq l$*) and for* $v = u$.

REMARKS. 1. In [1] it is shown that assumption a) is satisfied for $l = 0$ in the interesting cases $V = W^m_{p,0}(\Omega)$ and $V = W^m_p(\Omega)$. Further if $l < m - \frac{n}{p}$ and the boundary of Ω is sufficiently smooth then the assumption a) is trivially satisfied because of the imbedding $W^m_p(\Omega) \subset C^l_*(\Omega)$ where $C^l_*(\Omega)$ denotes the set of l times continuously differentiable functions u on Ω with the property that $\partial^\alpha u$ is bounded for $|\alpha| \leqq l$.

2. In [2], [4] and [5] there are formulated simple algebraic conditions which imply c) resp. d) and e).

3. The assumption (1.7) is satisfied if and only if

$$p_\alpha(x, \xi') \geqq 0 \quad \text{for} \quad \xi_\alpha \geqq 0$$
$$p_\alpha(x, \xi') \leqq 0 \quad \text{for} \quad \xi_\alpha < 0.$$

4. Let $l^* < m - \frac{n}{p}$ and suppose that for all $s > 0$ and all ν there is a function $g^0_{\nu,s} \in L^1(\Omega)$ such that for almost all $x \in \Omega$

$$\sup\{|g_\alpha(x, \xi')| : |\xi_\alpha| \leqq s \text{ for } \alpha \in A_\nu, |(\xi_0, \ldots, \xi_\gamma, \ldots)| \leqq s \text{ for } |\gamma| \leqq l^*\} \leqq$$
$$\leqq g^0_{\nu,s}(x).$$

Then by the boundedness of the imbedding $W^m_p(\Omega) \subset C^{l^*}_*(\Omega)$ the assumption h) is satisfied.

2. The proof of the existence theorem

Further suppose that the assumptions a)–h) are satisfied. For any $u, v \in V$ let

$$\langle T(u), v\rangle = \sum_{|\alpha| \leqq m} \int_\Omega f_\alpha(x, u, \ldots, D^\beta u, \ldots) D^\alpha v dx.$$

Then from the assumptions b), c) it follows that $T(u) \in V^*$. (See also [2], [4]). First we shall prove several lemmas. (Similar lemmas are proved in [1] and [3].)

LEMMA 1. $T: V \to V^*$ is a bounded (nonlinear) operator. Further T is continuous in finite dimension, i.e. if the sequence of $c^j = (c^j_1, \ldots, c^j_r) \in \mathbf{R}^r$ converges to $c = (c_1, \ldots, c_r) \in \mathbf{R}^r$ then for any fixed $u_1, \ldots, u_r, v \in V$

$$\lim_{j \to \infty} \langle T(c^j_1 u_1 + \ldots + c^j_r u_r), v\rangle = \langle T(c_1 u_1 + \ldots + c_r u_r), v\rangle. \tag{2.1}$$

PROOF. The boundedness of T follows from (1.3) and the inequality

$$|\langle T(u), v\rangle \leqq \sum_{|\alpha| \leqq m} \left\{\int_\Omega |f_\alpha(x, u, \ldots, D^\beta u, \ldots)|^q dx\right\}^{1/q} \left\{\int_\Omega |D^\alpha v|^p dx\right\}^{1/p}.$$

Denote $c^j_1 u_1 + \ldots + c^j_r u_r$ by u^j and $c_1 u_1 + \ldots + c_r u_r$ by u°. By assumption b) the sequence of functions

$$f_\alpha(\cdot, u^j, \ldots, D^\beta u^j, \ldots) D^\alpha v$$

converges a.e. in Ω to the function

$$f_\alpha(\cdot, u^\circ, \ldots, D^\beta u^\circ, \ldots) D^\alpha v$$

as $j \to \infty$. Moreover, for any measurable set $E \subset \Omega$

$$\int_E |f_\alpha(x, u^j, \ldots, D^\beta u^j, \ldots) D^\alpha v| \leqq \int_E |f_\alpha(x, u^j, \ldots, D^\beta u^j, \ldots)|^q dx \cdot \int_E |D^\alpha v|^p dx.$$

By virtue of (1.3) $\int_E |f_\alpha(x, u^j, \ldots, D^\beta u^j, \ldots)|^q \, dx$ is bounded thus by the Vitali convergence theorem we obtain (2.1).

LEMMA 2. The operator $T: V \to V^*$ is pseudomonotone, i. e. whenever a sequence (u_j) converges to u weakly in V, $T(u_j)$ converges to y weakly in V^* and

$$\limsup_{j\to\infty} \langle T(u_j), u_j - u\rangle \leqq 0$$

then

$$y = T(u) \text{ and } \lim_{j\to\infty} \langle T(u_j), u_j - u\rangle = 0.$$

PROOF. Since by Lemma 1 T is continuous in finite dimension thus assumption e) implies that T is pseudomonotone (see [3]).

For any number $\mu > 0$ let

$$g_{\alpha,\mu}(x, \xi') = \chi_\mu(x) p_{\alpha,\mu}(x, \xi') + r_\alpha(x, \xi') \tag{2.2}$$

where

$$p_{\alpha,\mu}(x, \xi') = \begin{cases} p_\alpha(x, \xi') \text{ if } |p_\alpha(x, \xi')| \leqq \mu, \\ \mu \dfrac{p_\alpha(x, \xi')}{|p_\alpha(x, \xi')|} \text{ otherwise} \end{cases} \tag{2.3}$$

and

$$\chi_\mu(x) = \begin{cases} 1 \text{ if } x \in \Omega, \ |x| \leqq \mu, \\ 0 \text{ if } x \in \Omega, \ |x| > \mu. \end{cases}$$

Then by (1.8)

$$|g_{\alpha,\mu}(x, \xi')| \leqq \mu\chi_\mu(x) + h_\alpha(x); \ g_{\alpha,\mu} \in L^q(\Omega). \tag{2.4}$$

LEMMA 3. If $\lim_{j\to\infty} D^\alpha u_j = D^\alpha u$ a.e. in Ω for $|\alpha| \leqq m-1$ then the sequence of functions $g_{\alpha,\mu}(\cdot, u_j, \ldots, D^\gamma u_j, \ldots)$ converges to $g_{\alpha,\mu}(\cdot, u, \ldots, D^\gamma u, \ldots)$ in $L^q(\Omega)$.

PROOF. In virtue of the assumption b) $g_{\alpha,\mu}(\cdot, u_j, \ldots, D^\gamma u_j, \ldots)$ converges to $g_{\alpha,\mu}(\cdot, u, \ldots, D^\gamma u, \ldots)$ a.e. as $j \to \infty$. By the estimation (2.4)

$$|g_{\alpha,\mu}(x, u_j, \ldots, D^\gamma u_j, \ldots) - g_{\alpha,\mu}(x, u, \ldots, D^\gamma u, \ldots)|^q \leqq c[\mu\chi_\mu(x) + h_\alpha(x)]^q$$

where the constant c does not depend on x and j, thus Lebesgue's dominated convergence theorem implies the assertion of Lemma 3.

By the inequality (2.4)

$$\langle S_\mu(u,)v\rangle = \sum_{|\alpha|\leqq l} \int_\Omega g_{\alpha,\mu}(x, u, \ldots, D^\gamma u, \ldots) D^\alpha v dx, \ v \in V \tag{2.5}$$

defines a continuous linear functional $S_\mu(u)$ on V.

LEMMA 4. The (nonlinear) operator S_μ: $V \to V^*$ is bounded and continuous in finite dimension.

PROOF. The boundedness of S_μ follows from the estimation (2.4). Lemma 3 implies that S_μ is continuous in finite dimension.

LEMMA 5. $T+S_\mu$ is pseudomonotone operator.

PROOF. Suppose that (u_j) converges to u weakly in V, $((T+S_\mu)(u_j))$ converges to $y\in V^*$ weakly in V^* and

$$\limsup_{j\to\infty} \langle (T+S_\mu)(u_j),\ u_j-u\rangle \leqq 0. \tag{2.6}$$

Then there is a subsequence (u'_j) of (u_j) such that

$$\lim_{j\to\infty} (D^\alpha u'_j) = D^\alpha u \text{ a.e. in } \Omega \text{ if } |\alpha|\leqq m-1$$

(see e.g. [6]). Thus by Lemma 3

$$\lim_{j\to\infty} \|g_{\alpha,\mu}(\cdot\, u'_j, \ldots, D^\gamma u'_j, \ldots) - g_{\alpha,\mu}(\cdot\,, u, \ldots, D^\gamma u, \ldots)\|_{L^q(\Omega)} = 0 \tag{2.7}$$

whence

$$\lim_{j\to\infty} S_\mu(u'_j) = S_\mu(u) \text{ weakly in } V^*$$

and

$$\lim_{j\to\infty} T(u'_j) = y - S_\mu(u) \text{ weakly in } V^*. \tag{2.8}$$

From equality

$$\langle S_\mu(u'_j), u'_j-u\rangle = \langle S_\mu(u'_j)-S_\mu(u), u'_j-u\rangle + \langle S_\mu(u), u'_j-u\rangle$$

it follows that

$$\lim_{j\to\infty} \langle S_\mu(u'_j),\ u'_j-u\rangle = 0 \tag{2.9}$$

because by (2.7), the boundedness of $\|u'_j-u\|_V$ and Hölder's inequality

$$\lim_{j\to\infty} \langle S_\mu(u'_j)-S_\mu(u),\ u'_j-u\rangle = 0.$$

Therefore (2.6) implies

$$\limsup_{j\to\infty} \langle T(u'_j),\ u'_j-u\rangle \leqq 0. \tag{2.10}$$

Since T is pseudomonotone (see Lemma 2), by (2.8), (2.10)

$$T(u) = y - S_\mu(u),$$

i. e.

$$(T+S_\mu)(u) = y;$$

further

$$\lim_{j\to\infty} \langle T(u'_j),\ u'_j-u\rangle = 0$$

and so by (2.9)

$$\lim_{j\to\infty} \langle (T+S_\mu)(u'_j),\ u'_j-u\rangle = 0. \tag{2.11}$$

(2.11) is valid also for the sequence (u_j) because else by the above arguments we get to a contradiction and so the proof is complete.

LEMMA 6. Assume that (u_j) converges weakly to u in V and there is a constant c such that

$$\sum_{|\alpha|\leqq l}\int_\Omega g_{\alpha,j}(x, u_j, \ldots, D_\gamma u_j, \ldots)D^\alpha u_j dx \leqq c. \tag{2.12}$$

Then for all α with $|\alpha|\leqq l$

$$g_\alpha(\cdot, u, \ldots, D^\gamma u, \ldots)\, D^\alpha u \in L^1(\Omega),\ g_\alpha(\cdot, u, \ldots, D^\gamma u, \ldots)\in L^1(\Omega) \tag{2.13}$$

and there exists a subsequence (u_{j_k}) of (u_j) with the properties

$$\lim_{k\to\infty} D^\gamma u_{j_k} = D^\gamma u \text{ a. e. in } \Omega \text{ for } |\alpha|\leqq m-1, \tag{2.14}$$

$$\lim_{k\to\infty}\|g_{\alpha,j_k}(\cdot, u_{j_k} \ldots, D^\gamma u_{j_k}, \ldots)-g_\alpha(\cdot, u, \ldots, D^\gamma u, \ldots)\|_{L^1(\Omega)} = 0. \tag{2.15}$$

PROOF. As (u_j) tends to u weakly in V thus there exists a subsequence (u_{j_k}) of (u_j) with the property (2.14) (see [6]). Let for $x\in\Omega$

$$\lim_{k\to\infty} D^\gamma u_{j_k}(x) = D^\gamma u(x) \text{ if } |\alpha|\leqq m-1.$$

We shall show that

$$\lim_{k\to\infty} g_{\alpha,j_k}(x, u_{j_k}(x), \ldots, D^\gamma u_{j_k}(x), \ldots) = g_\alpha(x, u(x), \ldots, D^\gamma u(x), \ldots) \tag{2.16}$$

and

$$\lim_{k\to\infty}\chi_{j_k}(x)p_{\alpha,j_k}(x, u_{j_k}(x), \ldots, D^\gamma u_{j_k}(x), \ldots) = \tag{2.17}$$

$$= p_\alpha(x, u(x), \ldots, D^\gamma u(x), \ldots).$$

Indeed,

$$\begin{aligned}|g_{\alpha,j_k}(x, u_{j_k}(x), \ldots, D^\gamma u_{j_k}(x), \ldots)-g_\alpha(x, u(x), \ldots, D^\gamma u(x), \ldots)| &\leqq \\ \leqq |g_{\alpha,j_k}(x, u_{j_k}(x), \ldots, D^\gamma u_{j_k}(x), \ldots)-g_\alpha(x, u_{j_k}(x), \ldots, D^\gamma u_{j_k}(x), \ldots)| &+ \\ + |g_\alpha(x, u_{j_k}(x), \ldots, D^\gamma u_{j_k}(x), \ldots)-g_\alpha(x, u(x), \ldots, D^\gamma u(x), \ldots)|.&\end{aligned} \tag{2.18}$$

Consider the neighbourhood B_ε of the point $(u(x), \ldots, D^\gamma u(x), \ldots)\in\mathbf{R}^M$ with radius ε. By virtue of the assumption f) and (2.2) there is j_0 such that

$$g_{\alpha,j_k}(x, \xi') = g_\alpha(x, u(x), \ldots, D^\gamma u(x), \ldots) \text{ if } \xi'\in B_\varepsilon,\ j_k\geqq j_0.$$

Thus for sufficiently large k the first term in the right of (2.18) equals to 0. The second term in the right of (2.18) converges to 0 as $k\to\infty$ by the assumption f). Therefore (2.18) implies (2.16). The relation (2.17) can be similarly proved.

By (2.2), (2.3), (2.12) and the assumption g)

$$\sum_{|\alpha|\leq l}\int_{\Omega}\chi_j(x)p_{\alpha,j}(x,u_j,\ldots,D^\gamma u_j,\ldots)D^\alpha u_j dx\leq c+\sum_{|\alpha|\leq l}\|h_\alpha\|_{L^q(\Omega)}\|u_j\|_V$$

thus (2.17) and Fatou's lemma implies

$$p_\alpha(\cdot,u,\ldots,D^\gamma u,\ldots)\,D^\alpha u\in L^1(\Omega).$$

Therefore in virtue of the assumption g) we have the first part of (2.13).

Now we shall prove the second part of (2.13) and (2.15), using the Vitali convergence theorem. For any fixed number $\delta>0$

$$|g_{\alpha,\mu}(x,\xi')|\leq\sup_{\substack{|\xi_\alpha|<\delta^{-1}\\ \alpha\in A_\nu}}|g_{\alpha,\mu}(x,\xi')|+\sum_{\alpha\in A_\nu}\delta|\xi_\alpha g_{\alpha,\mu}(x,\xi')|. \tag{2.19}$$

Since by (2.2), (2.3) and the assumption g)

$$|g_{\alpha,\mu}(x,\xi')|\leq|p_\alpha(x,\xi')|+h_\alpha(x)\leq|g_\alpha(x,\xi')|+2h_\alpha(x),$$

from (2.19) we obtain

$$|g_{\alpha,\mu}(x,\xi')|\leq g_{\nu,\delta^{-1}}(x,\xi'')+2h_\alpha(x)+\sum_{\alpha\in A_\nu}\delta|\xi_\alpha g_{\alpha,\mu}(x,\xi')| \tag{2.20}$$

(the definition of $g_{\nu,\delta^{-1}}$ see in the assumption h)).

(2.2), (2.3) and the assumption g) implies the estimation

$$|g_{\alpha,\mu}(x,\xi')\xi_\alpha|\leq g_{\alpha,\mu}(x,\xi')\xi_\alpha+2h_\alpha(x)|\xi_\alpha|.$$

Consequently, by (2.20) and the assumption h)

$$\begin{aligned}&|g_{\alpha,j_k}(x,u_{j_k}(x),\ldots,D^\gamma u_{j_k}(x),\ldots)|\leq\\ &\leq g^*_{\nu,\delta^{-1}}(x)+2h_\alpha(x)+2\delta\sum_{\alpha\in A_\nu}h_\alpha(x)|D^\alpha u_{j_k}(x)|+\\ &+\delta\sum_{\alpha\in A_\nu}g_{\alpha,j_k}(x,u_{j_k}(x),\ldots,D^\gamma u_{j_k}(x),\ldots)D^\alpha u_{j_k}(x).\end{aligned}$$

Hence by (2.12), the assumption g) and Hölder's inequality there exists a constant c' such that for any measurable set $E\subset\Omega$

$$\int_E|g_{\alpha,j_k}(x,u_{j_k},\ldots,D^\gamma u_{j_k},\ldots)|dx\leq\int_E[g^*_{\nu,\delta^{-1}}(x)+2h_\alpha(x)]dx+c'\delta. \tag{2.21}$$

Let $\varepsilon>0$ an arbitrary number and set $\delta=\dfrac{\varepsilon}{2c'}$. Then by (2.21) for sufficiently small meas E

$$\int_E|g_{\alpha,j_k}(x,u_{j_k},\ldots,D^\gamma u_{j_k},\ldots)|dx<\varepsilon$$

and there exists a set $A_\varepsilon \subset \Omega$ of finite measure such that

$$\int_{\Omega/A_\varepsilon} |g_{\alpha,j_k}(x, u_{j_k}, \ldots, D^\gamma u_{j_k}, \ldots)| dx < \varepsilon.$$

Thus (2.16) and Vitali's theorem imply the second part of (2.13) and (2.15) which completes the proof of Lemma 6.

LEMMA 7. For all $u \in V$

$$\langle (T+S_\mu)(u), u\rangle \geqq c_3 \|u\|_V^p - c_4 - \Big(\sum_{|\alpha|\leqq l} \|h_\alpha\|_{L^q(\Omega)}\Big)\|u\|_V \tag{2.22}$$

where c_3, c_4 are the constants of (1.4). Thus $T+S_\mu$ is coercive, i.e.

$$\lim_{\|u\|\to\infty} \frac{\langle (T+S_\mu)(u), u\rangle}{\|u\|} = +\infty.$$

PROOF. According to (1.4)

$$\langle T(u), u\rangle \leqq c_3 \|u\|_V^p - c_4. \tag{2.23}$$

Further from (2.2), (2.3), (2.5) and the assumption g) by Hölder's inequality we find

$$\langle S_\mu(u), u\rangle \geqq -\sum_{|\alpha|\leqq l} \|h\|_{L^q(\Omega)} |D^\alpha u\|_{L^p(\Omega)}. \tag{2.24}$$

The estimations (2.23), (2.24) imply (2.22). Finally, from (2.22) obviously follows that $T+S_\mu$ is coercive.

The PROOF of the THEOREM. By the Lemmas 1, 4, 5 and 7 the operator $T+S_j : V \to V^*$ is bounded, continuous in finite dimension, pseudomonotone and coercive for all $j = 1, 2, \ldots$ Using the well known theory of pseudomonotone operators in reflexive Banach spaces, we obtain that for any $F \in V^*$ there exists $u_j \in V$ such that

$$(T+S_j)(u_j) = F. \tag{2.25}$$

By Lemma 7 the sequence (u_j) is bounded in V. T is a bounded operator (Lemma 1) and so the sequence $(T(u_j))$ in bounded in V^*. Since V is a reflexive Banach space, there exist a subsequence (u_{j_k}) of (u_j'), $u \in V$ and $y \in V^*$ such that

$$\begin{cases} \lim\limits_{k\to\infty} (u_{j_k}) = u \text{ weakly in } V, \\ \lim\limits_{k\to\infty} T(u_{j_k}) = y \text{ weakly in } V^*. \end{cases} \tag{2.26}$$

Combining the definition of S_j with (2.25) we find that

$$\sum_{|\alpha|\leqq l} \int_\Omega g_{\alpha,j_k}(x, u_{j_k}, \ldots, D^\gamma u_{j_k}, \ldots) D^\alpha u_{j_k} dx =$$
$$= \langle S_{j_k}(u_{j_k}), u_{j_k}\rangle = \langle F, u_{j_k}\rangle - \langle T(u_{j_k}), u_{j_k}\rangle \leqq$$
$$\leqq \|F\|_{V^*} \|u_{j_k}\|_V + \|T(u_{j_k})\|_{V^*} \|u_{j_k}\|_V \leqq c.$$

Thus by Lemma 6 for all α with $|\alpha| \leqq l$

(2.27) $g_\alpha(\cdot, u, \ldots, D^\gamma u, \ldots) D^\alpha u \in L^1(\Omega), \quad g_\alpha(\cdot, u, \ldots, D^\gamma u, \ldots) \in L^1(\Omega)$

and there is a subsequence $(u_{j_k'})$ of (u_{j_k}) such that

$$\lim_{k \to \infty} D^\gamma u_{j_k'} = D^\gamma u \text{ a. e. in } \Omega \text{ for } |\alpha| \leqq m-1, \tag{2.28}$$

$$\lim_{k \to \infty} \|g_{\alpha, j_k'}(\cdot, u_{j_k'}, \ldots, D^\gamma u_{j_k'}, \ldots) - g_\alpha(\cdot, u, \ldots, D^\gamma u, \ldots)\|_{L^1(\Omega)} = 0. \tag{2.29}$$

From (2.25) it follows that for all $v \in V$ with the property $D^\alpha v \in L^\infty(\Omega)$ $(|\alpha| \leqq l)$

$$\langle (T + S_{j_k'})(u_{j_k'}), v \rangle = \langle F, v \rangle,$$

whence by making use of (2.26), (2.29) as $k \to \infty$ we find

$$\langle y, v \rangle + \sum_{|\alpha| \leqq l} \int_\Omega g_\alpha(x, u, \ldots, D^\gamma u, \ldots) D^\alpha v dx = \langle F, v \rangle. \tag{2.30}$$

Now we shall show that $y = T(u)$. Since T is pseudomonotone, it is sufficient to prove the inequality

$$\limsup_{k \to \infty} \langle T(u_{j_k'}), u_{j_k'} - u \rangle \leqq 0. \tag{2.31}$$

By (2.25)

$$\langle T(u_{j_k'}), u_{j_k'} - u \rangle = \langle F, u_{j_k'} \rangle - \langle S_{j_k'}(u_{j_k'}), u_{j_k'} \rangle - \langle T(u_{j_k'}), u \rangle$$

thus (2.26) implies that

$$\limsup_{k \to \infty} \langle T(u_{j_k'}), u_{j_k'} - u \rangle = \langle F - y, u \rangle - \liminf_{k \to \infty} \langle S_{j_k'}(u_{j_k'}), u_{j_k'} \rangle. \tag{2.32}$$

By making use of (2.2) the expression $\langle S_{j_k'}(u_{j_k'}), u_{j_k'} \rangle$ in the right of (2.32) can be written in the form

$$\langle S_{j_k'}(u_{j_k'}), u_{j_k'} \rangle = \tag{2.33}$$

$$= \sum_{|\alpha| \leqq l} \int_\Omega \chi_{j_k'} p_{\alpha, j_k'}(x, u_{j_k'}, \ldots, D^\gamma u_{j_k'}, \ldots) D^\alpha u_{j_k'} dx +$$

$$- \sum_{|\alpha| \leqq l} \int_\Omega r_\alpha(x, u_{j_k'}, \ldots, D^\gamma u_{j_k'}, \ldots) D^\alpha u_{j_k'} dx.$$

By (1.7), (2.3), (2.28) and the assumption f) Fatou's lemma yields that

$$\liminf_{k \to \infty} \sum_{|\alpha| \leqq l} \int_\Omega \chi_{j_k'} p_{\alpha, j_k'}(x, u_{j_k'}, \ldots, D^\gamma u_{j_k'}, \ldots) D^\alpha u_{j_k'} dx \geqq \tag{2.34}$$

$$\geqq \sum_{|\alpha| \leqq l} \int_\Omega p_\alpha(x, u, \ldots, D^\gamma u, \ldots) D^\alpha u dx.$$

Further by (1.8), (2.28), the assumption f) and the boundedness of $\|u_{j_k'}\|_V$ Hölder's inequality and Vitali's convergence theorem imply

$$\lim_{k\to\infty} \sum_{|\alpha|\leqq l} \int_\Omega r_\alpha(x, u_{j_k'}, \ldots, D^\nu u_{j_k'}, \ldots) D^\alpha u_{j_k'} dx =$$

$$= \sum_{|\alpha|\leqq l} \int_\Omega r_\alpha(x, u, \ldots, D^\nu u, \ldots) D^\alpha u dx.$$

Thus from (2.33), (2.34) we obtain

$$\liminf_{k\to\infty} \langle S_{j_k'}(u_{j_k'}), u_{j_k'}\rangle \geqq \sum_{|\alpha|\leqq l} \int_\Omega g_\alpha(x, u, \ldots, D^\nu u, \ldots) D^\alpha u dx$$

and so

$$\limsup_{k\to\infty} \langle T(u_{j_k'}), u_{j_k'} - u\rangle \leqq \langle F - y, u\rangle - \tag{2.35}$$

$$- \sum_{|\alpha|\leqq l} \int_\Omega g_\alpha(x, u, \ldots, D^\nu u, \ldots) D^\alpha u dx.$$

In virtue of the assumption a) there exist constants $c>0$, $c'>0$ and a sequence of functions $w_j \in V \subset L^\infty(\Omega)$ such that (1.1) and (1.2) hold. By use of (2.30) and (2.35) we find the inequality

$$\limsup_{k\to\infty} \langle T(u_{j_k'}), u_{j_k'} - u\rangle \leqq \langle F - y, u - w_j\rangle + \tag{2.36}$$

$$+ \sum_{|\alpha|\leqq l} \int_\Omega g_\alpha(x, u, \ldots, D^\nu u, \ldots)(D^\alpha w_j - D^\alpha u) dx.$$

From (1.1) we have

$$\lim_{j\to\infty} \langle F - y, u - w_j\rangle = 0$$

further as for a subsequence (w_j') of (w_j)

$$\lim_{j\to\infty} D^\alpha w_j' = D^\alpha u \text{ a. e. in } \Omega \text{ if } |\alpha| \leqq m-1$$

(see [6]) thus (1.2), (2.27) and Lebesgue's dominated convergence theorem imply

$$\lim_{j\to\infty} \sum_{|\alpha|\leqq l} \int_\Omega g_\alpha(x, u, \ldots, D^\nu u, \ldots)(D^\alpha w_j' - D^\alpha u) dx = 0.$$

Therefore from (2.36) we obtain (2.31), consequently, $y = T(u)$ and by (2.30) the equality (1.11) is valid for all $v \in V$ with $\partial^\alpha v \in L^\infty(\Omega)$ (for $|\alpha| \leqq l$).

Since (1.11) is true for $v = w_j'$ thus letting $j \to \infty$ we find that (1.11) is true also for $v = u$. The proof of the existence theorem is complete.

References

[1] WEBB, J. R. L.: Boundary value problems for strongly nonlinear elliptic equations, *J. London Math. Soc.* (2), **21** (1980), 123–132.

[2] MICHAEL, F. H.: An application of the method of monotone operators to nonlinear elliptic boundary value problems in unbounded domains, *Annales Univ. Sci. Budapest*, Sectio Mathematica, **25** (1982), 69–84.

[3] MICHAEL, F. H.: An elliptic boundary value problem for nonlinear equations in unbounded domains, *Annales Univ. Sci. Budapest*, Sectio Mathematica **26** (1983), 125–139.

[4] ШИМОН Л.: Аппроксимация решения задачи Дирихле для нелинейного эллиптического уравнения в неограниченной области, *Доклады Акад. Наук СССР*, **287** (1986), 1334–1337.

[5] ДУБИНСКИЙ, Ю. А.: Нелинейные эллиптические и параболические уравнения, *Современные проблемы математики*, том 9, Москва, 1976.

[6] EDMUNS, D. E. – WEBB, J. R. L.: Quasilinear elliptic problems in unbounded domains, *Proc. Roy. Soc. London*, Ser. A, **337** (1973), 397–410.

ON THE SUMMABILITY OF EIGENFUNCTION EXPANSIONS III.

By

I. JOÓ

Institute for Analysis of the L. Eötvös University, Budapest

(Received September 19, 1983)

Let Ω be an arbitrary bounded domain in $\mathbf{R}^N (N>3)$ having C^∞ – smooth boundary and q be a function of the form

$$q(x) = \frac{a(|x-x_0|)}{|x-x_0|},$$

where for the function $a(t)$ the following conditions are fulfilled: there exists $\omega(t) \in C(0, \infty)$ such that

$$(1) \qquad \begin{cases} |a(t)| \leq \omega(t)\ (t>0)^*, \\ \omega(t) \text{ increases, } \omega(t)/t \text{ decreases on } (0, 1), \\ \forall\, \delta>0\, \exists\, t_0(\delta): t^\delta \omega\left(\dfrac{1}{t}\right) \text{ increases for } t \geq t_0(\delta), \\ \displaystyle\int_{+0} \frac{\omega(t)}{t}\, dt < \infty. \end{cases}$$

Denote by L an arbitrary positive selfadjoint extension of the Schrödinger operator $L_0 = L_0(x, D) = -\Delta + q(x)$ from the domain $C_0^\infty(\Omega)$ with discrete spectrum. E.g. if $q(x) \geq 0$, $(x \in \Omega)$, then according to a theorem of K. O. FRIEDRICHS [2], there exists such a selfadjoint extension. Denote $L = \int_m^\infty \lambda dE_\lambda$ the spectral expansion of L and for any $f \in L_2(\Omega)$ consider the expansion $E_\lambda f$.

The aim of the present paper is to prove the following

---

* E. g. $\omega(t) = 1/\log^2 t$.

THEOREM. *Suppose the above conditions are fulfilled. Then for any* $f\in C_0^\infty(\Omega)$ *the expansion* $E_\lambda f(x)$ *tends to* $f(x)$ *(as* $\lambda\to\infty$*), uniformly on every compact subset of* Ω.

For the proof of the Theorem we need some lemmas.

Consider the functions

$$\nu(t)\overset{\text{def}}{=}\omega\left(\frac{1}{\sqrt{t}}\right),\qquad \varphi(\lambda)\overset{\text{def}}{=}\int_0^\infty\frac{t^{-\tau}\nu(t)}{\lambda+t}\,dt,\qquad \psi(t)\overset{\text{def}}{=}t^{-\tau}\nu(t),$$

$$(0<\tau<1).$$

Then by Lemma 1.1 of [14] we have

$$\varphi(\lambda)=\lambda^{-\tau}\nu(\lambda)[c_\tau+o(1)],$$

hence

$$c_1\lambda^{-\tau}\nu(\lambda)\le\varphi(\lambda)(\lambda>0). \tag{2}$$

Consider the operator

$$\varphi(L)\overset{\text{def}}{=}\int_m^\infty\varphi(\lambda)dE_\lambda$$

Obviously

$$\varphi(L)=\int_m^\infty\left(\int_0^\infty\frac{\psi(t)}{t+\lambda}\,dt\right)dE_\lambda=\int_0^\infty\psi(t)R_{-t}(L)dt,$$

where

$$R_t(L)\overset{\text{def}}{=}(L-tI)^{-1}.$$

Denote $\Phi(x,y)$ the kernel function of the operator $\varphi(L)$, i.e.

$$\Phi(L)f(x)=\int_\Omega\Phi(x,y)f(y)dy,\ (f\in D(\varphi(L)),\ x\in\Omega).$$

Define

$$z_0\overset{\text{def}}{=}\left\{z\in\mathbf{C}:\left|\arg z-\frac{\pi}{2}\right|\le\frac{\pi}{2}-\varepsilon,\ \varepsilon>0\ \text{fixed}\right\},$$

$$B=\{x\in\mathbf{R}^n:|x-x_0|<R\},$$

$$B_1=\left\{x\in\mathbf{R}^n:|x-x_0|<\frac{2}{3}R\right\},$$

$$B_0=\left\{x\in\mathbf{R}^n:|x-x_0|<\frac{1}{3}R\right\},\ r=|x-y|,$$

$$0<R<\operatorname{dist}(x_0,\partial\Omega)\overset{\text{def}}{=}R^*,$$

denote $E(x,y,\mu)$ the exponentially decreasing (for $\operatorname{Im}\mu>0$) fundamental solution of the operator $L_\mu=L-\mu^2 I$ (this was constructed and estimated in

[7]), $G_\mu = G(x, y, \mu)$ the Green function of L_μ, i. e. the kernel function of the operator $(L-\mu^2 I)^{-1}$. At last define

$$\Phi_0(x, y) \overset{\text{def}}{=} \int_0^\infty \psi(t) H^*(x, y, i\sqrt{t})dt,$$

where $\eta \in C_0^\infty(B)$ is such that $\eta(x) = 1$ for $x \in B_1$ and

$$H(x, y, \mu) \overset{\text{def}}{=} \xi(x)E(x, y, \mu),$$

$$H^*(x, y, \mu) \overset{\text{def}}{=} \xi(y)E(y, x, \mu) = \xi(y)E^*(x, y, \mu).$$

LEMMA 1. For $x \in B_0$ and $y \in \Omega$

$$|\Phi_0(x, y)| \leq C_\tau |x-y|^{2\tau-N} \nu(|x-y|^{-2}), \quad (0<\tau<1), \tag{3}$$

PROOF. In [7] is proved the estimate

$$|E(x, y, \mu)| \leq c|x-y|^{2-N} e^{-|\operatorname{Im}\mu|\,|x-y|}, \quad (x, y \in B;\ \mu \in \mathbf{C}), \tag{4}$$

hence we get

$$|\Phi_0(x, y)| = \left|\int_0^\infty \psi(t) H^*(x, y, i\sqrt{t}dt\right| \leq c|x-y|^{2-N} \int_0^\infty t^{-\tau} \nu(t) e^{-\sqrt{t}|x-y|} dt. \tag{5}$$

Using the substitution $s = r\sqrt{t}$ and the notation

$$I = I(1/r^2) = \int_0^\infty s^{1-2\tau} \nu\left(\frac{s^2}{r^2}\right) e^{-s} ds,$$

we obtain

$$|\Phi_0(x, y)| \leq c r^{2\tau-N} I.$$

We must estimate I. To this first consider the function

$$g_\lambda(s) \overset{\text{def}}{=} s^{1-2\tau} e^{-s} \left[\frac{\nu(\lambda s^2)}{\nu(\lambda)} - 1\right].$$

Now we prove the existence of a function $g(s)$ such that

$$|g_\lambda(s)| \leq g(s), \qquad \int_0^\infty g(s) ds < \infty. \tag{6}$$

Let $\delta > 0$ be arbitrary. Taking into consideration (1), $\lambda^\delta \nu(\lambda)$ increases for for $\lambda \geq t_0 = t(\delta)$. Hence

1. if $s^2 \geq 1$, then $\nu(\lambda s^2)/\nu(\lambda) \leq 1$;

2. if $\dfrac{t_0}{\lambda} \leq s^2 \leq 1$, then

$$\nu(\lambda s^2) = (\lambda s^2)^\delta \frac{\nu(\lambda s^2)}{(\lambda s^2)^\delta} \leq \frac{\lambda^\delta \nu(\lambda)}{(\lambda s^2)^\delta} = \frac{\nu(\lambda)}{s^{2\delta}},$$

i. e.

$$\frac{\nu(\lambda s^2)}{\nu(\lambda)} \leq s^{-2\delta};$$

3. if $s^2 \leq \frac{t_0}{\lambda}$, then for $\lambda \geq t_0$ we have

$$t_0^\delta \nu(t_0) \leq \lambda^\delta \nu(\lambda),$$

i.e.

$$\nu(\lambda) \geq c\lambda^{-\delta},$$

consequently

$$\frac{\nu(\lambda s^2)}{\nu(\lambda)} \leq \frac{\nu(o)}{\nu(\lambda)} \leq c\lambda^\delta \leq c(s^{-2}t_0)^\delta \leq cs^{-2\delta}.$$

From 1., 2. and 3. we obtain:

$$|g_\lambda(s)| \leq cs^{1-2\tau}e^{-s}(1+s^{-2\delta}).$$

Because $0<\tau<1$, we can choose $\delta<1-\tau$, and the desired estimate (6) follows.

Now we estimate I. Obviously

$$I(\lambda) = \nu(\lambda)\left\{\int_0^\infty s^{1-2\tau}e^{-s}ds + \int_0^\infty s^{1-2\tau}e^{-s}\left[\frac{\nu(\lambda s^2)}{\nu(\lambda)} - 1\right]ds\right\} =$$

$$= c_\tau \nu(\lambda) + \nu(\lambda)\int_0^\infty g_\lambda(s)ds.$$

It is easy to see that for every $s \in (0, \infty)$ $g_\lambda(s) \to 0$ as $\lambda \to \infty$ and by Lebesgue's dominated convergence theorem

$$\int_0^\infty g_\lambda(s)ds \to 0 \quad (\lambda \to \infty).$$

We obtain for $\lambda = 1/r^2$: $I \leq c_\tau\, \nu(1/r^2)$.

Lemma 1 is proved.

LEMMA 2. For arbitrary compact set $K \subset \Omega$ we have

$$\int_\Omega |\Phi_0(x,y)|^q dy \leq c(K) < \infty, \quad \left(x \in K,\ q = \frac{N}{N-2\tau}\right). \tag{7}$$

PROOF. Using polar coordinates and applying (3) we get

$$\int_\Omega |\Phi_0(x,y)|^q dy \leq c\int_{+0} r^{-N}r^{N-1}\nu^q\left(\frac{1}{r^2}\right)dr =$$

$$= c\int_{+0} \nu^q\left(\frac{1}{r^2}\right)\frac{dr}{r} = c\int_{+0} \omega^q(r)\frac{dr}{r} < \infty.$$

Lemma 2 is proved.

Define

$$Mf(x) \overset{\text{def}}{=} \int_0^\infty \psi(t)\hat{K}^*_{i\sqrt{t}}\,\hat{G}_{i\sqrt{t}}f(x)dt,$$

where

$$K^*_\mu(x,y) \overset{\text{def}}{=} 2\nabla_y\xi(y))\cdot\nabla_y E^*(x,y,\mu)+(\Delta_y\xi(y))E^*(x,y,\mu).$$

LEMMA 3. For any $f\in L_2(\Omega)$ and $x\in B_0$

$$|Mf(x)|\le c\|f\|_{L_2}. \tag{8}$$

PROOF. In [7] we proved the estimates

$$|G_\mu(x,y)|\le c_1|x-y|^{2-N}e^{-c_2|\mu|\,|x-y|},\ (\mu\in Z_0;\ x,y\in B_0), \tag{9}$$

$$|\nabla_x E(x,y,\mu)|\le c(R)e^{-|\operatorname{Im}\mu|\,|x-y|},\ \left(\frac{3}{4}R\le|x|<R,\ |y|\le\frac{1}{4}R,\ \mu\in\mathbf{C}\right). \tag{10}$$

Consequently

$$|Mf(x)|\le c_3\int_0^\infty\psi(t)e^{-c_4t}\|\hat{G}_{i\sqrt{t}}f\|_{L_2}dt\le c_5\|f\|_{L_2}\int_0^\infty\psi(t)e^{-c_6t}dt\le c\|f\|_{L_2}.$$

Lemma 3 is proved.

LEMMA 4. For any $f\in L_p(\Omega)$ $(p=N/2\tau)$ and $x\in B_0$

$$|\varphi(L)f(x)|\le c\|f\|_{L_p}. \tag{11}$$

PROOF. It is proved in [7] that for any $f\in L_2(\Omega)$ and $x\in B_0$

$$\hat{G}_\mu f(x)-\hat{H}^*f(x)=\hat{K}^*\hat{G}_\mu f(x),\ (\mu\in\mathbf{C}\backslash\{\lambda_n\}), \tag{12}$$

hence

$$[\varphi(L)f](x)=\hat{\Phi}_0f(x)+Mf(x),$$

and taking into consideration (7), (8) the desired estimate (12) follows.
Lemma 4 is proved.

LEMMA 5. For any $f\in L_2(\Omega)$

$$\|\varphi(L)L^{-\sigma}f\|_{L_\infty(B_0)}\le c\|f\|_{L_2(\Omega)}, \tag{13}$$

$$\left(\sigma=\frac{N}{4}-\tau,\ \tau\ge\frac{1}{2}\left(\frac{N}{2}-\left[\frac{N}{2}\right]\right)\right).$$

PROOF. Using Lemma 4 and the imbedding $W_2^{2\sigma}\to L_p$, $\left(\frac{N}{2}-2\sigma=\frac{N}{p}=\right.$ $\left.=2\tau\right)$ we obtain

$$\|\varphi(L)L^{-\sigma}f\|_{L_\infty(B_0)}\le c\|L^{-\sigma}f\|_{L_p(B_0)}\le c\|L^{-\sigma}f\|_{W_2^{2\sigma}}\le c\|f\|_{L_2},\ (cf[8]).$$

Lemma 5 is proved.

COROLLARY. For any compact set $K\subset\Omega$ there exists a constant $C(K)$, such that

$$\sum_{k=1}^{N}\frac{|v_k(x)|^2}{\lambda_k^{N/2}}\nu^2(\lambda_k)\leq C(K),\ (x\in K,\ N\geq 1). \tag{14}$$

Indeed, using (2) and (13) we obtain for any $x\in E_0$

$$\sum_{k=1}^{N}\frac{|u_k(x)|^2}{\lambda_k^{N/2}}\nu^2(\lambda_k)\leq c\sum_{k=1}^{N}\frac{|u_k(x)|^2}{\lambda_k^{N/2}}\varphi^2(\lambda_k)\lambda_k^{2\tau}=$$

$$=c\sum_{k=1}^{N}\frac{|u_k(x)|^2}{\lambda_k^{\left(\frac{N}{4}-\tau\right)\cdot 2}}\varphi^2(\lambda_k)=c\sup_{\substack{f\in L_2(\Omega)\\ \|f\|\leq 1}}\left|\sum_{k=1}^{N}\frac{(f,u_k)u_k(x)}{\lambda_k^{\frac{N}{4}-\tau}}\varphi(\lambda_k)\right|\leq$$

$$\leq c\sup_{\substack{f\in L_2(\Omega)\\ \|f\|\leq 1}}\sup_{x\in B_0}|\ \ |=c\sup_{\substack{f\in L_2(\Omega)\\ \|f\|\leq 1}}\|\varphi(L)L^{-\sigma}f\|_{L_\infty(B_0)}<\infty.$$

The Corollary is proved.

LEMMA 6. Suppose $f\in C_0^\infty(\Omega)$, $x_0=0$ and $f(0)=0$. Then

$$\|L^\sigma f\|_{L_2(\Omega)}<\infty,\ \left(\sigma=\frac{1}{2}\left(\frac{N}{2}+1\right)\right). \tag{15}$$

PROOF. Let $\dfrac{N+2}{4}=m+\delta$, $0<\delta\leq 1$. It is proved in [8] that

$$L^m f(x)=(-\triangle)^m f(x)+\sum_\alpha c_\alpha^{(x)}D^\alpha f(x)$$

and

$$|\nabla L^m f(x)|\leq c\omega(|x|)\sum_\alpha|x|^{|\alpha|-2m-1}|D^\alpha f(x)|\leq c\omega(|x|)|x|^{-2m},$$

because

$$|D^\alpha f(x)|\leq\text{const},\ \text{if}\ |\alpha|>0,$$

$$|D^\alpha f(x)|\leq\text{const}\cdot|x|,\ \text{if}\ |\alpha|=0.$$

We obtain

$$\|\nabla L^m f(x)\|_{L_p}<\infty,\ \text{if}\ 2mp=N,\ \text{i. e.}\ p=\frac{N}{2m}.$$

Taking into consideration the imbedding $W_p^1\to W_2^{2\delta}$,

$$\left(\frac{N}{p}-1=\frac{N}{2}-2\delta,\ \text{i. e.}\ 2\delta=\frac{N}{2}+1-\frac{N}{p}=\frac{N}{2}+1-2m\right)$$

it follows

$$\|L^m f\|_{W_p^1}<\infty\ \text{if}\ \frac{N}{p}=2m,$$

hence

$$\|L^{m+\delta}f\|_{L_2} = \|L^{\delta}L^m f\|_{L_2} \leq c\|L^m f\|_{W_2^{2\delta}} \leq c\|L^m f\|_{W_p^1} < \infty .$$

Lemma 6 is proved.

COROLLARY. For any $f \in C_0^\infty(\Omega)$ with $f(0) = 0$, the spectral expansion $E_\lambda f(x)$ tends to $f(x)$ absolutely and uniformly on every compact set $K \subset \Omega$.

PROOF. It is enough to prove the Corollary for $K = \overline{B}_0$. Using (14) and (15) we obtain (taking into account the spectral theorem):

$$\sum_{k=n}^{n+p} |(f, u_k)u_k(x)| \leq \left(\sum_{k=n}^{n+p} |(f, u_k)|^2 \lambda_k^{2\sigma}\right)^{1/2} \left(\sum_{k=n}^{n+p} |u_k(x)|^2 \lambda_k^{-2\sigma}\right)^{1/2} \leq$$

$$\leq \varepsilon \left(\sum_{k=1}^{\infty} |u_k(x)|^2 \lambda_k^{-2\sigma}\right)^{1/2} = c(k)\varepsilon$$

if n and p is large enough.

The Corollary is proved.

LEMMA 7. Let $\chi(|x|) \in C_0^\infty(\Omega)$ such that $\chi(|x|) = 1$ for $|x| \leq \frac{R}{2}$ and $\chi(|x|) = 0$ if $|x| \geq R$. Then we have

$$|\chi, u_n)| \leq c \frac{|u_n(0)|}{\lambda_n^{N/2}} \omega\left(\frac{1}{\sqrt{\lambda_n}}\right), \quad (n = 1, 2, \ldots). \tag{16}$$

PROOF. First consider the case $N \equiv 0 \pmod 2$ and use the notation $m = N/2$.

Obviously

$$(\chi, u_n) = \frac{1}{\lambda_n^m}(\chi, L^m u_n) = \frac{1}{\lambda_n^m}(L^m \chi, u_n).$$

On the other hand (cf. [8])

$$|L^m \chi(x)| \leq c \frac{\omega^m(|x|)}{|x|^{2m}}.$$

Taking into account that q is spherically symmetrical, it is enough to estimate the integral

$$I = \int_0^R \chi(r) \frac{\omega^m(r)}{r^N} \left| \int_\Theta u_n(0 + r\Theta) d\Theta \right| r^{N-1} dr.$$

We proved in [8] the following "generalized" Titchmarsh formula

$$\int_\Theta u_n(0 + r\Theta) d\Theta = u_n(0) \left[c_N \frac{J_p(r\sqrt{\lambda_n})}{(r\sqrt{\lambda_n})^p} + \alpha(r, \sqrt{\lambda_n}) \right],$$

where

$$p=\frac{N}{2}-1, \quad |\alpha(r,\sqrt{\lambda_n})|\leq cb(r)h(r\sqrt{\lambda_n}),$$

$$b(r)\overset{\text{def}}{=}\int_0^r\frac{\omega(t)}{t}dt, \quad h(t)\overset{\text{def}}{=}\min\{1,t^{-p-\frac{1}{2}}\}, \ (t>0).$$

We obtain

$$I\leq c|u_n(0)|\left\{\int_0^R\chi(r)\frac{\omega^m(r)}{r}\left|\frac{J_p(r\sqrt{\lambda_n})}{(r\sqrt{\lambda_n})^p}\right|dr+\right.$$

$$\left.+\int_0^R\chi(r)\frac{\omega^m(r)}{r}b(r)h(r\sqrt{\lambda_n})dr\right\}=c|u_n(0)|\{I_1+I_2\}.$$

For the estimation of I_1 we prove

$$I_1=\int_0^R\left|\frac{J_p(r\sqrt{\lambda_n})}{(r\sqrt{\lambda_n})^p}\right|\frac{\omega^m(r)}{r}dr\leq c\omega\left(\frac{1}{\sqrt{\lambda_n}}\right).$$

For $N\geq 3$ the well known estimates

$$|J_p(x)|\leq\frac{c}{\sqrt{x}}\ (x\geq\delta>0), \quad \left|\frac{J_p(x)}{x^p}\right|\leq c, \ (0<x\leq\delta)$$

hold. We have

$$I_1'\overset{\text{def}}{=}\int_0^{1/\sqrt{\lambda_n}}\leq\int_0^{1/\sqrt{\lambda_n}}\frac{\omega^2(r)}{r}dr\leq c\omega\left(\frac{1}{\sqrt{\lambda_n}}\right),$$

and

$$I_1''\overset{\text{def}}{=}\int_{1/\sqrt{\lambda_n}}^R\leq\int_{1/\sqrt{\lambda_n}}^R\frac{1}{r\sqrt{\lambda_n}}\frac{\omega^2(r)}{r}dr\leq\frac{c}{\sqrt{\lambda_n}}\int_{1/\sqrt{\lambda_n}}^R\frac{\omega(r)}{r}\frac{\omega(r)}{r}dr\leq$$

$$\leq\frac{c}{\sqrt{\lambda_n}}\int_{1/\sqrt{\lambda_n}}^R\frac{\omega(1/\sqrt{\lambda_n})}{1/\sqrt{\lambda_n}}\frac{\omega(r)}{r}\leq c\omega\left(\frac{1}{\sqrt{\lambda_n}}\right).$$

Now estimate I_2 in a similar way. Denote

$$I_2=\int_0^R=\int_0^{1/\sqrt{\lambda_n}}+\int_{1/\sqrt{\lambda_n}}^R=I_2'+I_2''.$$

Obviously

$$I_2'\leq\int_0^{1/\sqrt{\lambda_n}}\frac{\omega^2(r)}{r}dr\leq c\omega\left(\frac{1}{\sqrt{\lambda_n}}\right),$$

$$I_2''\leq c\int_{1/\sqrt{\lambda_n}}^R(r\sqrt{\lambda_n})^{-\frac{N}{2}-\frac{1}{2}}\frac{\omega^2(r)}{r}dr\leq c\int_{1/\sqrt{\lambda_n}}^R(r\sqrt{\lambda_n})^{-1}\frac{\omega^2(r)}{r}dr\leq c\omega\left(\frac{1}{\sqrt{\lambda_n}}\right).$$

Now consider the case $N \equiv 1 \pmod 2$. In this case define $m = (N-1)/2$ and repeat the argument above. We have in this case

$$|L_0^m \chi(x)| \leq c \cdot \frac{\omega^m(|x|)}{|x|^{N-1}},$$

further

$$I_1 = \int_0^R \left| \frac{J_p(r\sqrt{\lambda_n})}{(r\sqrt{\lambda_n})^p} \right| \omega^m(r)dr = \int_0^{1/\sqrt{\lambda_n}} + \int_{1/\sqrt{\lambda_n}}^R = I_1' + I_1'',$$

$$I_1' \leq \int_0^{1/\sqrt{\lambda_n}} \omega^m(r)dr \leq c\omega\left(\frac{1}{\sqrt{\lambda_n}}\right),$$

$$I_1'' \leq \int_{1/\sqrt{\lambda_n}}^R \frac{1}{r\sqrt{\lambda_n}} \omega(r) \leq \frac{c}{\sqrt{\lambda_n}} \frac{\omega\left(\frac{1}{\sqrt{\lambda_n}}\right)}{1/\sqrt{\lambda_n}} \int_0^R 1dr \leq c\omega\left(\frac{1}{\sqrt{\lambda_n}}\right),$$

$$I_2 \leq c \int_0^R \omega^m(r)h(r\sqrt{\lambda_n})dr = \int_0^{1/\sqrt{\lambda_n}} + \int_{1/\sqrt{\lambda_n}}^R = I_2' + I_2'',$$

$$I_2' \leq c \int_0^{1/\sqrt{\lambda_n}} \omega(r)dr \leq c\omega\left(\frac{1}{\sqrt{\lambda_n}}\right),$$

$$I_2'' \leq c \int_{1/\sqrt{\lambda_n}}^R \frac{1}{r\sqrt{\lambda_n}} \omega(r)dr \leq \frac{c}{\sqrt{\lambda_n}} \frac{\omega\left(\frac{1}{\sqrt{\lambda_n}}\right)}{1/\sqrt{\lambda_n}} \int_0^R 1dr \leq c\omega\left(\frac{1}{\sqrt{\lambda_n}}\right).$$

Summarising our estimates (16) follows. Lemma 7 is proved.

PROOF of the THEOREM. Let $f \in C_0^\infty(\Omega)$ be arbitrary. We may suppose $\operatorname{supp} f \subset \overline{B}_0$ (according to the Corollary after Lemma 6). Let χ be the function satisfying the requirements of Lemma 7. Obviously

$$f(x) = f(o)\chi(x) + \chi(x)(f(x) - f(o)),$$

and hence: it is enough to prove the uniform convergence of $E_\lambda \chi$ on $\overline{B}_0$. To this first remark the inequality

$$\sum_{n=N_1}^{N_2} |u_n(0)| \frac{\omega\left(\frac{1}{\sqrt{\lambda_n}}\right)}{\lambda_n^{N/2}} |u_n(x)| \leq \left(\sum_{n=N_1}^{N_2} \frac{u_n^2(o)}{\lambda_n^{N/2}} \omega\left(\frac{1}{\sqrt{\lambda_n}}\right)\right)^{1/2} \left(\sum_{n=1}^{\infty} \frac{u_n^2(x)}{\lambda_n^{N/2}} \omega\left(\frac{1}{\sqrt{\lambda_n}}\right)\right)^{1/2}$$

which is a corollary of (14) and take into account that for $\mu \geq 1$

$$\sum_{|\sqrt{\lambda_n} - \mu| \leq 1} u_n^2(o) \leq c\mu^{N-1}, \text{ (cf. [8])},$$

which implies

$$\sum_{n=1}^{\infty} \frac{u_n^2(o)}{\lambda_n^{N/2}} \omega\left(\frac{1}{\sqrt{\lambda_n}}\right) = \sum_{k=1}^{\infty} \sum_{k \leq \sqrt{\lambda_n} \leq k+1} {}_n \leq c \sum_{k=1}^{\infty} \omega \frac{1}{k}\left(\frac{1}{k}\right) \leq c \int_{+0} \frac{\omega(t)}{t} dt < \infty .$$

The Theorem is proved.

References

[1] G. Alexits, *Convergence problems of orthogonal series*, Akadémiai Kiadó, Budapest, 1961.

[2] F. Riesz et B. Sz.-Nagy, *Leçons d'analyse fonctionnelle*, Akadémiai Kiadó, Budapest, 1952.

[3] Ш. А. Алимов, Равномерная сходимость и суммируемость спектральных разложений, *Дифференц. Уравнения*, **9** (1973), 669–681.

[4] Ш. А. Алимов, О спектральных разложениях функции из H_p^α, *Матем. Сборник*, **101 (143)** (1976), 3–20.

[5] В. А. Ильин, О сходимости разложений по собственным функциям оператора Лапласа, *Успехи Матем. Наук*, **13** (1958), 87–180.

[6] Š. A. Alimov and I. Joó, On the Riesz summability of eigenfunction expansions, *Acta Sci. Math.*, **45** (1983), 5–18.

[7] I. Joó, Estimation of the Green function of the singular Schrödinger operator, *Acta Mat. Acad. Sci. Hung.*, **46** (1985), 275–284.

[8] I. Joó, On the summability of eigenfunction expansions II, *Annales Univ. Sci. Budapest*, Sect. Math., **27** (1984), 167–184.

[9] H. Bateman and A. Erdélyi, *Higher transcendental functions* 2, McGraw Hill Book Company (New York–Toronto–London, 1953).

[10] L. Hörmander, *On the Riesz means of spectral function and eigenfunction expansion for elliptic differential operators*, Lecture at the Belfer Graduate School, Yeshiva University, 1966.

[11] L. Hörmander, The spectral function of an elliptic operator, *Acta Math.* (Sweden), **121** (1968), 193–218.

[12] Б. М. Левитан, О разложении по собственным функциям уравнения $\Delta u + \{\lambda - q(x_1 \ldots x_m)\} n = 0$, *Изв. АН. СССР*, сер. Матем., **20** (1956), 437–468.

[13] С. М. Никольский, *Приближение функций многих переменных и теоремы вложения*, Наука (Москва, 1969).

[14] В. С. Серов, Обобщенные ядра дробного порядка, *Дифференц. Уравления*, **12** (1976), 1892–1902.

[15] E. C. Titchmarsh, *Eigenfunction expansions associated with second order differenctial equations*, Clarendon Press (Oxford, 1958).

[16] H. Triebel, *Interpolation theory – function spaces – differential operators*, VEB Deutscher Verlag der Wissenschaften (Berlin, 1978).

EINIGE BEMERKUNGEN BEZÜGLICH DER STRUKTUR VON ENDLICHEN BOLYAI – LOBATSCHEFSKY EBENEN

Von

F. KÁRTESZI und T. HORVÁTH

Lehrstuhl für Darstellende und Projektive Geometrie der L. Eötvös Universität und Student der L. Eötvös Universität, Budapest

(Eingegangen am 24. November 1982)

Man kann eine endliche Punktmenge mit Hilfe von ihren bestimmten Teilmengen zu einer Struktur, die Ebene genannt wird, organisieren. Die oben erwähnten Teilmengen werden Geraden genannt. Die Geraden in der klassischen hyperbolischen Geometrie haben die folgende Eigenschaft: Durch einen beliebigen Punkt gehen wenigstens drei Geraden, die mit einer nicht durch diesen Punkt verlaufende Gerade keinen gemeinsamen Punkt haben. Diese Eigenschaft gab die Idee, die Geraden der endlichen hyperbolischen Ebene zu definieren.

Es sei **S** eine endliche Punktmenge. Ihre bestimmten Teilmengen werden Geraden genannt, wenn die folgenden Erforderungen (Axiomen) erfüllt sind:

B1. *Zwei beliebige verschiedene Elemente von* **S** *werden von genau einer Gerade genannten Teilmenge von* **S** *enthält.*

B2. *Es gilt für jedes Paar, das aus einem beliebigen Punkt und aus einer nicht durch diesen Punkt verlaufenden Gerade von* **S** *besteht, daß m schneidende bzw. n nicht schneidende Geraden durch den Punkt eines Paares zu dem Geradenelement dieses Paares geben.*

B3. *Es gelten* $m>2$, $n>2$.

Mit der letzten Erforderung können wir die nichtssagenden Fällen ausschließen.

Wir nennen diese den vorigen drei kombinatorischen Erfordernissen entsprechende Struktur eine BOLYAI–LOBATSCHEFSKY Ebene (kurz **BL** Ebene) bei Charakter $\langle m, n\rangle$. Diese Struktur wird mit $\mathbf{S}\langle m, n\rangle$ bezeichnet.

Es ist offenbar, daß jede Gerade wegen **B1** und **B2** m Punkte enthält. Wenn wir die Zahl der Punkte bzw. der Gerade der Ebene mit ω bzw. λ bezeichnen, dann gilt

$$m\lambda = (m+n)\omega.$$

Auf jeder Gerade m Punkte gerechnet, haben wir nämlich jeden Punkt $(m+n)$-mal in Betracht gezogen. Nun bestimmen wir ω. Die Geraden, die durch einen bestimmten Punkt hindurchgehen, überdecken die Punkte der Ebene, mit Ausnahme des bestimmten Punktes, einfach. Deshalb gilt

$$\omega = (m+n)\,(m-1)+1,$$

und so bekommen wir für

$$\lambda = [(m+n)^2(m-1)+(m+n)]/m.$$

Die Tatsache, daß m der Teiler des Dividenden ist, ist nur die notwendige Bedingung der Existenz von $\mathbf{S}\langle m, n\rangle$. Die Nichtexistenz einiger Fälle ergibt sich sofort. Z. B. $\mathbf{S}\langle 4, 3\rangle$ existiert nicht, weil sich keine ganze Zahl für λ ($\lambda = 154/4 = 38{,}5$) ergibt.

Die Existenz der Ebene $\mathbf{S}\langle 3, 3\rangle$, die ein nichttrivialer Fall mit der kleinsten Elementenzahl ist, wurde von F. KÁRTESZI im Jahre 1969 mit Konstruktion eines entsprechenden Modells bewiesen. Ein zu dem vorigen *nicht isomorphes* Modell wurde von T. HORVÁTH konstruiert. Wir analisieren diese Modelle in dieser Arbeit, weiterhin untersuchen wir auch die **BL** Ebene, die den Fall $\mathbf{S}\langle 3, 4\rangle$ realisiert.

Die drei Beispiele, die wir im Rahmen dieser Arbeit vorlegen, decken interessante Eingenschaften auf.

1. In den Fällen der **BL** Ebenen $\mathbf{S}\langle 3, 3\rangle$ und $\mathbf{S}'\langle 3, 3\rangle$ sei

$$\mathbf{S} = \mathbf{S}' = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13\}$$

die Menge der Punkte, die die Ebenen bilden. Die Geraden der Ebene $\mathbf{S}\langle 3, 3\rangle$ seien die folgenden Teilmengen:

$\{1, 2, 3\}$, $\{1, 4, 5\}$, $\{1, 6, 7\}$, $\{1, 8, 9\}$, $\{1, 10, 11\}$, $\{1, 12, 13\}$,
$\{2, 4, 6\}$, $\{2, 5, 7\}$, $\{2, 8, 10\}$, $\{2, 9, 12\}$, $\{2, 11, 13\}$, $\{3, 4, 8\}$,
$\{3, 5, 9\}$, $\{3, 6, 11\}$, $\{3, 7, 13\}$, $\{3, 10, 12\}$, $\{4, 7, 10\}$, $\{4, 9, 13\}$,
$\{4, 11, 12\}$, $\{5, 6, 12\}$, $\{5, 8, 11\}$, $\{5, 10, 13\}$, $\{6, 8, 13\}$, $\{6, 9, 10\}$,
$\{7, 8, 12\}$, $\{7, 9, 11\}$.

Man kann leicht sehen, daß die 26 Teilmengen unserer Menge mit 13 Elementen so ausgewählt wurden, daß jede Teilmenge drei Elemente hat. Jede von den 13 Elementen tritt in sechs Teilmengen auf und eine beliebige Teilmenge $\{x, y\}$ $(x \neq y)$ ist der Teil genau einer von den oben aufgeführten Teilmengen mit drei Elementen, d. h., *die Axiomen* **B1**, **B2**, **B3** *gelten*.

Deshalb haben wir ein Modell der Ebene $\mathbf{S}\langle 3, 3\rangle$ angegeben.

Ein anderes Modell bekommen wir aus denselben Elementen, wenn wir die Teilmengen auf einer anderen Weise organisieren:

$\{1, 2, 3\}$, $\{1, 4, 5\}$, $\{1, 6, 7\}$, $\{1, 8, 9\}$, $\{1, 10, 11\}$, $\{1, 12, 13\}$,
$\{2, 4, 6\}$, $\{2, 5, 7\}$, $\{2, 8, 10\}$, $\{2, 9, 12\}$, $\{2, 11, 13\}$, $\{3, 4, 8\}$,
$\{3, 5, 9\}$, $\{3, 6, 10\}$, $\{3, 7, 13\}$, $\{3, 11, 12\}$, $\{4, 7, 12\}$, $\{4, 9, 11\}$,
$\{4, 10, 13\}$, $\{5, 6, 11\}$, $\{5, 8, 13\}$, $\{5, 10, 12\}$, $\{6, 8, 12\}$, $\{6, 9, 13\}$,
$\{7, 8, 11\}$, $\{7, 9, 10\}$.

Es ist leicht einzusehen, daß *die Axiomen* **B1**, **B2**, **B3** auch in diesem Fall *gelten*. Dieses letzte Modell wird mit $\mathbf{S}'\langle 3, 3\rangle$ bezeichnet.

2. Die **BL** Ebenen $\mathbf{S}\langle m, n\rangle$ und $\mathbf{S}'\langle m, n\rangle$ sind im wesentlichen nicht verschieden voneinander, sie sind *isomorphe Strukturen*, wenn es eine bijektive Abbildung zwischen den Punkten von $\mathbf{S}\langle m, n\rangle$ und $\mathbf{S}'\langle m, n\rangle$ gibt, die die Geraden von $\mathbf{S}\langle m, n\rangle$ in die Geraden von $\mathbf{S}'\langle m, n\rangle$ überführt. Mit Hilfe der folgenden Konfiguration können wir entscheiden, daß die oben erwähnten zwei Modelle nicht isomorph sind.

Wir betrachten eine beliebige Gerade von $\mathbf{S}\langle 3, 3\rangle$ und projizieren die Punkte dieser Gerade von einem Punkt, der nicht auf unserer Gerade liegt. Auf diesen drei Projektionsgeraden betrachten wir die dritten, von den vorigen verschiedenen Punkte. Diese drei Punkte liegen auf einer Gerade oder nicht. Die zwei vorkommenden Fälle kann man auf der Figur 1. sehen.

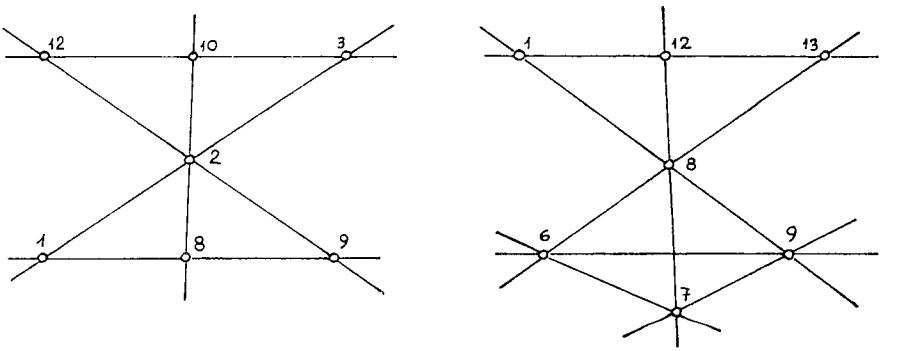

Fig. 1.

Die erste von diesen zwei Konfigurationen hat eine wichtige Rolle. Diese Konfiguration wird ein *perspektives Dreierpaar* und der Punkt 2 den *Mittelpunkt* der Konfiguration genannt. Es gibt auch solchen Mittelpunkt, der der Mittelpunkt von zwei verschiedenen Dreierpaaren ist. Dieser Punkt wird *2-fachen Mittelpunkt* genannt. Wir sprechen über einen *gewöhnlichen Punkt*, wenn der Punkt weder einfach, noch 2-fach ist. Mit Π bezeichnen wir ein perspektives Dreierpaar. Wir müssen die Konfiguration Π von den $26 \cdot 10 = 260$ Gerade-Punktpaaren auswählen, wo der Punkt und die Gerade miteinander nicht inzident sind. So bekommen wir, daß

die gewöhnliche Punkte: 4, 5, 8, 9, 10, 12,
die einfachen Mittelpunkte: 2, 3, 11, 13,
die zweifachen Mittelpunkte: 1, 6, 7

sind. Es gibt also 10 verschiedene Konfiguration Π in der Ebene $\mathbf{S}\langle 3, 3\rangle$. Auch die Mittelpunkte bilden eine Konfiguration Π, weil wir die Punkte der Gerade $\{1, 6, 7\}$ vom Punkt 3 mit Hilfe der Geraden $\{1, 2, 3\}$, $\{3, 7, 13\}$, $\{3, 6, 11\}$ in die Punkte der Gerade $\{2, 11, 13\}$ projizieren können. Diese sg. *Hauptkonfiguration* Π kann man auf der Figur 2. sehen.

Mit den Zeichnen ○, ● und ⊙ bezeichnen wir nacheinander den gewöhnlichen, den einfachen und den zweifachen Mittelpunkt.

Wenn wir die dritten Punkte der Geraden {1, 11}, {11, 7}, {7, 2}, {2, 6}, {6, 13}, {13, 1} betrachten, sind diese Punkte 10, 9, 5, 4, 8, 12, d. h. die gewöhnlichen Punkte. Die Menge dieser dritten Punkte und der diese Punkte enthaltenden Geraden bilden die Konfiguration Ω.

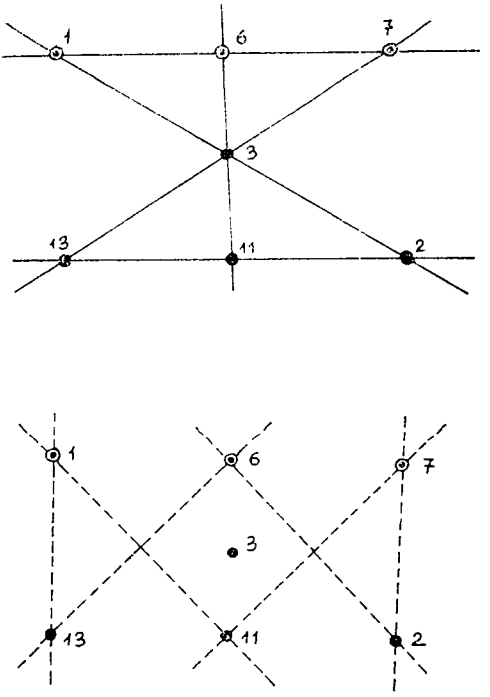

Fig. 2.

Jetzt stellen wir die Liste der Geraden zusammen, die gemeinsamen Punkt mit der Punktmenge {4, 5, 8, 9, 10, 12} haben.

Wir können sehen, daß jede von den Geraden

{4, 5, 1}, {4, 8, 3}, {4, 9, 13}, {4, 10, 7}, {4, 12, 11}, {5, 8, 11}, {5, 9, 3}, {5, 10, 13}, {5, 12, 6}, {8, 9, 1}, {8, 10, 2}, {8, 12, 7}, {9, 10, 6}, {9, 12, 2}, {10, 12, 3}

mit der Menge Ω zwei gemeinsame Punkte hat. Jede von den Geraden

{4, 2, 6}, {5, 2, 7}, {8, 6, 13}, {9, 7, 11}, {10, 1, 11}, {12,1,13}

hat einen einzigen gemeinsamen Punkt mit Ω, d. h., diese Geraden berühren die Menge Ω.

Das Oval ist eine wohlbekannte Konfiguration in der Theorie der endlichen Ebenen. Das Oval ist solche aus meisten Elementen bestehende Teilmenge der Punkte der Ebene, die höchstens zwei gemeinsame Punkte mit einer Gerade hat und die in jedem von ihren Punkten genau eine Gerade (Tangente) hat, die keinen weiteren gemeinsamen Punkt mit der obigen Teilmenge hat.

Wir bemerken, daß *die Ebene* $\mathbf{S}\langle 3, 3\rangle$ *aus dem Oval* Ω *und aus der Hauptkonfiguration* Π *besteht. Die Tangenten von* Ω *sind die Seitengeraden des Sechsecks, dessen Ecken zu dreien zwei Geraden bilden.*

3. Die Untersuchung der Ebene $\mathbf{S}'\langle 3, 3\rangle$ geht ähnlicherweise, wie bei $\mathbf{S}\langle 3, 3\rangle$.

Wenn wir die 260 miteinander nicht inzidenten Gerade-Punktpaare betrachten, können wir sehen, daß die Ebene keine Konfiguration Π enthält, d. h., jede Punkt der Ebene gewöhnlicher Punkt ist. Daraus folgt schon, daß $\mathbf{S}\langle 3, 3\rangle$ *und* $\mathbf{S}'\langle 3, 3\rangle$ *keine isomorphe Modelle sind.*

Im folgenden beweisen wir auf einer anderen Weise, daß $\mathbf{S}\langle 3, 3\rangle$ und $\mathbf{S}'\langle 3, 3\rangle$ nicht isomorph sind. Wir zeigen nämlich, daß es eine solche Konfiguration in $\mathbf{S}'\langle 3, 3\rangle$ gibt, die man in $\mathbf{S}\langle 3, 3\rangle$ nicht finden kann. Nun betrachten wir die Punkte 2, 3, 4, 6, 8, 10 von $\mathbf{S}'\langle 3, 3\rangle$. Diese Punkte sind die Eckpunkte eines *vollständigen Vierseits.* Die Seiten dieses vollständigen Vierseits sind die Punktmenge {2, 4, 6}, {2, 8, 10}, {3, 4, 8}, {3, 6, 10}. Die Punkte 1, 2, 4, 5, 6, 7 bzw. 1, 3, 4, 5, 8, 9 sind die Eckpunkte je eines vollständigen Vierseits, dessen Seiten die Punktmengen {1, 4, 5}, {1, 6, 7}, {2, 4, 6}, {2, 5, 7} bzw. {1, 4, 5}, {1, 8, 9}, {3, 4, 8}, {3, 5, 9} sind. Die obigen drei Vierseite ordnen sich entsprechend der Figur 3. Mit Σ bezeichnen wir diese aus drei vollständigen Vierseiten bestehende Konfiguration.

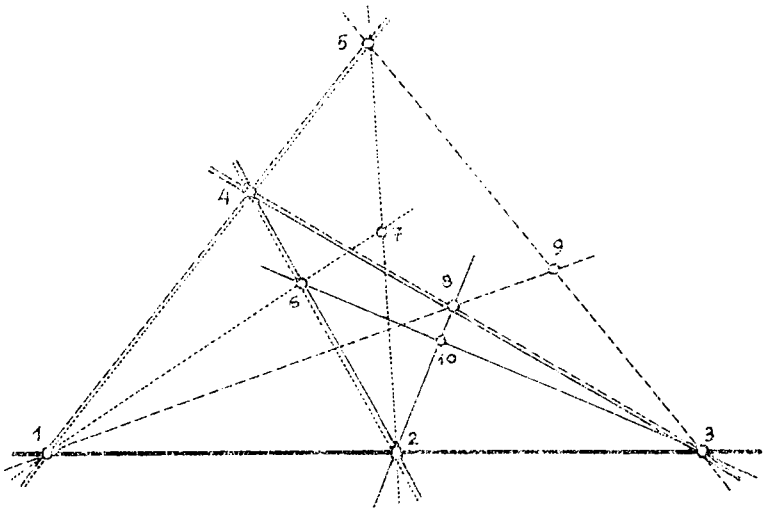

Fig. 3.

Es ist zu sehen, daß die Punkte 2 und 3, 1, und 2 bzw. 1 und 3 die Gegenecken je eines vollständigen Vierseites sind und diese Punkte auf einer Gerade liegen. Man kann auch das sehen, daß der Punkt 4 der gemeinsame Punkt der drei vollständigen Vierseite ist.

Nun betrachten wir die Ebene $\mathbf{S}\langle 3,3\rangle$. A und B seien beliebige Punkte der Ebene. Nehmen wir eine von den durch A verlaufenden Geraden, die den Punkt B nicht enthält. C und D seien die weiteren zwei Punkte dieser Gerade. Wir betrachten die Geraden BC und BD. E bzw. F sei der dritte Punkt auf der Gerade BC bzw. BD. Wenn die Punkte A, E und F auf derselben Gerade liegen, dann bekommen wir ein vollständiges Vierseit, dessen Eckpunkte A, B, C, D, E, F sind. Nun probieren wir dieses vollständige Vierseit zur Konfiguration Σ ergänzen. Ein von den Punkten C, D, E, F soll dem Punkt 4 der Ebene $\mathbf{S}'\langle 3,3\rangle$ entsprechen. Wir wählen einen von den Punkten C, D, E, F aus. Dann ist die Ergänzung der Konfiguration schon eindeutlich und es ist einzusehen, daß wir keine Konfiguration Σ bekommen. Insgesamt $390+3t$ Fälle müssen wir untersuchen, wo t die Zahl der vollständigen Vierseite ist.

Nach der Durchführung der Rechnungen *bekommen wir die Konfiguration Σ in der Ebene* $\mathbf{S}\langle 3,3\rangle$ *niemals*. Daraus folgt, daß *die Ebenen* $\mathbf{S}\langle 3,3\rangle$ *und* $\mathbf{S}'\langle 3,3\rangle$ *nicht isomorph sind.*

Nun betrachten wir die Ebene $\mathbf{S}'\langle 3,3\rangle$. Man kann Ovale auch in dieser Ebene finden. Das Oval enthält höchstens 6 Punkte auch hier, weil keine Tangente im entgegengesetzten Fall gibt.

Wir betrachten die Liste der Geraden, die durch je zwei Elemente der Punktmenge $\Omega' = \{1, 5, 7, 9, 11, 13\}$ bestimmt sind:

$$\{1, 5, 4\}, \{1, 7, 6\}, \{1, 9, 8\}, \{1, 11, 10\}, \{1, 13, 12\}, \{5, 7, 2\},$$
$$\{5, 9, 3\}, \{5, 11, 6\}, \{5, 13, 8\}, \{7, 9, 10\}, \{7, 11, 8\}, \{7, 13, 3\},$$
$$\{9, 11, 4\}, \{9,13, 6\}, \{11, 13, 2\}.$$

Wir können sehen, daß jede von diesen Geraden nur 2 Punkte von Ω' enthält.

Die Tangenten von Ω' sind die folgenden Geraden:

$$\{1, 2, 3\}, \{5, 10, 12\}, \{7, 4, 12\}, \{9, 2, 12\}, \{11, 3, 12\}, \{13, 4, 10\}.$$

Im Punkt 12 treffen sich vier Tangenten. Die Konfiguration der Tangenten können wir auch folgenderweise kennzeichnen. *Sie bilden die Seitengeraden der Dreiecke, die eine gemeinsame Ecke haben.* Es handelt sich um die Dreiecke $\{2, 3, 12\}$ und $\{4, 10, 12\}$.

Nun betrachten wir die Punktmenge $\Omega'' = \{1, 5, 7, 9, 11, 12\}$. Wie im vorigen können wir einsehen, daß auch Ω'' ein Oval ist und Ω'' gleiche Eigenschaften hat, wie Ω'. Dasselbe wissen wir auch über die Punktmenge $\Omega''' = \{5, 7, 9, 11, 12, 13\}$. Man kann sehen, daß sich diese drei Ovale nur in je einem Punkt voneinander unterscheiden. Das bedeutet, daß *fünf Punkte in der Ebene* $\mathbf{S}'\langle 3,3\rangle$ *ein Oval nicht eindeutlich bestimmt.*

4. Die in der Einführung erwähnte Ebene $\mathbf{S}\langle 3,4\rangle$ besteht aus 15 Punkten und 35 Geraden. Jede von den Geraden besteht aus drei Punkten und durch jeden Punkt geht sieben Geraden hindurch.

Bezeichne 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 die Punkte. Die Liste der Geraden ist:

{1, 2, 9}, {1, 3, 10}, {1, 4, 11}, {1, 5, 12}, {1, 6, 13}, {1, 7, 14},
{1, 8, 15}, {2, 3, 15}, {2, 4, 14}, {2, 5, 13}, {2, 6, 12}, {2, 7, 11},
{2, 8, 10}, {3, 4, 9}, {3, 5, 14}, {3, 6, 11}, {3, 7, 12}, {3, 8, 13},
{4, 5, 15}, {4, 6, 10}, {4, 7, 13}, {4, 8, 12}, {5, 6, 9}, {5, 7, 10},
{5, 8, 11}, {6, 7, 15}, {6, 8, 14}, {7, 8, 9}, {9, 10, 11}, {9, 12, 13},
{9, 14, 15}, {10, 12, 15}, {10, 13, 14}, {11, 12, 14}, {11, 13, 15}.

Wir bemerken, daß die Punkte 9, 10, 11, 12, 13, 14, 15 und die aus diesen Punkten bekommenden Punktdreier (s. die letzten sieben Dreier der vorigen Liste) als Geraden eine wohlbekannte Konstruktion repräsentieren. Das ist eine *Galois*-Ebene zweiter Ordnung, anders eine *Fano*-Ebene. Mit Φ bezeichnen wir die obige Teilebene.

Greifen wir den Punkt vom Index 8 der Ebene $\mathbf{S}\langle 3,4\rangle$ heraus und betrachten die dritten Punkte auf der Geraden, die den Punkt 8 und die Punkte von Φ verbinden. Diese dritten Punkte sind 1, 2, 3, 4, 5, 6, 7, weil es sich um die Geraden {1, 8, 15}, {2, 8, 10}, {3, 8, 13}, {4, 8, 12}, {5, 8, 11}, {6, 8, 14}, {7, 8, 9} handelt. Diese Geraden sind die Tangenten der Punktmenge $\Omega = \{1, 2, 3, 4, 5, 6, 7\}$, weil jede von den obigen Geraden nur durch einen einzigen Punkt von Ω hindurchgeht. Weiterhin ist Ω ein Oval nach den Geraden, die je zwei Punkte von Ω verbinden:

{1, 2, 9}, {1, 3, 10}, {1, 4, 11}, {1, 5, 12}, {1, 6, 13}, {1, 7, 14},
{2, 3, 15}, {2, 4, 14}, {2, 5, 13}, {2, 6, 12}, {2, 7, 11}, {3, 4, 9},
{3, 5, 14}, {3, 6, 11}, {3, 7, 12}, {4, 5, 15}, {4, 6, 10}, {4, 7, 13},
{5, 6, 9}, {5, 7, 10}, {6, 7, 15}.

Wenn es einen solchen Punkt gibt, in dem sich die sämtlichen Tangenten eines Ovals treffen, dann wird dieser Punkt Kernpunkt genannt. Der folgende Satz faßt unsere bisherige Ergebnisse zusammen:

Die Ebene $\mathbf{S}\langle 3,4\rangle$, *hinsichtlich ihrer Punkte, besteht aus einem Kernpunkt, aus einem Oval und aus einer Fano-Ebene. Die Tangenten stellen die folgende bijektive Abbildung zwischen dem Oval und der Fano-Ebene her:*

$$\Pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 15 & 10 & 13 & 12 & 11 & 14 & 9 \end{pmatrix}$$

* * *

Die hier erörterten Bemerkungen zeigen, daß die endlichen **BL** Ebenen schöne Struktureigenschaften haben. Die Untersuchungen der Fragen, die die allgemeinen Fälle betreffen, versprechen interessant zu werden.

Literatur

[1] CROWE, D. W., Projective and inversive models for finite hyperbolic planes, *Mich. Math. J.*, **13** (1966), 251–256.
[2] GRAVES, L. M., A finite Bolyai–Lobachevsky plane, *Amer. Math. Monthly*, **69** (1962), 130–132.
[3] HENDERSON, M., Finite Bolyai–Lobachevsky k-spaces, *Coll. Math.*, **15** (1966), 205–210.
[4] KÁRTESZI F., Egy legkisebb véges reguláris hiperbolikus sík, *MTA. III. Oszt. Közl.*, **19** (1969), 5–7.
[5] PUHAREV, N. K., Some properties of finite Lobachevsky planes, *Perm. Gos. Univ. Uchen. Zap. Mat.*, **103** (1963), 61–63.
[6] SZAMKOLOWICZ, J., On the problem of existence of finite regular planes, *Coll. Math.*, **9** (1962), 245–250.

SOME REMARKS ON q-ADDITIVE FUNCTIONS

By

J. FEHÉR and I. KÁTAI

Teacher's Training College, Pécs and
Department of Numerical Analysis and Computer Science of the
L. Eötvös University, Budapest

(Received November 26,1983)

1. Let $q \geq 2$ be an integer. We shall say that a function f defined on the set of nonnegative integers is q-additive if $f(a+mq^s) = f(a)+f(mq^s)$ holds for $0 \leq a < q^s$, $m \geq 0$, $s = 1, 2, \ldots$. Let $\mathcal{A}_q$ be the class of realvalued q-additive functions.

The function $f(n) = cn$ belongs to $\mathcal{A}_q$, and behaves very regularly. It plays a similar role as $c \log n$ does for additive functions.

THEOREM 1. *Let $f \in \mathcal{A}_q$. The condition*

$$\Delta f(n) := f(n+1)-f(n) = O(1) \tag{1.1}$$

holds if and only if f can be written in the form

$$f(n) = cn+g(n), \tag{1.2}$$

where

$$g(aq^j) = O(1), \quad (a = 1, \ldots, q-1;\ j = 0, 1, 2, \ldots) \tag{1.3}$$

$$\sum_{l=0}^{s} g((q-1)q^j) = O(1), \quad (s \to \infty). \tag{1.4}$$

It is obvious that $f \in \mathcal{A}_q$ is bounded if and only if

$$\sum_{j=0}^{\infty} \sum_{a=1}^{q-1} |f(aq^j)| < \infty. \tag{1.5}$$

THEOREM 2. *Let $f_0, f_1, \ldots, f_k \in \mathcal{A}_q$, $l(n) = f_0(n)+\ldots+f_k(n+k)$. The condition*

$$l(n) = O(1), \quad (n = 0, 1, 2, \ldots) \tag{1.6}$$

holds if and only if

$$f_i(n) = c_i n+g_i(n), \quad (i = 0, 1, \ldots, k), \tag{1.7}$$

$$g_i(aq^j) = O(1), \quad (a = 1, 2, \ldots, q-1; j = 0, 1, 2, \ldots), \tag{1.8}$$

$$\sum_{l=0}^{s} g_i((q-1)q^l) = O(1), \ (s \to \infty), \tag{1.9}$$

$$\sum_{j=0}^{\infty} \sum_{a=1}^{q-1} \left| \sum_{i=0}^{R} g_i(aq^j) \right| < \infty, \tag{1.10}$$

and

$$c_0 + c_1 + \ldots + c_k = 0. \tag{1.11}$$

Let us consider now Theorem 2 in the special case. Let $P(z) = \alpha_0 + \alpha_1 z + \ldots + \alpha_k z^k$, E denote the shift operator defined as $Ey_n = y_{n+1}$, $E^0 y_n = y_n$.

THEOREM 3. *Let $f \in \mathcal{A}_q$. The relation $P(E)f(n) = O(1)$ holds if $f(n)$ is bounded. There exists no more solution if $P(1) \neq 0$ while in the case $P(1) = 0$ the relation $P(E)f(n) = O(1)$ holds if and only if $\Delta f(n) = O(1)$.*

LEMMA 1. Let $f \in \mathcal{A}_q$. The relation

$$\frac{1}{x} \sum_{n \leq x} |f(n)| \ll 1, \quad (x \to \infty) \tag{1.12}$$

holds if and only if

$$f(aq^j) = O(1), \quad (a = 1, \ldots, q-1; j = 0, 1, 2, \ldots) \tag{1.13}$$

$$\alpha_N = \sum_{j=0}^{N} \sum_{a=1}^{q-1} f(aq^j) \tag{1.14}$$

is bounded,

$$\sum_{j=0}^{\infty} \sum_{a=1}^{q-1} f^2(aq^j) < \infty. \tag{1.15}$$

We shall use Lemma 1 to prove Theorem 6.

THEOREM 4. *Let $f \in \mathcal{A}_q$. and*

$$\sum_{n \leq x} |\Delta f(n)| = O(x). \tag{1.16}$$

Then f can be written in the form

$$f(n) = cn + g(n), \tag{1.17}$$

furthermore with the notations

$$\begin{cases} \Delta_s = g(q^{s+1} - 1), \ \lambda_s = \dfrac{1}{q^{s+1}} (\Delta_{s+1} - q\Delta_s) \\ \eta_{s,a} = f(aq^s) - a\Delta_{s-1}, \quad (a = 1, \ldots, q-1) \end{cases} \tag{1.18}$$

we have

$$\frac{\Delta s}{q^s} \to 0, \quad (s \to \infty), \tag{1.19}$$

$$\sum_{j=0}^{\infty} |\lambda_j| < \infty, \tag{1.20}$$

$$\sum_{s=0}^{\infty} \sum_{a=1}^{q-1} |\eta_{s,a}| \frac{1}{q^s} < \infty. \tag{1.21}$$

If $f \in \mathcal{A}_q$ satisfies the conditions (1.17)–(1.21), *then* (1.16) *holds.*

THEOREM 5. *Let $f_0, \ldots, f_k \in \mathcal{A}_q$, $l(n) = f_0(n) + \ldots + f_k(n+k)$, and*

$$\sum_{n \leq x} |l(n)| = O(x). \tag{1.22}$$

Then

$$\sum_{n \leq x} |\Delta f_i(n)| = O(x), \quad (i = 0, \ldots, k), \tag{1.23}$$

$$\sum_{n \leq x} |F_k(n)| = O(x), \tag{1.24}$$

where

$$F_k(n) = f_0(n) + \ldots + f_k(n).$$

The fulfilment of (1.23), (1.24) *involves* (1.22).

From this and the earlier theorems it follows immediately

THEOREM 6. *Let $f_0, \ldots, f_k \in \mathcal{A}_q$, $l(n) = f_0(n) + \ldots + f_k(n+k)$. The relation*

$$\sum_{n \leq x} |l(n)| = O(x)$$

holds if and only if the following assertions are true:

$$f_i(n) = c_i n + g_i(n), \quad (i = 0, \ldots, k), \tag{1}$$

$$g_i(q^{s+1} - 1) = o(q^s), \quad (s \to \infty;\ i = 0, \ldots, k), \tag{2}$$

$$\sum_{i=0}^{k} \sum_{s=0}^{\infty} q^{-s} |g_i(q^{s+1} - 1) - q g_i(q^s - 1)| < \infty, \tag{3}$$

$$\sum_{i=0}^{k} \sum_{s=0}^{\infty} \sum_{a=1}^{q-1} q^{-s} |g_i(aq^s) - a g_i(q^s - 1)| < \infty, \tag{4}$$

$$\sum_{j=0}^{N} \sum_{a=0}^{q-1} \sum_{i=0}^{k} f_i(aq^j) = \gamma_N \tag{5}$$

is bounded as $N \to \infty$,

$$\sum_{j=0}^{\infty} \sum_{a=1}^{q-1} \left(\sum_{i=0}^{k} f_i(aq^j) \right)^2 < \infty, \tag{6}$$

$$c_0 + c_1 + \ldots + c_k = 0. \tag{7}$$

As a special case of Theorem 5 we have

THEOREM 7. *Let* $f \in \mathcal{A}_q$. *The relation*

$$\sum_{n \leq x} |P(E)f(n)| = O(x) \tag{1.25}$$

obviously holds if

$$\sum_{n \leq x} |f(n)| = O(x).$$

There exists no more solution if $P(1) \neq 0$. *If* $P(1) = 0$, *then* (1.25) *holds if and only if*

$$\sum_{n \leq x} |\Delta f(n)| = O(x).$$

2. PROOF of THEOREM 1. Let $n+1 = b + mq^s$, $0 < b < q^s$. Then $\Delta f(n) = \Delta f(b-1)$. If $n+1 = mq^s$, $m \geq 1$, then $n = (q-1)(1+q+\ldots+q^{s-1}) + (m-1)q^s$, and so

$$\Delta f(n) = f(mq^s) - f((m-1)q^s) - f((q-1)(1+q+\ldots+q^{s-1})).$$

Let us denote by A_l the sum

$$A_l = f(q-1) + f((q-1)q) + \ldots + f((q-1)q^l).$$

Assume that (1.1) holds. By putting $n = mq^s - 1$, we get that

$$|f(mq^s) - f((m-1)q^s) - A_{s-1}| \leq K \tag{2.1}$$

for every $m \geq 1$, $s \geq 1$. Here and as follows K is a suitable large constant. Summing these inequalities for $m = 1, 2, \ldots, q-1$, we deduce that

$$|f((q-1)q^s) - (q-1)A_{s-1}| \leq K(q-1),$$

whence we have

$$|A_s - qA_{s-1}| \leq K(q-1),$$

and so for $\beta_s = q^{-s}A_s$ we get

$$|\beta_s - \beta_{s-1}| \leq \frac{K}{q^{s-1}}. \tag{2.2}$$

From (2.2) if follows immediately that there exists

$$\lim_{s \to \infty} \beta_s = \gamma,$$

and that

$$|\gamma-\beta_s|\leq\frac{K}{q^s}. \tag{2.3}$$

Let us consider the function $g(n):=f(n)-cn$, $c=\frac{\gamma}{q}$. Then $\Delta g(n)=O(1)$, consequently (2.2), (2.3) holds with $\gamma=0$ and β'_s instead of β_s, where

$$\beta'_s=q^{-s}\sum_{j=0}^{s}g((q-1)q^j).$$

$\beta'_s=O(q^{-s})$ involves (1.4). By considering (2.1) with g instead of f we get (1.3) immediately. So we proved that (1.1) involves (1.2), (1. 3), (1.4). It is obvious that for a $g\in\mathcal{A}_q$ satisfying (1.3), (1.4) the relation $\Delta g(n)=O(1)$ holds. Consequently (1.2), (1.3), (1.4) involve (1.1). ■

3. Proof of Theorem. **2.** Let

$$F_r(n)=\sum_{i=0}^{r}f_i(n),\quad S_r(n)=\sum_{i=r}^{k}f_i(n).$$

Assume that (1.6) holds.

Let $n=a+mq^s$, $a<q^s-k$. Then

$$l(a+mq^s)=l(a)+\sum_{i=0}^{k}f_i(mq^s)=l(a)+F_k(mq^s). \tag{3.1}$$

By putting $a=0$, we get that $F_k(mq^s)=:t(m)$ is bounded. Observing that $t\in\mathcal{A}_q$, by our earlier remark we get

$$\sum_{j=0}^{\infty}\sum_{a=1}^{q-1}|F_k(aq^j)|<\infty. \tag{3.2}$$

Let now $n=mq^s-r$, $0<r<k$. We have

$$\begin{aligned}l(mq^s-r)&=f_0(mq^s-r)+\ldots+f_{r-1}(mq^s-1)+f_r(mq^s)+\ldots+\\&+f_k(mq^s+k-r)=F_{r-1}((m-1)q^s)+S_r(mq^s)+f_0(q^s-r)+\ldots+\\&+f_{r-1}(q^s-1)+f_r(0)+\ldots+f_k(k-r)=F_{r-1}((m-1)q^s)+\\&+S_r(mq^s)-S_r(q^s)+l(q^s-r)=F_k(mq^s)+F_{r-1}((m-1)q^s)-F_{r-1}(mg^s)-\\&-S_r(q^s)+l(q^s-r).\end{aligned}$$

Hence it follows that $F_{r-1}(mq^s)-F_{r-1}((m-1)q^s)=O(1)$, $(m\to\infty)$. By using this for $r=1,2,\ldots$ we deduce that $f_i(mq^s)-f_i((m-1)q^s)=O(1)$, $(m\to\infty)$. This last relation involves that $\Delta f_i(n)=0$. Really, for $1+n=mq^s+a$, $0>a>q^s$ we get $\Delta f_i(n)=\Delta f_i(a-1)$ while for $n+1=mq^s$ we have

$$f_i(mq^s)-f_i(mq^s-1)=f_i(mq^s)-f_i((m-1)q^s)-f_i(q^s-1)=O(1).$$

Consequently the relations (1.7), (1.8), (1.9) hold with suitable constant c_i. Now we prove (1.11).

Let $\delta = c_0+c_1+\ldots+c_k$. (1.8) involves that $F_k(aq^j) = \delta aq^j+O(1)$. From (3.2) we get that $\delta = 0$. The necessity of the conditions (1.7)–(1.11) has been proved. It is obvious that these conditions are sufficient for (1.6). ■

4. Proof of Lemma 1. The assertion is an easy consequence of basic probabilistic theorems. Let $\xi_0, \xi_1, \ldots$ be independent random variables with the distribution

$$P(\xi_i = f(aq^i)) = \frac{1}{q}, \quad (a = 0, 1, \ldots, q-1).$$

Let $\eta_N = \xi_0+\ldots+\xi_N$. It is obvious that

$$M|\eta_{N-1}| = q^{-N} \sum_{N<q^N} |f(n)|.$$

Let us assume now that (1.12) holds. Then

$$M|\eta_N| < c, \quad (N \to \infty). \tag{4.1}$$

Consequently $|M\eta_N| < c$, and hence by $\alpha_N = M\eta_N$ we get (1.14). Since

$$M|\eta_N| = \frac{1}{q} \sum_{a=0}^{q-1} M|\eta_{N-1}+f(aq^j)|,$$

from (4.1) we get (1.13).

Let

$$m_k = M\xi_k,\ D_k^2 = M|\xi_k-m_k|^2,\ H_k^3 = M|\xi_k-m_k|^3,$$

$$S_n^2 = \sum_{k=0}^{n} D_k^2,\ K_n^3 = \sum_{k=0}^{n} H_k^3.$$

Since ξ_k, consequently m_k are bounded, we get that $H_k^3 \ll D_k^2$. Hence we have $K_n^3 \ll S_n^2$.

We shall prove that $S_n = O(1)$. Let us assume that $S_n \to \infty$. Then $\frac{K_N}{S_N} \to 0$, i.e. the Liapunov condition for the central limit theorem is fulfilled, consequently

$$\lim_N P\left(\frac{|\eta_N-\alpha_N|}{S_N} < y\right) = \Phi(y)-\Phi(-y)$$

holds for every $y \geq 0$. Hence we get that

$$\lim_N P(|\eta_N| > H) = 1$$

for every H, that involves that $M|\eta_N| \to \infty$, but this is impossible. So we have $S_n^2 = O(1)$. Since

$$M(\xi_k-m_k)^2 = \frac{1}{q} \sum_{a=1}^{q-1} (f(aq^k)-m_k)^2 \geq \frac{1}{q}(f(0)-m_k)^2 \geq \frac{m_k^2}{q},$$

we get that $\sum_{k=0}^{\infty} m_k^2 < \infty$, and so $S_n^2 = O(1)$ involves (1.15).

Let us assume now that (1.13), (1.14), (1.15) hold. Then $D^2\eta_N = O(1)$, $M|\eta_N-\alpha_N|^2 = O(1)$, $M|\eta_N-\alpha_N| = O(1)$, and so $M|\eta_N| = O(1)$. Let now $q^N \leq x \leq q^{N+1}-1$. Since

$$\frac{1}{x}\sum_{n\leq x}|f(n)| \leq qM|\eta_N|,$$

(1.12) is proved. ■

5. Proof of Theorem 4. Let $\Lambda_s = f(q^{s+1}-1)$, $s = -1, 0, 1, \ldots$. Assume that (1.16) holds. If $n+1 = aq_s+mq^{s+1}$, $0<a<q$, then $n = (q_s-1)+$ $+(a-1)q_s+mq^{s+1}$, and so

$$\Delta f(n) = f(aq^s)-f\big((a-1)q^s\big)-\Lambda_{s-1}.$$

The number of $n<q^N$ having the above form with fixed a and s is q^{N-s-1} consequently

$$q^{-N}\sum_{s=0}^{N-1}\sum_{a=1}^{q-1}|f(aq^s)-f\big((a-1)q^s\big)-\Lambda_{s-1}|\cdot q^{N-s-1} \ll 1,$$

and so

$$\sum_{s=0}^{\infty}\sum_{a=1}^{q-1}|f(aq^s)-f\big((a-1)q^s\big)-\Lambda_{s-1}|\cdot q^{-s} < \infty. \tag{5.1}$$

Hence we get that

$$\sum_{s=0}^{\infty}\sum_{a=1}^{q-1}|f(aq^s)-a\Lambda_{s-1}|q^{-s} < \infty, \tag{5.2}$$

and by observing that $f\big((q-1)q^s\big) = \Lambda_s-\Lambda_{s-1}$, we have

$$\sum_{s=1}^{\infty}|\Lambda_s-q\Lambda_{s-1}|q^{-s} < \infty. \tag{5.3}$$

Let $\beta_s = q^{-s}\Lambda_s$. From (5.3) we get that $\lim\limits_{s\to\infty}\beta_s = \gamma$ exists. Substituting $f(n)$ with $f(n)-cn$, $c = \gamma/q$, if needed, we may assume that $\gamma = 0$. The relations (1.18)–(1.21) immediately follow from (5.2), (5.3).

The converse assertion is obvious. ■

6. Proof of Theorem 5. Let us assume that (1.22) holds. By using the notations of section 3, from (3.1) we deduce that

$$\sum_{m<q^N}|F_k(mq^s)| = O(q^N) \tag{6.1}$$

and from (3.3) that

$$\sum_{m<q^N}|F_{r-1}(mq^s)-F_{r-1}\big((m-1)q^s\big)| = O(q^N),\ (r = 1, \ldots, k). \tag{6.2}$$

Here s being fixed so that $k<q^s$.

By repeating the argument used in section 3, from (6.1) we deduce that

$$\sum_{n<x} |F_k(n)| = O(x), \tag{6.3}$$

and from (6.2) that

$$\sum_{n<x} |\Delta f_i(n)| = O(x), \quad (i = 0, 1, \ldots, k). \tag{6.4}$$

The necessity of (1.23), (1.24) is proved.

The sufficiency of these conditions is almost obvious. We split the integers $n<q^N-k$ into disjoint subclasses $n = a+mq^s$, $(a = 0, \ldots, q^s-k-1)$, $n = mg^s-r$, $(r = 0, \ldots, k-1)$.

Assume that (6.3), (6.4) hold. From (3.1), (3.3), (6.3), (6.4) we deduce immediately that

$$\sum_{n<q^N-k} |l(n)| = O(q^N), \quad (N \to \infty)$$

and so (1.22) holds. ■

DIE BIKONJUGIERTEN PUNKTSYSTEME

Von

SZ. MÁRTA HOLLAI

Lehrstuhl für Darstellende Geometrie der L. Eötvös Universität, Budapest

(Eingegangen am 17 April 1983)

In dieser Arbeit geben wir eine hinreichende Bedingung dafür an, daß ein diskretes Punktsystem des n dimensionalen euklidischen Raumes bikonjugiert ist. (S. Definition 2., SATZ 1.) Eine notvendige Bedingung wird nur für Punktgitter angegeben. (SATZ 2.) Wir beschäftigen uns ferner mit den bikonjugierten Gittertypen von E^2 und E^3.

1. Die binonjugierten Punktsysteme im E^n

DEFINITION 1. Eine Punktmenge $\mathfrak{M}'$ wird konjugiert zur diskreten Punktmenge $\mathfrak{M}$ genannt, wenn $\mathfrak{M}'$ aus den Eckpunkten der Dirichletschen-Voronoischen Zelle von $\mathfrak{M}$ besteht [1]. (D–V Zelle: [3].)

DEFINITION 2. Ist $(\mathfrak{M}')' = \mathfrak{M}$, dann wird $\mathfrak{M}$ bikonjugiert genannt.

Im weitern spielt der Begriff der Stütz- oder Leerkugel eines Punktsystems eine wichtige Rolle. Ein solche liegt vor, wenn eine Kugeloberfläche mindestens $n+1$ Punkte eines Punktsystems enthält, diese Punkte nicht in einer Hyperebene liegen und sich kein Punkt des Systems in ihrem Inneren befindet.

Nimmt man die konvexe Hülle der auf einer Stützkugel liegenden Punkte, so bekommen wir eine sogenannten L-Polytop.

Bildet man alle Stützkugeln und L-Polytopen eines diskreten Punktsystems, so bekommen wir die Delaunay–Zerlegung des Raumes, kurtz: L-Zerlegung. Das ist eine duale Zerlegung der Dirichletschen–Voronoischen–Zerlegung. Die Ecke der D-V-Zellen sind die Mittelpunkte der Stützkugeln von Punktsystem.

SATZ 1. *Wenn die Radien jeder Stützkugel einer diskreten Punktmenge $\mathfrak{M}$ in E^n gleich sind, dann ist $\mathfrak{M}$ eine bikonjugierte Punktmenge.*

BEWEIS. Wir bezeichen die Punkte von $\mathfrak{M}$ mit P_i, von $\mathfrak{M}'$ mit P'_j, den Radius der Stützkugeln mit R und mit L_i^* die Gesamtheit der Polytope, die P_i als Eckpunkt besitzen. Die konvexe Hülle der Mittelpunkte der Um-

kugeln der L-Polytope in L_i^* ist die D-V-Zelle von P_i. Da nach Voraussetzungen die Radien aller Stützkugeln gleich sind, so ist diese Zelle ein n-dimensionaler Polytop mit wenigstens $n+1$ Ecken, die einer Kugel mit dem Radius R und mit dem Mittelpunkt P_i einbeschrieben ist. Nach der Voraussetzung ist auch $P_iP'_j \geq R$. Daher ist die Umkugel dieser D-V-Zellen eine Stützkugel von $\mathfrak{M}'$. Es ist also die D-V-Zerlegung des Raumes nach $\mathfrak{M}$ der L-Zerlegung nach $\mathfrak{M}'$ gleich. Diese Zerlegungen sind eindeutig ([3]), deshalb sind alle Punkte von $\mathfrak{M}$ und nur diese die Ecken der D-V-Zellen von $\mathfrak{M}'$.

SATZ 2. *Wenn $\{P_i\} := \Gamma$ ein bikonjugiertes Punktgitter des n-dimensionalen euklidischen Raumes ist, dann sind die Radien der Stützkugeln von Γ gleich.*

BEWEIS. Es gibt nur endlich viele inkongruente L-Polytope in einem Gitter [3]. Sei R der Radius der Stützkugel mit kleinstem Radius. Nach der Bezeichnung im ersten Satz sind

$$P_iP'_j \geq R. \tag{1}$$

(2) Die Gleicheit gilt für jedes i,

weil Γ gitterförmig ist. Nach der Voraussetzung bilden die Punkte von Γ und nur diese Punkte die Mittelpunkte der Stützkugeln von Γ. Nach (1) sind die Radien dieser Stützkugeln mindestens gleich R. Nach (2) sind sie nicht größer als R, sonst wäre mindestens ein Punkt von P'_j im Inneren, d.h. Die Kugeln wäre nicht leer.

SATZ 3. *Wenn die Radien aller Stützkugeln einer diskreten Punktmenge $\mathfrak{M}$ in E^n-gleich sind, dann liegen die Mittelpunkte der Umkugeln aller* L-*Polytope von $\mathfrak{M}$ im Inneren der* L-*Polytope.*

BEWEIS (indirekt). Nehmen wir an, daß L_m von L-Polytopen seinen Mittelpunkt nicht im Inneren enthält. So gibt es wenigstens eine $n-1$ dimensionale Seitenfläche von $L_m(:=l_m)$, deren Hyperebene den Mittel punkte P'_m und der P_m trennt, wobei P_m der nicht in dieser Hyperebene liegende Punkt von L_m ist. Die Zerlegung ist normal [3]. Daher gehört noch eine Stützkugel zu dieser Seitenfläche l_m. Ihr Radius ist auch R nach Voraussetzung, aber die beiden Kugeln sind nicht identisch. So ist ihr Mittelpunkt im gleichen Halbraum wie P_m. In diesem Fall enthält die zweite Kugel P_m im Inneren, d.h. sie ist nicht leer.

2. Die bikonjugierten Gittertypen von E^2 und E^3

Fjodorow hat die Gitter nach L-Zerlegung klassifiziert [4]:

In E^2:

1. *Rechtecksgitter:* die L-Zellen sind Rechtecke, (die D-Zelle ist auch Rechteck);
2. *Dreiecksgitter* (das primitive Gitter): die L-Zellen sind Dreiecke (die D-Zelle ist Sechseck).

In E^3:

1. *Quadergitter:* alle 8 Ecken des Grundparallelepipeds des Gitters befinden sich auf einer Stützkugel.

2. *"Prismengitter mit 6 Punkten"*: 6 Gitterpunkten liegen auf einer Stützkugel und bilden ein Prisma.

3. *"Oktaedergitter":* das Gitter solche Stützkugeln hat, auf denen genau 6 Punkte des Gitters sind, die aber ein Oktaeder bilden.

4. *"Pyramidengitter mit 5 Punkten":* das Gitter hat solche Stützkugeln, auf denen sich genau 5 Punkte des Gitters befinden. Diese Punkte bilden eine Pyramide mit rechteckiger Grundfläche.

5. *"Tetraedergitter"* (das primitive Gitter): jede Stützkugel des Gitters hat genau 4 Punkte des Gitters.

Im ersten Teil haben wir gesehen, daß ein Gitter dann und nur dann bikonjugiert ist, wenn alle Stützkugelradien des Gitters gleich sind. Daraus folgt:

SATZ 4. *In der euklidischen Ebene sind alle Gitter bikonjugiert* (vgl. [1]).

SATZ 5. *Im dreidimensionalen euklidischen Raum sind die folgenden Gitter bikonjugiert:*

– *alle Quadergitter,*
– *alle "Prismengitter mit 6 Punkten" und*
– *solche Tetraedergitter, die zwei windschiefe Kanten haben, die rechte Keilwinkel tragen (Rechtwinkelkanten).*

BEWEIS. Es ist nach dem in [2] beweisenen Hilfssatz 2. gilt: wenn 4 L-Polyeder eines Gitters in einer Kante zusammankommen, sind die Radien der diesem L-Polyeder umschriebenen Kugeln dann und nur dann gleich, wenn die in dieser Kante zusammankommenden Seitenflächen zueinander senkrecht sind. Der Hilfssatz 3. von [2] schließt jedes "Oktaedergitter" und "Pyramidengitter mit 5 Punkten" aus.

Literatur

[1] HOLLAI, M., Die konjugierte gitterförmige inkongruente Kreispackung der Ebene, *Annales, Univ. Sci. Budapest*, Sectio Math., 21 (1978), 149–155.
[2] HOLLAI, M., Das dichteste gitterförmige ϱ-System der Kugeln, *Annales, Univ. Sci. Budapest*, Sectio Math., 24 (1981), 157–180.
[3] ДЕЛОНЕ, Б. Н.–САНДАКОВА, Н. Н., Теории стереоэдров, *Тр. Матем. ин-та АН СССР*, 64 (1961) 28–51.
[4] ФЕДОРОВ, Е. С., *Симметрия правильных систем фигур*, С. Петербург, 1890.

Corrections to

B. M. PÖTSCHER, Some Results on ω_μ-Metric Spaces, *Annales Univ. Sci., Budapest*, Sect. Math., **25** (1982), 3–18.

On p. 4, line 1 replace "cofinality ω_μ" by "cofinality $\leqq \omega_\mu$". On p. 4 replace the second sentence of footnote 2 by "If X is discrete any linearly ordered base of $\mathcal{U}_d$ has cofinality 1 if $\omega_\mu > |X|$ and either 1 or ω_μ if $\omega_\mu \leq |X|$; and for every $\omega_\mu \leq |X|$ there is an ω_μ-metric d such that the cofinality of $\mathcal{U}_d$ is ω_μ. On p. 7, line 9 replace the symbol $\underset{\neq}{\supset}$ by $\underset{\neq}{\subset}$. On p. 16, line 22 replace "cofinality ω_0" by cofinality $\leq \omega_0$". Note also that some proofs (especially in sect. 3) do not cover the trivial discrete case as they stand, but nevertheless the corresponding statements are also true in this case.

INDEX

Address:
MATHEMATICAL INSTITUE L. EÖTVÖS UNIVERSITY
BUDAPEST, MÚZEUM KRT. 6–8.
H–1088

ISSN 0524—9007

Technikai szerkesztő:
DR. SCHARNITZKY VIKTOR
A kiadásért felelős: az Eötvös Loránd Tudományegyetem rektora
A kézirat nyomdába érkezett: 1984. május. Megjelent: 1986. augusztus
Terjedelem: 24 A/5 ív. Példányszám: 1000
Készült monó- és kéziszedéssel, íves magasnyomással,
az MSZ 5601–59 és MSZ 5602–55 szabványok szerint
84.548., Állami Nyomda, Budapest
Felelős vezető: Mihalek Sándor igazgató