Rostocker Mathematisches Kolloquium

Heft 16

WILHELM-PIECK-UNIVERSITÄT
ROSTOCK

ROSTOCKER MATHEMATISCHES KOLLOQUIUM

Heft 16

1981

Wilhelm-Pieck-Universität Rostock
Sektion Mathematik

Herausgeber: Der Rektor der Wilhelm-Pieck-Universität Rostock

Schriftleitung: Prof. Dr. Wolfgang Engel, Direktor der Sektion Mathematik

Prof. Dr. Gerhard Maeß, Schriftleiter
Dr. Werner Plischke, Lektor
Dorothea Meyer, Herstellung der Druckvorlage

Sektion Mathematik der
Wilhelm-Pieck-Universität Rostock,
DDR - 2500 Rostock, Universitätsplatz 1

Das Rostocker Mathematische Kolloquium erscheint in der Regel dreimal im Jahr und ist im Rahmen des Schriftentausches über die Universitätsbibliothek, Tauschstelle, DDR - 2500 Rostock, Universitätsplatz 5, zu beziehen.

Veröffentlicht durch die Abt. Wissenschaftspublizistik der Wilhelm-Pieck-Universität Rostock, DDR - 2500 Rostock, Vogelsang 13/14
Fernruf: 369 577
Leiter: Dipl.-Ges.-Wiss. Bruno Schrage
Genehmigungs-Nr. C 169/80
Druck: Ostsee-Druck Rostock, Werk II Ribnitz

https://doi.org/10.18453/rosdok_id00003818

Inhalt

		Seite
Gronau, Hans-Dietrich O. F.	The 2-(10,4,2) Designs	5
Gronau, Hans-Dietrich O. F.	Coverings of the complete (di-)graph with n vertices by complete bipartite (di-)graphs with n vertices II	11
Weber, Karl	Domination number for almost every graph	31
Wildenhain, Günther	Über punktweise Abschätzungen von Lösungen linearer elliptischer Differentialgleichungen beliebiger Ordnung	45
Wildenhain, Günther	Über die Darstellbarkeit polyharmonischer Funktionen in der Kugel	59
Berg, Lothar	Ein Grenzfall beim Goursatschen Anfangswertproblem	75
Polonijo, Mirko	Associativity and n-ary cyclic bisymmetry	81
Thielcke, Helmut	Rechentechnische Aspekte bei der numerischen Auswertung von mehrfachen Integralen	87
Stade, Carsten	Ein Verfahren mit quadratischer Konvergenzgeschwindigkeit zur simultanen Berechnung zweier Nullstellen beliebiger Ordnung	95
Schott, Dieter	Endlich erzeugte Projektionsverfahren zur Lösung linearer Gleichungen im Hilbertraum	103

Rostock. Math. Kolloq. 16, 5 - 10 (1981)

Hans-Dietrich O. F. Gronau

The 2-(10,4,2) Designs

A t-(v,k,λ) design on v symbols (say the numbers $1,2,\ldots,v$) has b distinct blocks of k distinct symbols with every t-element subset of $\{1,2,\ldots,v\}$ appearing in exactly λ blocks. If and only if there is a permutation of the symbols which has the property that every block B_i of a t-(v,k,λ) design D_1 is mapped into some block B_j of a t-(v,k,λ) design D_2, the designs D_1 and D_2 are called isomorphic. The problem of determination of all nonisomorphic designs with given (small) parameters and $\lambda = 1$ was studied by numerous authors; see Doyen and Rosa /1/. Motivated by practical interests Rasch and Herrendörfer /2/ extended this problem to designs with $\lambda \geq 2$ and small v, k, and t. Several papers published in this journal solved this problem for some parameter-combinations. One of the open problems with small parameters was that with $v = 10$, $k = 4$, $t = 2$, $\lambda = 2$. In this paper we present the following

<u>Theorem:</u> There are exactly 3 nonisomorphic 2-(10,4,2) designs. Representatives of the 3 classes of nonisomorphic 2-(10,4,2) designs are:[1]

$$M_1 = \begin{pmatrix} 1&1&1&1&2&2&3&1&1&2&2&3&3&4&4 \\ 2&2&3&4&3&4&4&5&8&5&7&5&6&6&5 \\ 3&7&6&5&5&6&7&6&9&6&9&7&8&7&8 \\ 4&8&9&0&0&9&8&7&0&8&0&9&0&0&9 \end{pmatrix},$$

$$M_2 = \begin{pmatrix} 1&1&1&1&2&2&3&1&1&2&2&3&3&4&4 \\ 2&2&3&4&3&4&4&5&8&5&7&5&6&6&5 \\ 3&5&6&7&7&6&5&6&9&6&9&7&8&7&8 \\ 4&0&9&8&8&9&0&7&0&8&0&9&0&0&9 \end{pmatrix},$$

1 We always give the designs in terms of 4×15 - matrices, where the columns are the blocks. As symbol in the designs we replace 0 by 10.

$$M_3 = \begin{bmatrix} 1&1&1&1&2&2&3&1&1&2&2&3&3&4&4 \\ 2&2&3&4&3&4&4&5&8&5&7&5&6&6&5 \\ 3&5&7&6&6&7&5&6&9&6&9&7&8&7&8 \\ 4&0&8&9&9&8&0&7&0&8&0&9&0&0&9 \end{bmatrix}.$$

It can easily be checked that a 2-(10,4,2) desgin consists of ($b =$) 15 blocks and every symbol occurs exactly ($r =$) 6 times.

<u>Proof of the Theorem:</u> We construct all 2-(10,4,2) designs successively. For a particular block of a 2-(10,4,2) design M let n_i, $0 \le i \le 4$, be the number of blocks that intersect it in exactly i symbols. Then $n_4 = 1$ and

$$\begin{aligned} n_0 + n_1 + n_2 + n_3 &= 14, \\ n_1 + 2n_2 + 3n_3 &= 20, \\ n_2 + 3n_3 &= 6. \end{aligned}$$

These equations have an unique solution in integers:

$$n_0 = 0,\ n_1 = 8,\ n_2 = 6,\ n_3 = 0.$$

W.l.o.g. we may assume that 1234 is one of the blocks of M. Using $\lambda = 2$ M has the structure

$$M = \left(\begin{array}{c|cccccc|cc|cc|cc|cc} 1 & 1&1&1&2&2&3 & 1&1 & 2&2 & 3&3 & 4&4 \\ \cline{8-15} 2 & 2&3&4&3&4&4 & & & & & & & & \\ \cline{2-7} 3 & & & & & & & \multicolumn{2}{c|}{A} & \multicolumn{2}{c|}{B} & \multicolumn{2}{c|}{C} & \multicolumn{2}{c}{D} \\ 4 & a&b&c&d&e&f & & & & & & & & \end{array}\right).$$

In all empty cells the numbers 5,6,7,8,9,0 have to be placed. We distinguish two cases.

- In every of the matrices A,B,C,D every symbol of $\{5,6,7,8,9,0\}$ occurs exactly once. - Case 1.
- There is a matrix, say A, containing an element, say 5, twice. - Case 2.

<u>Case 1:</u> W.l.o.g. $A = \begin{pmatrix} 5 & 8 \\ 6 & 9 \\ 7 & 0 \end{pmatrix}$. Each of the matrices B,C,D contains (at least) one of the pairs $\binom{5}{6}$, $\binom{5}{7}$, $\binom{6}{7}$. Noting $\lambda = 2$ and that the design, so far as constructed now, is invariant with respect to permutations τ_1, τ_2, where τ_1 and τ_2 are

permutations of $\{2,3,4\}$ and $\{8,9,0\}$, respectively. Hence we may assume w.l.o.g.

$$B = \begin{pmatrix} 5 & 7 \\ 6 & 9 \\ 8 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 5 & 6 \\ 7 & 8 \\ 9 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 6 & 5 \\ 7 & 8 \\ 0 & 9 \end{pmatrix}.$$

Counting the frequency of every pair yields that

$$\left[\begin{array}{cccccc} 1 & 1 & 1 & 2 & 2 & 3 \\ 2 & 3 & 4 & 3 & 4 & 4 \\ \hline a & b & c & d & e & f \end{array}\right]$$

has to contain the pairs $\binom{x}{y}$, $x \in \{1,2,3,4\}$, $y \in \{5,6,7,8,9,0\}$, exactly once and the pairs $\binom{5}{0}$, $\binom{6}{9}$, $\binom{7}{8}$ twice.

Thus, a = f, b = e, and c = d, and we get 6 possible continuations to construct a design.

a b c d e f	result
$\begin{matrix} 5 & 6 & 7 & 7 & 6 & 5 \\ 0 & 9 & 8 & 8 & 9 & 0 \end{matrix}$	M_2
$\begin{matrix} 5 & 7 & 6 & 6 & 7 & 5 \\ 0 & 8 & 9 & 9 & 8 & 0 \end{matrix}$	M_3
$\begin{matrix} 6 & 5 & 7 & 7 & 5 & 6 \\ 9 & 0 & 8 & 8 & 0 & 9 \end{matrix}$	isomorphic to M_3 by (34)(56)(90)
$\begin{matrix} 6 & 7 & 5 & 5 & 7 & 6 \\ 9 & 8 & 0 & 0 & 8 & 9 \end{matrix}$	isomorphic to M_2 by (34)(56)(90)
$\begin{matrix} 7 & 5 & 6 & 6 & 5 & 7 \\ 8 & 0 & 9 & 9 & 0 & 8 \end{matrix}$	isomorphic to M_2 by (23)(67)(89)
$\begin{matrix} 7 & 6 & 5 & 5 & 6 & 7 \\ 8 & 9 & 0 & 0 & 9 & 8 \end{matrix}$	M_1

Now we show that M_1, M_2, and M_3 are mutually nonisomorphic. For any element j and any design M_i let $M_i(j) = \{X: X \cup \{j\} \in M_i,\ j \notin X\}$. The $M_i(j)$'s are 1-(9,3,2) designs, obviously. Some of them are resolvable, i.e., $M_i(j)$ can be partitioned in two 1-(9,3,1) designs.

E.g. $M_1(1) = \begin{pmatrix} 2 & 2 & 3 & 4 & 5 & 8 \\ 3 & 7 & 6 & 5 & 6 & 9 \\ 4 & 8 & 9 & 0 & 7 & 0 \end{pmatrix}$ can be partitioned in the

1-(9,3,1) designs

$$\begin{pmatrix}2&5&8\\3&6&9\\4&7&0\end{pmatrix} \text{ and } \begin{pmatrix}2&3&4\\7&6&5\\8&9&0\end{pmatrix}, \text{ while } M_2(5) = \begin{pmatrix}1&3&1&2&3&4\\2&4&6&6&7&8\\0&0&7&8&9&9\end{pmatrix} \text{ is not}$$

partitionable. Obviously, for every design M_i the number of resolvable $M_i(j)$'s is invariant with respect to applying of permutations. Thus, we complete our proof giving the following table:

M_i	M_1	M_2	M_3
Number of resolvable $M_i(j)$'s	10	6	4

Case 2: Since $n_3 = 0$, the remaining elements of A have to be different and we may assume w. l. o. g.

$$A = \begin{pmatrix}5&5\\6&7\\8&9\end{pmatrix}.$$

If $x \in \{a,b,\dots,f,A,B,C,D\}$ then x_i denotes the number of elements i in the matrix x. Then

$$d_5 + e_5 + B_5 = 2,$$
$$d_5 + f_5 + C_5 = 2,$$
$$e_5 + f_5 + D_5 = 2,$$

and

$$d_5 + e_5 + f_5 + B_5 + C_5 + D_5 = 4.$$

W. l. o. g. we may assume $d_5 \geq e_5 \geq f_5$ and obtain $d_5 = e_5 = 1$, $f_5 = 0$, $B_5 = 0$, $C_5 = D_5 = 1$.

Consider (a,b,c). From A it follows immediately that (a,b,c) contains 6,7,8,9 once each and 0 twice. Since $n_3 = 0$, $\binom{6}{8}$ or $\binom{7}{9}$ cannot occur in (a,b,c) and we may assume w. l. o. g. $(a,b,c) = \begin{pmatrix}6&8&9\\7&0&0\end{pmatrix}$. In analogy to x_5 we get $d_0 = e_0 = f_0 = 0$, $B_0 = 2$, $C_0 = D_0 = 1$. Moreover, C and D contain the pair $\binom{5}{0}$.

Thus, $B_i = 1$ for $i = 6,7,8,9$ and in (d,e) there occur 8 and 9. If $(d,e) = \binom{5\ 5}{8\ 9}$ then the remaining empty triples in C and D are $\begin{pmatrix}6\\7\\9\end{pmatrix}$ and $\begin{pmatrix}6\\7\\8\end{pmatrix}$, respectively. But now we have the pair $\binom{6}{7}$ 3 times. Hence, $(d,e) = \binom{5\ 5}{9\ 8}$. Then

$$\begin{aligned} f_8 + C_8 &= 1, \\ f_8 + D_8 &= 1, \\ f_8 + C_8 + D_8 &= 2, \end{aligned}$$

and consequently, $f_8 = 0$, $C_8 = D_8 = 1$. Analogously, $f_9 = 0$, $C_9 = D_9 = 1$. Hence, $f = \binom{6}{7}$. Then $\binom{6}{7}$ does not occur in B. Since the pairs $\binom{6}{8}$ and $\binom{7}{9}$ exist in A, the matrices C and D contain the triples $\begin{pmatrix}7\\8\\9\end{pmatrix}$ and $\begin{pmatrix}6\\8\\9\end{pmatrix}$.

Applying the permutation (34)(67)(89) in the second case we may assume that C contains $\begin{pmatrix}7\\8\\9\end{pmatrix}$. Then it follows uniquely

$$M = \left(\begin{array}{c|cccccc|cc|cc|cc|cc} 1 & 1 & 1 & 1 & 2 & 2 & 3 & 1 & 1 & 2 & 2 & 3 & 3 & 4 & 4 \\ \hline 2 & 2 & 3 & 4 & 3 & 4 & 4 & 5 & 5 & 6 & 7 & 5 & 7 & 5 & 6 \\ 3 & 6 & 8 & 9 & 5 & 5 & 6 & 6 & 7 & 9 & 8 & 6 & 8 & 7 & 8 \\ 4 & 7 & 0 & 0 & 9 & 8 & 7 & 8 & 9 & 0 & 0 & 0 & 9 & 0 & 9 \end{array}\right).$$

M is isomorphic to M_2, e. g. by the permutation (1598742063),

q. e. d.

References

/1/ Doyen, J., and Rosa, A.: A Bibliography and Survey of Steiner Systems. Boll. Un. Mat. Ital. 7, 392 - 419 (1973)

/2/ Rasch, D., und Herrendörfer, G.: Über die Anzahl elementarer BUB in eingeschränkten (v,k)-Familien. Rostock. Math. Kolloq. 8, 71 - 82 (1978)

received: May 23, 1980

Author's address:

Dr. rer. nat. Hans-Dietrich O. F. Gronau
Wilhelm-Pieck-Universität Rostock
Sektion Mathematik
Universitätsplatz 1
DDR-2500 Rostock

Rostock. Math. Kolloq. 16, 11 - 30 (1981)

Hans-Dietrich O. F. Gronau

Coverings of the complete (di-)graph with n vertices by complete bipartite (di-)graphs with n vertices II

1. Introduction

This paper continues the investigations of part I /1/. More precisely, we complete here the proof of Theorem 5. First we repeat briefly some notations to make this paper selfcontained. Let n be an integer, $n \geq 2$. Let $B^*(V,W)$ denote complete bipartite digraphs, where the set V of vertices has size n, $W \subseteq V$, and the set of edges is $W \times (V - W)$. Let K_n^* denote the complete digraph with n vertices. A set $\mathfrak{M}$ of $B^*(V,W)$'s is called a critical covering of K_n^*, if and only if

$G = (V, \bigcup W \times (V \setminus W))$ - the union taken on all B^*'s of $\mathfrak{M}$ - is equal to K_n^* and no proper subset $\mathfrak{M}'$ of $\mathfrak{M}$ has that property.

Let $g^*(n)$ denote the maximum cardinality of a critical covering M of K_n^*.

We will number the vertices by 1, 2, ..., n and the digraphs of $\mathfrak{M}$ by 1, 2, ...

In /1/, Theorem 5, the author stated the

Theorem:

$$g^*(n) = h(n) = \begin{cases} 2n - 2 & \text{if } 2 \leq n \leq 6, \\ \left\lfloor \frac{n^2}{4} \right\rfloor^{1} & \text{if } n \geq 7. \end{cases}$$

There was shown $g^*(n) \geq h(n)$ by an example of a critical covering. Moreover, we introduced the concept of essential edges. Let $\mathfrak{M} = \{B_1^*(V,W_1), \ldots, B_t^*(V,W_t)\}$ be a maximal critical covering of K_n^*.

---

1 $\lfloor x \rfloor$ denotes the greatest integer not exceeding x.

We call a (directed) edge of a digraph $B^* \in \mathfrak{M}$ essential, if and only if this edge occurs only in this B^*.
Let T_i be the set of essential edges of the digraph $B_i^*(V,W_i)$, $i = 1, 2, \ldots, t$.
Since $\mathfrak{M}$ is critical, the sets T_i are nonempty and, obviously, the T_i's are mutually disjoint. If $R(T_1,\ldots,T_t)$ is any system of representatives, there are 2 possible cases:

Case 1: There is no pair $x,y \in V$ with $(x,y) \in R$ and $(y,x) \in R$.

Case 2: There is a pair $x,y \in V$ with $(x,y) \in R$ and $(y,x) \in R$.

If R belongs to the first case, then $t = h(n)$ or there is another system of representatives R' belonging to case 2.
This was proven in /1/.
This paper is reserved for the proof in case 2.
In section 2 we will give an equivalent description of our problems by means of matrices and some useful notations. In order to prove the Theorem we need some lemmata which will be given in sections 3, 4, 5, and 6. Finally in section 7 we prove the Theorem.

2. The matrix concept

To every covering $\mathfrak{M} = \{B_1^*(V,W_1), \ldots, B_t^*(V,W_t)\}$ we associate uniquely an $n \times t$ - matrix $M = (a_{ij})$ defined by

$$a_{ij} = \begin{cases} 1 & \text{if } i \notin W_j, \\ 0 & \text{if } i \in W_j. \end{cases}$$

If $a_{ij} = 0$ and $a_{kj} = 1$, $1 \leq i,k \leq n$, $i \neq k$, $1 \leq j \leq t$, we say, the j-th column of M contains the zop - 0,1-pair - (i,k).
At all there are $n(n-1)$ distinct zops. $\mathfrak{M}$ is a covering of K_n^* if and only if M contains all these zops. Analogously, a zop is called essential, iff it belongs only to one of the columns of M. Thus, iff $\mathfrak{M}$ is a critical covering, every column of M contains at least one essential zop.

A matrix M which is associated with a critical covering is called cc-matrix, i. e., M contains all possible zops and every column contains (at least) one essential zop. According to case 2, we may assume w. l. o. g., M contains the essential zops (1,2) and (2,1) and M has the structure:

$$M = \left(\begin{array}{c|cc|cc} 10 & \multicolumn{2}{c|}{000\ldots0} & \multicolumn{2}{c}{111\ldots1} \\ 01 & \multicolumn{2}{c|}{000\ldots0} & \multicolumn{2}{c}{111\ldots1} \\ \hline & & & & \\ N & M_2 & \multicolumn{1}{|c|}{N_2} & M_3 & \multicolumn{1}{|c}{N_3} \\ & & & & \end{array}\right) ,$$

where N has the form $N = (M_1 | N_1)$. Let the matrix $(M_1 | M_2 | M_3)$ be a cc-matrix with $n - 2$ rows. Such a submatrix exists certainly. N_2 and N_3 guarantee that M contains all zops (i,j) and (j,i), $i \in \{1,2\}$ and $j \in \{3,4,\ldots,n\}$. Every column of

$$\left(\begin{array}{c|c} 000\ldots0 & 111\ldots1 \\ 000\ldots0 & 111\ldots1 \\ \hline & \\ N_2 & N_3 \\ & \end{array}\right)$$

contains at least one of these zops as an essential zop which does not appear in other columns of M and, consequently, in

$$\left(\begin{array}{c|c} 000\ldots0 & 111\ldots1 \\ 000\ldots0 & 111\ldots1 \\ \hline & \\ M_2 & M_3 \\ & \end{array}\right) .$$

E. g., if

$$\left(\begin{array}{c} 000\ldots0 \\ 000\ldots0 \\ \hline \\ N_2 \\ \\ \end{array}\right)$$

contains the essential zop $(1,i)$, $3 \le i \le n$, the i-th[2] row of M_2 consists only of 0's.

---

2 The numbering of rows of submatrices are the same as in M, i. e., the numbers of rows of N_2 are 3, 4, ..., n.

3. M_2^0 and M_3^1

Let M_2^0 denote the set of numbers of those rows of M_2 which consist of 0's only. Analogously, let M_3^1 denote the set of numbers of those rows of M_3 which consist of 1's only. Obviously, there are not two distinct columns in

$$\begin{pmatrix} 000\ldots0 \\ \underline{000\ldots0} \\ \\ N_2 \end{pmatrix}$$

such that one has the essential zop (1,i) and the other has the essential zop (2,i) for some i. Hence it follows

<u>Lemma 1:</u>

(i) $|M_2^0| \geq |N_2|$,[3]

(ii) $|M_3^1| \geq |N_3|$.

<u>Lemma 2:</u> Let $M_2^0 \neq \emptyset$ and $M_3^1 \neq \emptyset$. Then

(i) $|M_2^0 \cap M_3^1| \leq 2$,

(ii) $|M_2^0 \cap M_3^1| = 2$ implies $|M_1| = 2$,

(iii) $|M_2^0 \cap M_3^1| = 1$ and $|M_2^0 \setminus M_3^1| \geq 1$ and $|M_3^1 \setminus M_2^0| \geq 1$ imply $|M_1| = 2$,

(iv) $|M_2^0 \setminus M_3^1| \geq 1$ or $|M_3^1 \setminus M_2^0| \geq 1$ implies $|M_1| \geq 1$.

3 $|M_i|$ and $|N_i|$ denote the numbers of columns of M_i and N_i, respectively.

Proof: (i) Consider the rows $M_2^0 \cap M_3^1$ of $(M_1|M_2|M_3)$. These rows form a matrix P. The columns of P, excepted those, deduced from M_1, consist only of 0's or only of 1's. Hence, the first two columns of P form a cc-matrix. It can be checked easily that P consists of at most 2 rows.

(ii) We note that $|M_1| \leq 2$ and that the only cc-matrix with 2 rows and at most 2 columns are $\binom{10}{01}$ and $\binom{01}{10}$.

(iii) Let $i \in M_2^0 \cap M_3^1$, $j \in M_2^0 \setminus M_3^1$, and $h \in M_3^1 \setminus M_2^0$. Then $i \neq j \neq h \neq i$ and $(M_2|M_3)$ does not contain the zops (i,j) and (h,i). Since $(M_1|M_2|M_3)$ is a cc-matrix, M_1 contains these both zops. A column cannot contain the zops (i,j) and (h,i), i. e. $|M_2| = 2$.

(iv) W. l. o. g. we may assume $|M_2^0 \setminus M_3^1| \geq 1$. Let $i \in M_2^0 \setminus M_3^1$ and $j \in M_3^1$. (M_3^1 is nonempty by the assumption.) Then $i \neq j$ and $(M_2|M_3)$ does not contain the zop (j,i). This zop must lie in M_1, i. e., $|M_1| \geq 1$, q. e. d.

4. M's with great N_2's

Lemma 3: If $|N_2| = n-2$, then $|M| \leq 2n-2$.

Proof: By Lemma 1 and $|M_2^0| \leq n-2$ it follows $|M_2^0| = n-2$ and $M_2 = \emptyset$. Furthermore, every column of

$$\begin{pmatrix} 000\ldots0 \\ 000\ldots0 \\ \hline N_2 \end{pmatrix}$$

contains at least one essential zop (i,j) with $i \in \{1,2\}$ and $j \in \{3,4,\ldots,n\}$. A column contains both or none of the zops $(1,i)$ and $(2,i)$. Hence, the zops $(1,i)$ and $(2,i)$ cannot be essential in different columns. Every column contains (at least) one essential zop (i,j), $i \in \{1,2\}$, $j \in \{3,4,\dots,n\}$, where different j's are associated to different columns.

Now $|N_2| = n-2$ implies that N_2 is (suitable ordered) the identity matrix. In

$$\left(\begin{array}{c|c} 10 & 000\dots0 \\ 01 & 000\dots0 \\ \hline N & N_2 \end{array}\right)$$

there do not appear at most the following zops: (i,j), $i \in \{3,4,\dots,n\}$, $j \in \{1,2\}$. Every column of

$$\left(\begin{array}{c|c} \multicolumn{2}{c}{111\dots1} \\ \multicolumn{2}{c}{111\dots1} \\ \hline M_3 & N_3 \end{array}\right)$$

containing the zop $(i,1)$ contains the zop $(2,i)$ too and conversely, for any $i \in \{3,4,\dots,n\}$. Since every column contains an essential zop, it follows

$$|(M_3|N_3)| \le n-2$$

and

$$|M| = |N| + |M_2| + |N_2| + |M_3| + |N_3| \le 2n-2,$$

q. e. d.

Lemma 4: If $|N_2| = n-3$, then $|M| \le 2n-2$.

Proof: By Lemma 1 we obtain $|M_2^0| \ge n-3$. We distinguish two cases.

1. $|M_2^0| = n-2$, i. e. $M_2 = \emptyset$. We assume that the n-th row of N_2 contains exactly d 0's. Then

$$M' = \left(\begin{array}{c|c} 10 & 000\dots0 \\ 01 & 000\dots0 \\ \hline N & N_2 \end{array}\right)$$

has w. l. o. g. the following structure

$$\left(\begin{array}{c|l}
10 & 000\ldots000\ldots0 \\
01 & 000\ldots000\ldots0 \\
\hline
 & 100\ldots000\ldots0 \\
 & 010\ldots000\ldots0 \\
 & 001\ldots000\ldots0 \\
 & \cdots\cdots\cdots \\
 & \cdots\cdots\cdots \\
N & 000\ldots100\ldots0 \\
 & 000\ldots010\ldots0 \\
 & 000\ldots001\ldots0 \\
 & \cdots\cdots\cdots \\
 & \cdots\cdots\cdots \\
 & 000\ldots000\ldots1 \\
 & \underbrace{000\ldots0}_{d}\underbrace{11\ldots1}_{n-d-3}
\end{array}\right).$$

M' does not contain at most the following zops:

a) $d = n - 3$: (i,n) with $i \in \{3,4,\ldots,n-1\}$ and
$(i,1)$, $(i,2)$ with $i \in \{3,4,\ldots,n\}$.

b) $d = n - 4$: $(n-1,n)$, $(n,n-1)$ and
$(i,1)$, $(i,2)$ with $i \in \{3,4,\ldots,n\}$.

c) $d \leq n - 5$: (n,i) with $i \in \{d+3, d+4,\ldots,n-1\}$ and
$(i,1)$, $(i,2)$ with $i \in \{3,4,\ldots,n\}$.

Every column of

$$\left(\begin{array}{c|c}
\multicolumn{2}{c}{111\ldots1} \\
\multicolumn{2}{c}{111\ldots1} \\
\hline
M_3 & N_3
\end{array}\right)$$

containing the zop $(i,1)$ contains $(i,2)$ certainly. We discuss the cases separately.

a) Every column containing the zop (i,n) contains also the zops $(i,1)$ and $(i,2)$. Hence $|(M_3|N_3)| \leq n-2$ and $|M| \leq 2n-3$.

b) Every column containing the zop $(n-1,n)$ contains also the zop $(n-1,1)$ (and $(n-1,2)$), while every column containing the zop $(n,n-1)$ contains the zop $(n,1)$ (and $(n,2)$). Thus we obtain $|(M_3|N_3)| \leq n-2$ and consequently $|M| \leq 2n-3$.

c) Let $s_1, s_2, \ldots, s_p$ be exactly those columns of

$$\left(\begin{array}{c|c} \multicolumn{2}{c}{111\ldots1} \\ \multicolumn{2}{c}{111\ldots1} \\ \hline M_3 & N_3 \end{array}\right)$$

which contain at least one essential zop (n,i) with $i \in \{d+3, d+4, \ldots, n-1\}$. We are going to show that $S \doteq (s_1, s_2, \ldots, s_p)$ contains at least p mutually different zops $(i,1)$, $i \in \{d+3, d+4, \ldots, n-1\}$. Then we have $|(M_3|N_3)| \leq n-2$ and $|M| \leq 2n-3$.

c1) $p = 0$. Then nothing is to prove.

c2) $p = 1$. s_1 contains an essential zop (n,i) with $i \in \{d+3, d+4, \ldots, n-1\}$. This column contains also the zop $(n,1)$.

c3) $p \geq 2$. Suppose that s_j contains the essential zop (n,i_j) with $i_j \in \{d+3, d+4, \ldots, n-1\}$, $i_j \neq i_{j'}$ for $j \neq j'$ and $j,j' = 1, 2, \ldots, p$.
Since these zops are essential in M, S has the following structure:
The n-th row of S consists of 0's only. The i_j-th row, $j = 1, 2, \ldots, p$, of S has exactly in the j-th column a 1 and 0's in all other columns. Since the first (and second) row of S consists of 1's only, S contains all p zops $(i_j,1)$ with $j = 1, 2, \ldots, p$.

2. $|M_2^0| = n-3$. Then M_2 contains at most one row which consists of 0's only. Since no column appears twice in M and M does not contain a column consisting of 0's only, M_2 has exactly one column, and we may assume w. l. o. g. that M' has the following structure

$$\left(\begin{array}{c|c|c}
10 & 000\ldots000\ldots0 & 0 \\
01 & 000\ldots000\ldots0 & 0 \\
\hline
 & 100\ldots000\ldots0 & 0 \\
 & 010\ldots000\ldots0 & 0 \\
 & 001\ldots000\ldots0 & 0 \\
 & \ldots\ldots\ldots\ldots & . \\
 & \ldots\ldots\ldots\ldots & . \\
 & 000\ldots100\ldots0 & 0 \\
 & 000\ldots010\ldots0 & 0 \\
 & 000\ldots001\ldots0 & 0 \\
 & \ldots\ldots\ldots\ldots & . \\
 & \ldots\ldots\ldots\ldots & . \\
 & 000\ldots000\ldots1 & 0 \\
 & \underbrace{000\ldots}_{d}\underbrace{011\ldots1}_{n-d-3} & 1
\end{array}\right) .$$

M' does not contain at most the following zops:
(n,i) with $i \in \{d+3, d+4, \ldots, n-1\}$ and $(i,1)$, $(i,2)$ with $i \in \{3,4,\ldots,n\}$.
Also in this case at most those zops do not appear in M' which do not appear in M' in the first case, and we get analogously $|(M_3|N_3)| \leq n-2$ and by $|M_2| = 1$ finally $|M| \leq 2n-2$,

q. e. d.

5. A lemma

Lemma 5: Let M be a cc-matrix with n rows. Let M' be a matrix which is obtained from M by omitting some columns. If M' has exactly $n-l$ rows consisting of 0's only, then

$$|M'| \leq \begin{cases} 1 & \text{if } l = 1, \\ g^*(l) & \text{if } l \geq 2. \end{cases}$$

Proof: We prove the lemma by induction on l.
1. $l = 1$. Since M contains mutually different columns only (and consequently M' too) and no column consisting of 0's only, M' can consist of only one column. This column contains exactly one 1.

2. $l \geq 2$. Every column of M (and consequently M' too) contains at least one essential zop (with respect to M).

a) Let each column of M' contain an essential zop (i,j) with $i,j \in \{n-l+1, n-l+2, \ldots, n\}$. (W. l. o. g. we may assume that the rows $1,2,\ldots,n-l$ of M' consist of 0's only.) Let M'' denote the matrix which is obtained from M' by omitting the rows $1, 2, \ldots, n-l$. Analogously let M''' denote the matrix which is obtained from M by omitting the rows $1, 2, \ldots, n-l$. Certainly, M'' is a submatrix of M'''. M''' is a matrix which contains all possible zops (on the set $\{n-l+1, n-l+2, \ldots, n\}$, of course). Hence M''' contains a cc-matrix as a submatrix (getting by omitting of columns, no rows). Since every column of M'' contains an essential zop with respect to M''', all columns of M'' are contained in the cc-matrix which is a submatrix of M'''. Thus

$$|M'| = |M''| \leq g^*(l).$$

b) There is a column of M' having an essential zop (i,j) with $i \in \{1,2,\ldots,n-l\}$ and $j \in \{n-l+1, n-l+2, \ldots, n\}$. (W. l. o. g. we assume also in this case that the first $n-l$ rows consist of 0's only.) W. l. o. g. let $j = n - l + 1$. Then the $(n-l+1)$-th row of M' consists of exactly one 1 which is w. l. o. g. in the first column of M'. We omit the first column of M' and obtain a matrix M'''' having at least $n - l + 1$ rows consisting of 0's only. By the hypothesis it follows

$$|M'| = 1 + |M''''| \leq 1 + \begin{cases} 1 & \text{if } l = 2, \\ g^*(l-1) & \text{if } l \geq 3. \end{cases}$$

$g^*(2) = 2$ can be verified easily. We note $g^*(l-1) + 1 \leq g^*(l)$, since, if S' is a cc-matrix with $l - 1$ rows and $|S'| = g^*(l-1)$, the matrix

$$S'' = \left(\begin{array}{c|c} 1 & 000\ldots0 \\ \hline 0 & \\ \vdots & S' \\ 0 & \end{array}\right)$$

is a cc-matrix with l rows and consequently $|S''| \leq g^*(l)$. Hence, $|M'| \leq g^*(l)$ is proven for $l \geq 2$,

q. e. d.

6. A lemma on h(n)

The function $h(n)$ is defined for $n \geq 2$ in the Theorem. Additionally we set $h(1) = 1$.

Lemma 6: Let a and b be integers with $a \geq b \geq 2$. Then

$$h(a+1) + h(b-1) \geq h(a) + h(b).$$

Proof: We distinguish 4 cases.

1. $a \geq b \geq 8$. Then, using $\frac{a^2-1}{4} \leq \left\lfloor \frac{a^2}{4} \right\rfloor \leq \frac{a^2}{4}$,

$$h(a+1) + h(b-1) - h(a) - h(b)$$

$$\geq \frac{(a+1)^2-1}{4} + \frac{(b-1)^2-1}{4} - \frac{a^2}{4} - \frac{b^2}{4} = \frac{a-b}{2} \geq 0.$$

2. $a \geq 7 \geq b \geq 3$. Then

$$h(a+1) + h(b-1) - h(a) - h(b)$$

$$\geq \frac{(a+1)^2-1}{4} - \frac{a^2}{4} + (2(b-1)-2) - (2b-2) = \frac{a}{2} - 2 > 0,$$

noting $\left\lfloor \frac{7^2}{4} \right\rfloor = 2 \cdot 7 - 2$.

3. $6 \geq a \geq b \geq 3$. Then

$$h(a+1) + h(b-1) - h(a) - h(b)$$

$$= 2(a+1)-2 + 2(b-1)-2 - (2a-2) - (2b-2) = 0.$$

4. $a \geq b = 2$.

$$h(a+1) + h(b-1) - h(a) - h(b) = h(a+1) - h(a) - 1 \geq 0,$$

obviously,

q. e. d.

7. Proof of the Theorem

W. l. o. g. we may assume $|N_2| \geq |N_3|$.

The cases with $n \leq 8$ will be discussed separately. For all other n we prove the Theorem by induction on n.

1. $n = 2$. $g^*(2) = 2$ as noted at the end of section 5.
2. $n = 3$. Since $(M_2|N_2)$ and $(M_3|N_3)$ consist of at most one column, respectively, by arguments which we used above, it follows $|M| \leq 4$ (i. e., $g^*(3) = 4$).
3. $n = 4$. If $|N_2| = 2$ or 1, then by Lemma 3 or 4, respectively, $|M| \leq 2 \cdot 4 - 2 = 6$. If $|N_2| = 0$, then $|N_3| = 0$ and by $|(M_2|M_3)| \leq g^*(2) = 2$ it follows $|M| \leq 4$.
4. $n = 5$. $|N_2| \geq 2$ implies $|M| \leq 2n - 2$ accordingly to Lemma 3 and 4. If $|N_2| \leq 1$, then $|N_3| \leq 1$ and $|M| = 2 + |(M_2|M_3)| + |N_2| + |N_3| \leq g^*(3) + 4 = 8$.
5. $n = 6$. $|N_2| \geq 3$ implies $|M| \leq 2n - 2$ accordingly to Lemma 3 and 4. Now let $|N_2| \leq 2$. We distinguish 3 cases.

a) $|N_2| = |N_3| = 2$. Then $|M_2^0| \geq 2$ and $|M_3^1| \geq 2$, i. e., M_2 has at most 2 rows which do not consist of 0's only. By Lemma 5 we obtain $|M_2| \leq g^*(2) = 2$. Analogously, $|M_3| \leq 2$ and consequently $|M| \leq 10$.

b) $|N_2| = 2$, $|N_3| = 1$. Then $|M_2^0| \geq 2$ and $|M_3^1| \geq 1$ by Lemma 1 and furthermore $|M_1| \geq 1$ by Lemma 2. Hence $|M| = 2 + |(M_2|M_3)| + 3 \leq 5 + g^*(4) - |M_1| \leq 10$.

c) $|N_2| + |N_3| \leq 2$. Then $|M| = 2 + |(M_2|M_3)| + |N_2| + |N_3| \leq 4 + g^*(4) = 10$.

6. $n = 7$. $|N_2| \geq 4$ implies $|M| \leq 2n - 2$ accordingly to Lemma 3 and 4. We distinguish the following 6 cases.

a) $|N_2| = |N_3| = 3$. Then $|M_2^0| \geq 3$, $|M_3^1| \geq 3$ by Lemma 1 and $|M_2| \leq g^*(2) = 2$ by Lemma 5. In analogy it holds $|M_3| \leq 2$. Thus, $|M| = 2 + |M_2| + |M_3| + |N_2| + |N_3| \leq 12$.

b) $|N_2| = 3$, $|N_3| = 2$. Then $|M_2^0| \geq 3$ and $|M_3^1| \geq 2$ by Lemma 1 and $|M_2| \leq g^*(2) = 2$ as well as $|M_3| \leq g^*(3) = 4$ by Lemma 5. Accordingly to the last case $|M| \leq 13$ follows immediately. We will show that $|M| = 13$ is impossible. Assume the contrary, i. e., there is a cc-matrix with 7 rows and 13 columns. Then in all inequalities above equality must hold, i. e.,

$|N_2| = |M_2^0| = 3$, $|N_3| = |M_3^1| = 2$, and $|M_2| = 2$, $|M_3| = 4$.

W. l. o. g. let $M_3^1 = \{3,4\}$. Every column of M_3 contains at least one essential zop (i,j) with $i,j \in \{3,4,\ldots,7\}$, $i \neq j$. If a column has only essential zops (i,j) with $i \in \{5,6,7\}$ and $j \in \{3,4\}$, $|M_3| \leq 1 + g^*(2) = 3$ follows by Lemma 5, contradicting our assumption.

Hence every column of M_3^* has an essential zop (with respect to M), where M_3^* denotes the matrix consisting of the last three rows of M_3. Analogously let M^* denote the matrix consisting of the last three rows of M. Since M^* is a matrix containing all possible zops, it contains a cc-matrix as submatrix with 3 rows. Let $\overline{M}$ denote this cc-matrix. M_3^* is a submatrix of $\overline{M}$, since M_3^* has an essential zop in each column. By $4 = |M_3| = |M_3^*| \leq |\overline{M}| \leq g^*(3) = 4$ it follows that M_3^* is a cc-matrix with 3 rows. The only cc-matrix with 3 rows and 4 columns is (applying suitable permutations of rows and columns):

$$\begin{pmatrix} 0101 \\ 1001 \\ 0110 \end{pmatrix} .$$

Consequently $(M_3|N_3)$ has the structure

$$\left(\begin{array}{c|c} 1111 & 01 \\ 1111 & 10 \\ \hline 0101 & .. \\ 1001 & .. \\ 0110 & .. \end{array}\right) .$$

Consider all zops (i,j) with $i,j \in \{3,4,\dots,7\}$, $i \neq j$. The only ones which could not appear in $(M_3|N_3)$ are $(3,5)$, $(3,6)$, $(3,7)$, $(4,5)$, $(4,6)$, and $(4,7)$.

Since the columns s_1 and s_2 of M_2 have essential zops (i,j) with $i,j \in \{3,4,\dots,7\}$, $i \neq j$, these ones must occur in the 6 cases given above.

b1) $(3,i) \in s_1$, $(4,i) \in s_2$ with $i \in \{5,6,7\}$.

W. l. o. g. let $i = 5$. Noting that the zops $(3,i)$ and $(4,i)$ are essential we obtain $M_2 = \begin{bmatrix} 01 \\ 10 \\ 11 \\ \cdot\cdot \\ \cdot\cdot \end{bmatrix}$ which contradicts $|M_2^0| = 3$.

b2) $(i_1,j_1) \in s_1$, $(i_2,j_2) \in s_2$ with $i_1,i_2 \in \{3,4\}$, $j_1,j_2 \in \{5,6,7\}$ and w. l. o. g. $j_1 < j_2$. $|M_2^0| = 3$ and the fact that (i_1,j_1) and (i_2,j_2) are essential zops imply

$$M_2 = \begin{bmatrix} 00 \\ 00 \\ 10 \\ 01 \\ 00 \end{bmatrix} \text{ if } \begin{cases} j_1 = 5 \\ j_2 = 6 \end{cases}, \quad M_2 = \begin{bmatrix} 00 \\ 00 \\ 10 \\ 00 \\ 01 \end{bmatrix} \text{ if } \begin{cases} j_1 = 5 \\ j_2 = 7 \end{cases},$$

and

$$M_2 = \begin{bmatrix} 00 \\ 00 \\ 00 \\ 10 \\ 01 \end{bmatrix} \text{ if } \begin{cases} j_1 = 6 \\ j_2 = 7 \end{cases}.$$

Since M_3 contains the essential zops $(5,6)$, $(5,7)$, $(6,5)$, and $(7,5)$, no one of these zops can occur in M_2, i.e., no one of the 3 matrices described above is possible. Hence, also in this case $|M| \leq 12$ is proven.

c) $|N_2| = 3$, $|N_3| \leq 1$. For convenience, in this case and only in this case, we assume that the numbers of columns of

M_2 and N_2 are exchanged[4], i. e., N_2 has the numbers of columns 3,4, and 5. Furthermore, we may assume that N_2 has the following structure:

$$\begin{bmatrix} 100 \\ 010 \\ 001 \\ \cdots \\ \cdots \end{bmatrix}.$$

Let $i \in \{6,7\}$. Let d_i denote the set of numbers of columns (with respect to M) which contain a "0" in the i-th row of N_2. Then $0 \leq |d_i| \leq 3$ and $d_i \subseteq \{3,4,5\}$. Every column of $(M_2|M_3)$ contains at least one essential zop (i,j) with $i,j \in \{3,4,\ldots,7\}$ and $i \neq j$.
Altogether there are exactly $5 \cdot 4 = 20$ different zops (i,j) with $i,j \in \{3,4,\ldots,7\}$ and $i \neq j$.
Now we determine the number $a(d_6,d_7)$ of different zops which are not contained in N_2. These ones cannot be essential zops of other columns. Thus,

$$|(M_2|M_3)| \leq 20 - a(d_6,d_7). \tag{1}$$

N_2 contains

1. all 6 zops (i,j) with $i,j \in \{3,4,5\}$, $i \neq j$,
2. the zops (6,v) with $v \in d_6$ and (7,v) with $v \in d_7$,
3. the zops (3,6), (4,6), (5,6) if $|d_6| \leq 1$,
   $(v_1,6)$, $(v_2,6)$ with $v_1,v_2 \in d_6$ if $|d_6| = 2$ and
   (3,7), (4,7), (5,7) if $|d_7| \leq 1$,
   $(v_1,7)$, $(v_2,7)$ with $v_1,v_2 \in d_7$ if $|d_7| = 2$.
4. at least the zop (6,7) or (7,6) if $|d_6| \neq |d_7|$.

Obviously, all these zops are mutually different.
W. l. d. g. let $|d_6| \leq |d_t|$. For all possible combinations of $|d_6|$ and $|d_7|$ we give a lower bound for $a(d_6,d_7)$ and using (1) an upper bound for $|(M_2|M_3)|$.

---

4 In this way the numbers of columns are constant; $|M_2|$ is not fixed.

Furthermore there are listed the numbers of zops which the sets 1. – 4. contribute in each case.

$\lvert d_6\rvert$	$\lvert d_7\rvert$	1.	2.	3.	4.	$a(d_6,d_7) \geq$	$\lvert(M_2\lvert M_3)\rvert \leq$
0	0	6	0	6	0	12	8
0	1	6	1	6	1	14	6
0	2	6	2	5	1	14	6
0	3	6	3	3	1	13	7
1	1	6	2	6	0	14	6
1	2	6	3	5	1	15	5
1	3	6	4	3	1	14	6
2	2	6	4	4	0	14	6
2	3	6	5	2	1	14	6
3	3	6	6	0	0	12	8

For $(\lvert d_6\rvert, \lvert d_7\rvert) \neq (0,0), (0,3), (3,3)$ we obtain

$$\lvert M\rvert = 2 + \lvert(M_2\lvert M_3)\rvert + \lvert N_2\rvert + \lvert N_3\rvert \leq 12,$$

which proves our statement.
Now we consider the 3 remaining cases.

c1) $\lvert d_6\rvert = \lvert d_7\rvert = 0$. Every column of $(M_2\lvert M_3)$ contains at least one essential zop of $\{(6,3), (7,3), (6,4), (7,4), (6,5), (7,5), (6,7), (7,6)\}$. Hence, $\lvert(M_2\lvert M_3)\rvert \leq 8$. If $\lvert(M_2\lvert M_3)\rvert \leq 6$, the statement follows immediately. Let $\lvert(M_2\lvert M_3)\rvert \geq 7$. Then at most one of the eight zops given above can be missed.
W. l. o. g. we may assume that at most the zop $(7,5)$ or $(7,6)$ does not occur in $(M_2\lvert M_3)$. Either there is no column in $(M_2\lvert M_3)$ which contains $(6,3)$ as well as $(7,3)$ and no column in $(M_2\lvert M_3)$ which contains $(6,4)$ as well as $(7,4)$, or there are two (not necessarily different) columns s_1 and s_2 in $(M_2\lvert M_3)$ with $(6,3), (7,3) \in s_1$ and $(6,4), (7,4) \in s_2$. In the first case the column containing $(6,3)$ contains $(6,7)$ and the column containing $(7,3)$ contains $(7,6)$. In the second case the columns s_i contain at least 2 of the 8 zops given above. Hence, $\lvert(M_2\lvert M_3)\rvert \leq 6$. Thus, in both cases it follows $\lvert M\rvert \leq 12$.

c2) $|d_6| = 0$, $|d_7| = 3$. Every column of $(M_2|M_3)$ contains at least one essential zop of $\{(6,3), (6,4), (6,5), (3,7), (4,7), (5,7), (6,7)\}$. If $(6,7)$ does not occur in a column of $(M_2|M_3)$, $|(M_2|M_3)| \leq 6$ implies the desired result. If $(6,7)$ is contained in a column of $(M_2|M_3)$, this column contains $(6,3)$ or $(3,7)$, i. e., $|(M_2|M_3)| \leq 6$ and $|M| \leq 12$.

c3) $|d_6| = |d_7| = 3$. Every column of $(M_2|M_3)$ contains at least one essential zop of $\{(3,6), (3,7), (4,6), (4,7), (5,6), (5,7), (6,7), (7,6)\}$. In analogy to the case $|d_6| = |d_7| = 0$ we obtain $|M| \leq 12$.

d) $|N_2| = 2$, $|N_3| = 2$. Then $|M_2^0| \geq 2$ and $|M_3^1| \geq 2$ and it follows $|M_2| \leq 4$ and $|M_3| \leq 4$ by Lemma 5. Then

$$|M| = 2 + |M_2| + |M_3| + 4 \leq 14.$$

If $|M_2| + |M_3| \leq 6$, then $|M| \leq 12$. In all what follows let $|M_2| + |M_3| \geq 7$, i. e., we may assume $|M_3| = 4$. Furthermore we may assume that $(M_3|N_3)$ has the same structure as in case 6.b), p. 23

$$\left[\begin{array}{c|c} 1111 & 01 \\ 1111 & 10 \\ \hline 0101 & .. \\ 1001 & .. \\ 0110 & .. \end{array}\right].$$

Since M_3 has the essential zops $(5,6)$, $(6,5)$, $(5,7)$, and $(7,5)$, all columns of M not belonging to M_3 are $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ or $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

In $(M_3|N_3)$ at most the following zops do not occur: $(3,5)$, $(3,6)$, $(3,7)$, $(4,5)$, $(4,6)$, and $(4,7)$. Every column of M_2 has at least one of these zops as an essential zop, consequently. Using the remarks above, it follows immediately that every column containing $(3,5)$ also contains $(3,6)$ and $(3,7)$, etc. The same argument is true concerning the zops $(4,i)$. Hence, $|M_2| \leq 2$, which contradicts $|M_2| + |M_3| \geq 7$.

Thus, also in this case $|M| \leq 12$ is proven.

e) $|N_2| = 2$, $|N_3| = 1$. Then $|M_2^0| \geq 2$ and $|M_3^1| \geq 1$. Then $|M_1| \geq 1$ by Lemma 2. Finally,

$|M| = 2 + |(M_2|M_3)| + 3 \leq 5 + g^*(5) - |M_1| \leq 12.$

f) $|N_2| + |N_3| \leq 2$. Then it follows immediately

$|M| = 2 + |(M_2|M_3)| + |N_2| + |N_3| \leq 4 + g^*(5) = 12.$

7. $n = 8$. If $|N_2| \geq 5$, then

$|M| \leq 2 \cdot 8 - 2 = 14 < 16$

by Lemma 3 and Lemma 4.

We distinguish the following cases now.

a) $|N_2| = 4$, $3 \leq |N_3| \leq 4$. Then $|M_2^0| \geq 4$ and $|M_3^1| \geq 3$.
It follows $|M_2| \leq g^*(2) = 2$ and $|M_3| \leq g^*(3) = 4$ by Lemma 5.
Hence, $|M| = 2 + |M_2| + |M_3| + |N_2| + |N_3| \leq 16.$

b) $|N_2| = 4$, $|N_3| = 2$. In analogy to the preceeding case we get $|M_3| \leq g^*(4) = 6$ and

$|M| = 2 + |M_2| + |M_3| + |N_2| + |N_3| \leq 16.$

c) $|N_2| = 4$, $|N_3| = 1$. Then $|M_2^0| \geq 2$ and $|M_3^1| \geq 1$. Hence, $|M_1| \geq 1$ by Lemma 2. Thus,

$|M| = 2 + |M_2| + |M_3| + |N_2| + |N_3| \leq 7 + g^*(6) - |M_1| \leq 16.$

d) $|N_2| = 3$, $|N_3| = 3$. In analogy to case 7. a) we have

$|M| \leq 2 + 2 \cdot g^*(3) + 2 \cdot 3 = 16.$

e) $|N_2| = 3$, $|N_3| = 2$. Then $|M_1| \geq 1$ by Lemma 2 and furthermore $|M| \leq 2 + g^*(6) - |M_1| + 3 + 2 \leq 16.$

f) $|N_2| + |N_3| \leq 4$. It follows immediately

$|M| = 2 + |(M_2|M_3)| + |N_2| + |N_3| \leq 6 + g^*(6) = 16.$

8. Up to here $g^*(n) = h(n)$ is proven for $n \leq 8$. Now we complete the proof by induction. It is sufficient to make the

induction step for $n \geq 9$. Let $n \geq 9$. Then for every cc-matrix M with n rows:

$$\begin{aligned} |M| &= 2 + |(M_2|M_3)| + |N_2| + |N_3| \\ &\leq 2 + g^*(n-2) + |N_2| + |N_3| - |M_1|. \end{aligned} \tag{2}$$

Using

$\left\lfloor\frac{(n-2)^2}{4}\right\rfloor + (n-1) = \left\lfloor\frac{n^2}{4}\right\rfloor$, (2), and $g^*(n-2) = h(n-2)$ it is sufficient to prove

$$|N_2| + |N_3| - |M_1| \leq n - 3. \tag{3}$$

By $|N_2| \leq |M_2^0|$, $|N_3| \leq |M_3^1|$, $|M_2^0 \cap M_3^1| \leq 2$, and $|M_2^0 \cup M_3^1| \leq n-2$ we obtain $|N_2| + |N_3| \leq n$.

W. l. o. g. let $|N_2| \geq |N_3|$.

We distinguish 4 cases.

a) $|N_2| + |N_3| \leq n - 3$. Then (3) follows immediately.

b) $|N_2| + |N_3| = n - 2$. If $|N_2| = n - 2$, then $|M| \leq 2n - 2 < h(n)$ by Lemma 3. If $|N_2| < n - 2$, then $|N_3| \geq 1$ and $|M_1| \geq 1$ by Lemma 2 and consequently (3).

c) $|N_2| + |N_3| = n - 1$. If $|N_2| = n - 2$, then $|M| \leq 2n - 2 < h(n)$ by Lemma 3. If $|N_2| < n - 2$, then $|N_3| \geq 2$ and either $|M_2^0 \cap M_3^1| = 2$ or $|M_2^0 \cap M_3^1| = 1$. In the last case we have additionally $|M_2^0 \setminus M_3^1| \geq 1$ and $|M_3^1 \setminus M_2^0| \geq 1$. In both cases it follows $|M_1| = 2$ by Lemma 2. Hence, (3) is true.

d) $|N_2| + |N_3| = n$. If $|N_2| = n - 2$, then $|M| \leq 2n - 2 < h(n)$ by Lemma 3.

Let $|N_2| \leq n - 3$. Then also $|N_3| \leq n - 3$. Since

$$\begin{aligned} n &= |N_2| + |N_3| \leq |M_2^0| + |M_3^1| \\ &= |M_2^0 \cup M_3^1| + |M_2^0 \cap M_3^1| \leq n - 2 + 2 = n \end{aligned}$$

we get $|N_2| = |M_2^0|$ and $|N_3| = |M_3^1|$. By the induction assumption and Lemma 5 we obtain

$$|M_2| \leq h(n-2-|N_2|) \text{ and } |M_3| \leq h(n-2-|N_3|).$$

Finally,

$$\begin{aligned} |M| &= 2 + |M_2| + |M_3| + |N_2| + |N_3| \\ &\leq 2 + h(n-2-|N_2|) + h(n-2-|N_3|) + n. \end{aligned} \tag{4}$$

By $n-3 \geq |N_2| \geq |N_3|$ and $1 \leq n-2-|N_2| \leq n-2-|N_3|$ it follows by repeated application ($(n-2-|N_2|-1)$-times) of Lemma 6 to (4):

$$|M| \leq 2 + h(1) + h(n-5) + n = \left\lfloor\frac{(n-5)^2}{4}\right\rfloor + n + 3 \leq \left\lfloor\frac{n^2}{4}\right\rfloor = h(n),$$

q. e. d.

Reference

/1/ Gronau, H.-D. O. F.: Coverings of the complete (di-)graph with n vertices by complete bipartite (di-)graphs with n vertices I. Discrete Math., to appear

received: July 15, 1980

Author's address:

Dr. Hans-Dietrich O. F. Gronau
Wilhelm-Pieck-Universität Rostock
Sektion Mathematik
Universitätsplatz 1
DDR-2500 Rostock

Rostock. Math. Kolloq. 16, 31 - 43 (1981)

Karl Weber

Domination number for almost every graph

Abstract: Best possible bounds for the independence number and the domination number for almost every graph are established. The possible cardinalities of independent dominating sets (maximal independent sets) are determined.

1. Introduction and terminology

We consider finite graphs with neither loops nor multiple edges. The set of vertices and edges of a graph G are denoted by V and E, respectively.
A set $M \subseteq V$ is called a dominating set if each vertex of $V-M$ is adjacent to some vertex of M. The domination number $\gamma(G)$ is the smallest cardinality of a dominating set of G. A set M is called independent if no two vertices of M are adjacent. The largest number of vertices in an independent set of G is called the independence number $\beta(G)$. The independent domination number i(G) is the smallest number of vertices in an independent dominating (or maximal independent) set of G.[1]
Let $\mathcal{G}_n$ be the set consisting of all 2^k, $k = \binom{n}{2}$, graphs with n fixed and labeled vertices. We assume that each graph $G_n \in \mathcal{G}_n$ occurs with the same probability $P(G_n) = 1/2^{\binom{n}{2}}$. This means we consider the discrete probability space defined over $\mathcal{G}_n$ containing all random graphs, in which each edge occurs with probability 1/2 independently of the presence or absence of any other edges. The parameters γ, B and i are random variables defined on $\mathcal{G}_n$. One says that a certain statement holds for

1 A motivation for investigations of graph-theoretic parameters concerning domination and independence one can find in many papers, for example in /2/.

almost every graph if as $n \to \infty$ the probability of the set of graphs, for which the assertion fails, tends to 0. Similarly we will say that a statement holds for almost every G_{n_i} if this limit condition is fulfilled for the subsequence $n_1 < n_2 < n_3 < \dots$ of natural numbers as $i \to \infty$.[2]

2. Preliminaries

We use Landau's notation $O(f(n))$ for a term that, when diveded by $f(n)$, remains bounded as $n \to \infty$. Similarly $o(f(n))$ denotes a term that, when diveded by $f(n)$, tends to 0 as $n \to \infty$. We write $f(n) \asymp g(n)$ if $f(n) = O(g(n))$ and $g(n) = O(f(n))$, $f(n) \sim g(n)$ if the quotient $f(n)/g(n)$ tends to 1 as $n \to \infty$.
We write c, c', c", ... for positive absolute constants. Different occurances of c may denote different constants. For any real x, $[x]$ and $]x[$ denote the greatest integer not greater than x and the least integer not less than x, respectively. For integers $s \le k$ we put $(k)_s = k(k-1)\dots(k-s+1)$. Furthermore, ln n and log n denote the natural and the dual logarithm, respectively.
We note two consequences of the inequalities of Čebyšev. If X is a non-negative random variable with expectation $\mu = EX > 0$ and $t > 0$, then

$$P(X \ge t\mu) \le 1/t. \tag{1}$$

If, furthermore, X has variance $\sigma^2 = D^2X = EX^2 - \mu^2$ and $t > 0$, then

$$P(|X-\mu| \ge t) \le \sigma^2/t^2. \tag{2}$$

In particular,

$$P(X=0) \le \sigma^2/\mu^2. \tag{3}$$

We use the following estimate:

$$\binom{n}{k} = \left[(en/k)^k/\sqrt{2\pi k}\right](1+O(k^2/n)),\ k^2 = o(n). \tag{4}$$

---

2 This concept of random graphs one can find, for example, in the paper /3/.

3. Results

Throughout the paper we use the following notations:

$$k^* = 2(\log n - \log\log n + \log e) - 1,\ k_1 = [k^*],$$

$$k^{**} = \log n - 2\log\log n + \log\log e,\ k_2 = [k^{**}] + 1,$$

$$\varrho_1 = k_1 + 1,\quad \varrho_2 = k_2 + 1,$$

$$n_{1,i} = \left[2^{2^i}\right],\ n_{2,i} = \left[2^{2^i+2}/e\right],\ n_{3,i} = \left[2^{2^i+1}\ln 2\right],$$

$$n_{4,i} = \left[2^{2^i+[2(\ln 2)i]}\ln 2\right],\ i = 1, 2, 3, \ldots .$$

Now we can list the results.

Theorem 1: For almost every graph the following inequalities hold:

$$k_1 \leq \beta \leq k_1 + 1.^3 \tag{6}$$

Furthermore, for almost every $G_{n_{1,i}}$ ($G_{n_{2,i}}$) it holds $\beta = k_1$ ($\beta = k_1 + 1$).

Theorem 2: For almost every graph the following inequalities hold:

$$k_2 \leq \gamma \leq k_2 + 1. \tag{7}$$

Furthermore, for almost every $G_{n_{1,i}}$ ($G_{n_{3,i}}$) it holds $\gamma = k_2$ ($\gamma = k_2 + 1$).

Theorem 3: For almost every graph there exist independent dominating k-sets if k satisfies

$$\varrho_2 < k < \varrho_1, \tag{8}$$

and there don't exist independent dominating k-sets if k satisfies

$$k < \varrho_2 \quad \text{or} \quad \varrho_1 < k. \tag{9}$$

3 This improves a result from /1/.

Theorem 4: For almost every graph the following inequalities hold:

$$k_2 + 1 \leq i \leq k_2 + 2. \tag{10}$$

For almost every $G_{n_{1,i}}$ ($G_{n_{3,i}}$) it holds $i = k_2 + 1$ ($i = k_2 + 2$).

Theorem 5: For almost every graph the inequalities

$$\gamma \leq i \leq \gamma + 1 \tag{11}$$

are satisfied. For almost every $G_{n_{4,i}}$ it holds $i = \gamma$. (This statement together with theorems 2 and 4 show that also (11) can not be improved.)

4. Proofs

(a) On $\mathcal{G}_n$ we define the random variables X_1, X_2 and X as numbers of independent k-sets, dominating k-sets and independent dominating k-sets, respectively. (In order to simplify the notations we do not indicate the dependence on k and n of these random variables.) Putting $\mu = EX$, $\sigma^2 = D^2X$, $\mu_i = EX_i$, $\sigma_i^2 = D^2X_i$, $i = 1, 2$, it is easy to check that

$$\mu_1 = \binom{n}{k} 2^{-\binom{k}{2}}, \tag{12}$$

$$\mu_2 = \binom{n}{k}(1-2^{-k})^{n-k}, \tag{13}$$

$$\mu = \binom{n}{k} 2^{-\binom{k}{2}} (1-2^{-k})^{n-k}. \tag{14}$$

In fact, if K is a fixed k-set of vertices, then one can easily evaluate the probabilities

P(no two vertices of K are adjacent) $= 2^{-\binom{k}{2}}$ and

P(K dominates a fixed vertex in V-K) $= (1-2^{-k})$. Thus, by (4)

$$\ln \mu_1 = k(\ln 2)(\log n - \tfrac{k-1}{2}) - k(\ln k) + k - \tfrac{\ln k}{2} + O(1), \tag{15}$$

$$\ln \mu_2 = k(\ln n) - n2^{-k} - k(\ln k) + k - \tfrac{\ln k}{2} + O(1), \tag{16}$$

$$\ln \mu = k(\ln 2)(\log n - \frac{k-1}{2}) - n2^{-k} - k(\ln k) + k - \frac{\ln k}{2} + O(1), \tag{17}$$

where $k \to \infty$ such that $k^2 = o(n)$ and $n = o(2^{2k})$.
Finally, we remark that the investigation of ratios $\mu_1(k+1)/\mu_1(k)$, $\mu_2(k+1)/\mu_2(k)$, and $\mu(k+1)/\mu(k)$ shows that μ_1, μ_2, and μ are concave functions in k with the maximal value for k near $\log n - \log\log n$, $n/2$, and $\log n$, respectively.

b) <u>Proof of theorem 1:</u> By (15) we calculate for $k = k^* + \varepsilon$, $\varepsilon = O(\frac{\log\log n}{\log n})$,

$$\log \mu_1 = - \varepsilon \frac{k^*}{2} + (2\log e - \tfrac{1}{2})\log\log n + O(1), \tag{18}$$

and we deduce that

$$\mu_1 \to \begin{cases} 0 \text{ if } \liminf\limits_{n\to\infty} \frac{(k-k^*)\log n}{\log\log n} > 2\log e - \frac{1}{2} \\ \\ \infty \text{ if } \limsup\limits_{n\to\infty} \frac{(k-k^*)\log n}{\log\log n} < 2\log e - \frac{1}{2}. \end{cases} \tag{19}$$

Hence, for $k = k_1 = [k^*]$ it holds $\mu_1 \to \infty$. Furthermore, it is easy to check that for $k = k_1 + 1$ and the sequences $n_{1,i}$ and $n_{2,i}$ hold $\mu_1 \to 0$ and $\mu_1 \to \infty$, respectively. Put $\varepsilon_n^* = k_1 + 1 - k$. Then

$$\varepsilon^*_{n_{1,i}} = [2\log e] + 1 - 2\log e > 0{,}1$$

and

$$\varepsilon^*_{n_{2,i}} \asymp \frac{1}{\log n_{2,i}} \qquad \text{as } i \to \infty,$$

using

$$\log n_{2,i} = 2^i + 2 - \log e - O(\frac{1}{n_{2,i}}),$$

$$i + \frac{c'}{2^i} < \log\log n_{2,i} < i + \frac{c''}{2^i}, \; 0 < c' < c'',$$

$$2^{i+1} - 2i + 4 - \frac{2c''}{2^i} < k^* + 1 < 2^{i+1} - 2i + 4 - \frac{2c'}{2^i},$$

and so

$$\frac{2c'}{2^i} < \varepsilon^*_{n_{2,i}} < \frac{2c''}{2^i}.$$

Now we obtain $\sigma_1^2 = o(\mu_1^2)$ for all k's with $\mu_1 \to \infty$. In order to achieve this, we have to estimate EX_1^2, since $\sigma_1^2 = EX_1^2 - \mu_1^2$. Obviously, $EX_1^2 = \sum_{s=0}^{k} \alpha_s P_s^{(1)}$, denoting by α_s the number of ordered pairs (K,K') of k-sets of vertices with $|K \cap K'| = s$ and by $P_s^{(1)}$ the probability for any such two fixed k-sets to be both independent. Then

$$\alpha_s = \binom{n}{k}\binom{k}{s}\binom{n-k}{k-s},$$

$$P_0^{(1)} = P_1^{(1)} = 2^{-2\binom{k}{2}} \text{ and in general } P_s^{(1)} = 2^{-2\binom{k}{2}+\binom{s}{2}}, \quad s \geq 2.$$

Hence $\alpha_0 P_0^{(1)} + \alpha_1 P_1^{(1)} \leq \binom{n}{k}^2 2^{-2\binom{k}{2}} = \mu_1^2$ and so $\sigma_1^2 \leq \sum_{s=2}^{k} \alpha_s P_s^{(1)}$. Now we will show that

$$\sum_{s=2}^{k} \alpha_s P_s^{(1)} = o(\mu_1^2) \text{ as } \mu_1 \to \infty.$$

Indeed, we may confine us to the evaluation of this sum up to $s = k - 1$ because $\alpha_k P_k^{(1)} = \mu_1 = o(\mu_1^2)$ as $\mu_1 \to \infty$. Using

$$\alpha_s P_s^{(1)} = \frac{\binom{k}{s}\binom{n-k}{k-s}}{\binom{n}{k}} 2^{\binom{s}{2}} \mu_1^2$$

we have with the notation $\gamma_s' = k \dfrac{(k)_s \binom{k}{s} 2^{\binom{s}{2}}}{n^s}$ that

$$\sum_{s=2}^{k-1} \alpha_s P_s^{(1)} \lesssim \left(\max_{2 \leq s \leq k-1} \gamma_s'\right) \mu_1^2.$$ [4]

In the following we will prove $\gamma_s' \to 0$ for $2 \leq s \leq k - 1$ and $k \leq k^* + c$, $c < 1$. Let us calculate $\log \gamma_s'$. For $k^* - s \asymp \log n$ the upper bound

---

[4] $f(n) \lesssim g(n)$ means that $\lim\sup(f(n)/g(n)) \leq 1$ as $n \to \infty$.

$$\log \gamma_s' \le -s(\log n - \tfrac{s-1}{2}) + (2s+1)\log k \tag{20}$$

implies $\log \gamma_s' \to -\infty$. For $s = k^* - u$, $u \ge 1 - c$, $u = o(\log n)$, we require a sharper bound. From

$$\log \gamma_s' = -s(\log n - \tfrac{s-1}{2}) + \log(k)_s + \log\tbinom{k}{s} + \log k, \text{ using}$$

$$\log(k)_s \le \log(k^*+c)! = (k^*+c)\log(k^*+c) - (k^*+c)\log e + O(\log k^*)$$

and $\log\binom{k}{s} \le (u+c)\log(k^*+c) = O(u\log k^*)$, we conclude

$$\log \gamma_s' \le -(k^*-u)(\log\log n + \tfrac{u}{2} - \log e + 1)$$

$$+ (k^*+c)\log(k^*+c) - (k^*+c)\log e + O(u\log k^*)$$

$$= -k^*\tfrac{u}{2} + O(u\log\log n) \to -\infty .$$

The theorem now follows by (1) and (3).

c) <u>Proof of theorem 2:</u>

From (16) we have for $k = k^{**} + \varepsilon$, $\varepsilon = O(\frac{\log\log n}{\log n})$,

$$\log \mu_2 = (\log n)^2(1-2^{-\varepsilon}) - 2(\log n)\log\log n - k^{**}\log k^{**} + O(\log n)$$

$$= (\ln 2)(\log n)^2\varepsilon - 3(\log n)\log\log n + O(\log n)$$

because $(1-2^{-\varepsilon}) = 1 - (1-\varepsilon\ln 2 + O(\varepsilon^2)) = \varepsilon\ln 2 - O(\varepsilon^2)$ as $\varepsilon \to 0$.

Thus, by (21)

$$\mu_2 \to \begin{cases} 0 & \text{if } \lim\limits_{n\to\infty} \sup \frac{(k-k^{**})\log n}{\log\log n} < 3\log e \\ \infty & \text{if } \lim\limits_{n\to\infty} \inf \frac{(k-k^{**})\log n}{\log\log n} > 3\log e \end{cases} . \tag{22}$$

Hence, for $k = k_2 + 1 = [k^{**}] + 2$ it holds $\mu_2 \to \infty$. Put $k = k_2$. Then for $n_{1,i}$ also it holds $\mu_2 \to \infty$, but for $n_{3,i}$ the expectation μ_2 tends to 0. Putting $\varepsilon_n^{**} = k_2 - k^{**}$,

$$\varepsilon_{n_{1,i}}^{**} = 1 - \log\log e > 0{,}4, \tag{23}$$

$$\varepsilon_{n_{3,i}}^{**} = O(\tfrac{1}{\log n_{3,i}}) \qquad \text{as } i \to \infty \tag{24}$$

we show as above.

Now let us estimate the second moment $\sigma_2^2 = EX_2^2 - \mu_2^2$ in order to show $\sigma_2^2 = o(\mu_2^2)$ for k with $k^{**} \le k \le \log n$ and thus, in particular, for $k = k_2$ and $k = k_2 + 1$. (In fact, only these two k's are interesting, since the existence of a dominating k-set for $k = k_2$ or $k = k_2 + 1$, of course, implies the existence of a dominating k-set for every k greater than $k_2 + 1$. But the sharper result - in which we are not interested here - that for almost every graph it holds $X_2 \sim \mu_2$ we may deduce from (2) only for those k's, for which we have shown $\sigma_2^2 = o(\mu_2^2)$.)

Clearly $EX_2^2 = \sum_{s=0}^{k} \alpha_s P_s^{(2)}$, where α_s is defined as above and $P_s^{(2)}$ is the probability that two fixed k-sets K and K' with $|K \cap K'| = s$ are dominating sets. Furthermore $\alpha_k P_k^{(2)} = \mu_2 = o(\mu_2^2)$ as $\mu_2 \to \infty$, $P_o^{(2)} \le (1 - \frac{1}{2^k})^{2(n-2k)}$, and so

$$\alpha_o P_o^{(2)} \le \binom{n}{k}^2 (1-2^{-k})^{2(n-k)} (1-2^{-k})^{-2k} = \mu_2^2 (1+O(k2^{-k})). \text{ Then}$$

$$\alpha_o P_o^{(2)} - \mu_2^2 = \mu_2^2\, O(k2^{-k}) = o(\mu_2^2)$$

because $k^{**} \le k$.

Let K and K' be two fixed k-sets of vertices with $|K \cap K'| = s$, $1 \le s \le k - 1$, and let P(a) denote the probability for a fixed vertex $a \in V-(K \cup K') = R$ to be dominated by both K and K'.
It is easy to check that

$$P(a) = (1-2^{-s}) + 2^{-s}(1-2^{-k+s})^2 = 1 - 2^{-2k}(2^{k+1}-2^{s})$$

because $P(S = K \cap K' \text{ dominates } a) = 1 - 2^{-s}$, but
P(both sets $K - S$ and $K' - S$ dominate a and S doesn't dominate a) $= 2^{-s}(1-2^{-k+s})^2$. Thus

$$P_s^{(2)} \le P(\text{both K and K' dominate R}) = (1-2^{-2k}(2^{k+1}-2^{s}))^{n-2k+s}.$$

Putting $S_s = (1-2^{-2k}(2^{k+1}-2^{s}))^{n-2k+s}$ and using

$\alpha_s = \binom{n}{k}\binom{k}{s}\binom{n-k}{k-s} \leqslant \binom{n}{k}^2 \frac{k^{2s}}{n^s}$ we have with the notation

$\gamma''_s = n^{-s}(1-2^{-k})^{-2(n-k)}k^{2s+1}\beta_s$ that

$$\sum_{s=1}^{k-1} \alpha_s P_s^{(2)} \lesssim (\max_{1\le s\le k-1} \gamma''_s)\mu_2^2.$$

Next we show that $\gamma''_s \to 0$ for $1 \le s \le k-1$. In order to achieve this, we estimate $\ln \gamma''_s$.

$$\begin{aligned}
\ln \gamma''_s &= -s \ln n + (n-2k+s)\ln(1-2^{-2k}(2^{k+1}-2^s)) \\
&\qquad - 2(n-k)\ln(1-2^{-k}) + (2s+1)\ln k \\
&= -s \ln n + (n-2k+s)(-2^{-2k}(2^{k+1}-2^s) - O(2^{-2k})) \\
&\qquad - 2(n-k)(-2^{-k}-O(2^{-2k})) + (2s+1)\ln k \\
&= -s \ln n - n2^{-2k}(2^{k+1}-2^s) + 2n2^{-k} + O(s \ln k) \\
&= -s \ln n + n2^{-2k}2^s + O(s\log\log n)
\end{aligned}$$

provided that $k^{**} \le k \le \log n$.

First, let $s = o(\log n)$. Then, using $k^{**} \le k$, we have

$$n2^{-2k+s} \le 2^s \frac{(\log n)^4}{(\log e)^2 n} = o(1)$$

and consequently $\ln \gamma''_s \to -\infty$.

Now, let $s = \log n - t$, $t = o(\log n)$. Since

$$k \ge k^{**} \doteq \log n - 2\log\log n + \log\log e \quad \text{and}$$

$$k \ge s + 1 = \log n - (t-1),$$

it follows

$$k \ge \log n - \min(t-1, 2\log\log n - \log\log e).$$

That means

$$n^2 2^{-2k-t} \le 2^{2\min(t-1,2\log\log n - \log\log e) - t}$$

$$\le \begin{cases} 2^{t-2} & \text{if } t \le 2\log\log n - \log\log e + 1, \\ 2^{4\log\log n - 2\log\log e - 2\log\log n + \log\log e - 1} & \text{if } t > 2\log\log n - \log\log e + 1, \end{cases}$$

and thus in both cases it holds

$$n^2 2^{-2k-t} \le \tfrac{1}{2}(\ln 2)(\log n)^2.$$

Therefore, we have also for those s

$$\ln \gamma_s'' = -(\ln 2)(\log n)^2 + n^2 2^{-2k-t} + t\ln n + O(\log n \operatorname{loglog} n)$$
$$\leq -\frac{\ln 2}{2}(\log n)^2 + o((\log n)^2) \to -\infty .$$

The theorem we deduce now by (1) and (3).

d) <u>Proof of theorem 3:</u> By (12), (13), and (14) we have that

$$\mu = \mu_1(1-2^{-k})^{n-k} = \mu_2 2^{-\binom{k}{2}} . \tag{25}$$

Then, for k-log $n \to \infty$ it holds $(1-2^{-k})^{n-k} > 1 - 2^{-k}(n-k) = 1 - o(1)$, i. e. $\mu \sim \mu_1$. Thus, by (19) for these k

$$\mu \to \begin{cases} 0 & \text{if } \liminf\limits_{n\to\infty} \frac{(k-k^*)\log n}{\operatorname{loglog} n} > 2\log e - 1/2, \\ \infty & \text{if } \limsup\limits_{n\to\infty} \frac{(k-k^*)\log n}{\operatorname{loglog} n} < 2\log e - 1/2. \end{cases} \tag{26}$$

For $k = k^{**} + 1 + \varepsilon$, $\varepsilon = O(\frac{\operatorname{loglog} n}{\log n})$, we have by (17)

$$\begin{aligned} \log \mu &= \frac{(\log n)^2}{2}(1-2^{-\varepsilon}) - k^{**} \log k^{**} + O(\log n) \\ &= \frac{\ln 2}{2}(\log n)^2 \varepsilon - (\log n) \operatorname{loglog} n + O(\log n) \end{aligned} \tag{27}$$

and by (27)

$$\mu \to \begin{cases} 0 & \text{if } \limsup\limits_{n\to\infty} \frac{(k-k^{**}-1)\log n}{\operatorname{loglog} n} < 2\log e, \\ \infty & \text{if } \liminf\limits_{n\to\infty} \frac{(k-k^{**}-1)\log n}{\operatorname{loglog} n} > 2\log e. \end{cases} \tag{28}$$

Now (9) already follows by (1), (26), and (28). In order to prove (8), we have to show that $\sigma^2 = o(\mu^2)$ as $\mu \to \infty$. We will obtain a little bit more in the following, namely that $\sigma^2 = o(\mu^2)$ is satisfied for $k^{**} + 1 \leq k \leq k^* + c$, $c < 1$.

Let P_s denote the probability that two fixed k-sets K and K' such that $|K \cap K'| = s$ are both independent dominating (maximal independent) sets. Using the notations from (b) and (c) we have in this case:

$$\alpha_o P_o \leq \binom{n}{k}^2 2^{-2\binom{k}{2}} (1-2^{-k})^{2(n-2k)} = \mu^2(1+O(k2^{-k})),$$

$$\alpha_o P_o - \mu^2 = o(\mu^2), \; P_s \leq \beta_s 2^{-2\binom{k}{2}+\binom{s}{2}},$$

$$\alpha_s P_s \leq \gamma_s \mu^2, \text{ where } \gamma_s = \gamma_s' \beta_s (1-2^{-k})^{-2(n-k)},$$

$$\sum_{s=1}^{k-1} \alpha_s P_s \leq \left(\max_{1 \leq s \leq k-1} \gamma_s\right) \mu^2.$$

Since by (25)

$$\ln \gamma_s = \ln \gamma_s' + n2^{-2k+s} + o(1) \tag{29}$$

and moreover $n2^{-2k+s} \leq n2^{-k-1} = O(1)$ for $\log n - c' < s$, we may deduce as under (b) that $\ln \gamma_s \to -\infty$ for $\log n - c' < s \leq k-1 \leq k^* - (1-c)$. Furthermore, by (20) and (29)

$$\ln \gamma_s = -s(\ln 2)(\log n - \tfrac{s-1}{2}) + n2^{-2k+s} + O(s \log\log n)$$

and therefore we have with the same arguments as in the discussion of $\ln \gamma_s''$ under (c) that for $k^{**} + 1 \leq k$ and

(I) $s = o(\log n)$ it holds $n2^{-2k+s} = o(1)$ and $-s(\ln n)$ is the main term of $\ln \gamma_s$,

(II) $s = \log n - t$, $t = o(\log n)$, it holds

$$\ln \gamma_s = -\frac{\ln 2}{4}(\log n)^2 + O(\log n \log\log n).$$

In both cases γ_s tends to $-\infty$ as $n \to \infty$.

Now (8) follows by (3).

(e) Theorem 4 is in part a corollary of theorem 3, namely (10) follows by (8) and (9). The second assertion of the theorem follows by (23), (24), and (28) because $\varepsilon_n^{**} = (k_2+1) - (k^{**}+1)$ and $\sigma^2 = o(\mu^2)$ for $k^{**} + 1 \leq k \leq k^* + c$, $c < 1$.

f) Proof of theorem 5: For every graph it holds clearly $\gamma \leq i$. By theorems 2 and 4 we know that almost every graph satisfies $k_2 \leq \gamma \leq i \leq k_2 + 2$. Now (22) and (28) imply that for almost every graph with $\gamma = k_2$ also it holds $i = k_2 + 1$. We already have seen above that for almost every $G_{n_{3,i}}$ it holds $\gamma = k_2 + 1$ and $i = k_2 + 2$. Finally we check that for almost every $G_{n_{4,i}}$ it holds $i = \gamma = k_2 + 1$. It is by setting of $n_{4,i}$

$$\log n_{4,i} = 2^i + [2(\ln 2)i] - \log\log e - O(1/n_{4,i}),$$

$$\log\log n_{4,i} = i + 2i/2^i - O(1/\log n_{4,i}),$$

$$k^{**} = 2^i + [2(\ln 2)i] - 2i - 4i/2^i + O(1/\log n_{4,i}),$$

and thus

$$\varepsilon_{n_{4,i}}^{**} = [k^{**}] + 1 - k^{**} = 4i/2^i - O(1/\log n_{4,i}) \sim 4 \frac{\log\log n_{4,i}}{\log n_{4,i}}$$

as $i \to \infty$. By (22) and (28) now the assertion follows, since $2\log e < 4 < 3\log e$. This completes the proof of theorem 5.

5. The case of arbitrary fixed p.

It is evident that theorems 1 - 5 also hold for such random graphs, in which each edge occurs independently of any other edges with arbitrary but fixed probability p, $0 < p < 1$. We have only to interpret log as logarithm with base $r = 1/p$, since in the general case

$$\mu_1 = \binom{n}{k} r^{-\binom{k}{2}},$$

$$\mu_2 = \binom{n}{k}(1-r^{-k})^{n-k},$$

and

$$\mu = \binom{n}{k} r^{-\binom{k}{2}} (1-r^{-k})^{n-k}$$

are the expectation values. As sequences $n_{1,i}, \ldots, n_{4,i}$ one can take in the general case

$$n_{1,i} = \left[r^{r^i}\right], \quad n_{2,i} = \left[r^{r^i} +]\log e[\, ^{i}/_{e}\right],$$

$$n_{3,i} = \left[r^{r^i} +]\log\log e[\,_{\ln r}\right],$$

$$n_{4,i} = \left[r^{r^i} +]1{,}25i[\,_{\ln r}\right], \quad i = 1, 2, 3, \ldots .$$

<u>References</u>

/1/ Korschunow, A. D.: About the interiorly stable number. Cybernetics 1, 17 - 28 (1974)

/2/ Cockayne, E. J., Favaron, O., Payan, C., and Thomason, A.: Contributions to the theory of domination, independence and irredundance in graphs. Preprint. University of Victoria, Department of Mathematics, Victoria 1980

/3/ Bollobás, B.: Degree sequences of random graphs. Mathematisk Institut, Aarhus Universitet, Preprint Series 1978/79 No. 9, Aarhus 1978

<u>received:</u> April 8, 1980

<u>Author's address:</u>

Dr. rer. nat. Karl Weber
Wilhelm-Pieck-Universität Rostock
Sektion Mathematik
Universitätsplatz 1
DDR-2500 Rostock

Rostock. Math. Kolloq. 16, 45 - 57 (1981)

Günther Wildenhain

Über punktweise Abschätzungen von Lösungen linearer elliptischer Differentialgleichungen beliebiger Ordnung

1. Der Beweis von apriori-Abschätzungen für Lösungen partieller Differentialgleichungen ist bekanntlich der Dreh- und Angelpunkt in der Theorie der Randwertprobleme. Während man im Rahmen der funktionalanalytischen Interpretation solcher Probleme in der Regel mit Abschätzungen in Integralnormen auskommt, benötigt man für feinere Untersuchungen (zum Beispiel potentialtheoretischen Charakters) punktweise Abschätzungen in C-Normen. Diese sind jedoch im allgemeinen schwieriger zu beweisen und daher in der Literatur seltener vorhanden. Neben den klassischen Maximum-Minimum-Abschätzungen für Lösungen elliptischer Gleichungen 2. Ordnung sind für lineare elliptische Gleichungen höherer Ordnung im wesentlichen die Schauder-Abschätzungen von S. Agmon, A. Douglis, L. Nirenberg /2/ und J. P. Krasowskij /5/ sowie die Agmon-Miranda-Abschätzungen (vgl. /1/, /8/) zu nennen. Die letzteren sagen aus, daß unter gewissen Glattheitsvoraussetzungen an den Rand $\partial\Omega$ des beschränkten Gebietes $\Omega \subset \mathbb{R}^n$ und an die Koeffizienten des elliptischen Operators L der Ordnung 2m für die Lösungen $u \in C^{2m}(\Omega) \cap C^{m-1}(\bar{\Omega})$ der homogenen Gleichung $Lu = 0$ für $x \in \Omega$ und alle Ableitungen $D^\alpha u$ mit $|\alpha| \leq m-1$ mit einer von u unabhängigen Konstanten $k > 0$ die Ungleichung

$$|D^\alpha u(x)| \leq k \sum_{|\alpha|\leq m-1} \sup_{y\in\partial\Omega} |D^\alpha u(y)| \tag{1}$$

besteht. In der vorliegenden Arbeit wollen wir Resultate von J. M. Berezanskij und J. A. Rojtberg (vgl. /3/, /4/, /7/) benutzen, um folgendes zu zeigen:

Sind $q > \frac{n}{2}$ und $l \geq 0$ ganze Zahlen, so gelten (unter gewissen Glattheitsvoraussetzungen) für die Lösungen $u \in C^\infty(\bar{\Omega})$ der homogenen elliptischen Gleichung $Lu = 0$ mit von u unabhängigen Konstanten $C > 0$ Ungleichungen der Gestalt

$$\|u\|_{C^1(\overline{\Omega})} \leq C \, \|u\|_{C^{q+1}(\partial\Omega)} := C \sum_{|\alpha| \leq q+1} \sup_{y \in \partial\Omega} |D^\alpha u(y)|. \tag{2}$$

Ein Vergleich zeigt, daß sich daraus für $\frac{n}{2} < q < m-1$ in gewissen Fällen Verschärfungen gegenüber (1) ergeben, z. B. für $n = 2,3$, $l = 0$, $m \geq 4$.

2. Es sei $\Omega \subset \mathbb{R}^n$ im folgenden stets ein beschränktes Gebiet mit beliebig oft differenzierbarem Rand. Für $s \geq 0$ ganz bezeichne $W_2^s(\Omega)$ den klassischen Sobolev-Raum mit der Norm

$$\|u\|_s := \Big(\sum_{|\alpha| \leq s} \int_\Omega |D^\alpha u|^2 dx \Big)^{\frac{1}{2}}.$$

Der Raum $W_2^{-s}(\Omega)$ ($s > 0$ ganz) sei als Vervollständigung von $L^2(\Omega)$ bezüglich der Norm

$$\|u\|_{-s} := \sup_{v \in W_2^s(\Omega)} \frac{|(u,v)|}{\|v\|_s}$$

erklärt, wobei (u,v) das gewöhnliche L^2-Skalarprodukt bedeutet. $W_2^s(\Omega)$ und $W_2^{-s}(\Omega)$ können als zueinander duale Räume aufgefaßt und das Skalarprodukt (u,v) kann im Sinne dieser Dualität für beliebige $u \in W_2^{-s}(\Omega)$, $v \in W_2^s(\Omega)$ erklärt werden. Es gilt

$$|(u,v)| \leq \|u\|_{-s} \, \|v\|_s. \tag{3}$$

Für ganzzahlige $s \geq 0$ bezeichnen wir weiter mit $W_2^{s-\frac{1}{2}}(\partial\Omega)$ den Vektorraum aller Funktionen φ auf $\partial\Omega$, die (im Sinne der Einbettungssätze, vgl. /3/, S. 86) Randwerte von Funktionen aus $W_2^s(\Omega)$ sind.

Durch

$$\|\varphi\|_{s-\frac{1}{2}} := \inf \|u\|_s,$$

wobei das Infimum über alle $u \in W_2^s(\Omega)$ mit $u|_{\partial\Omega} = \varphi$ erstreckt ist, wird $W_2^{s-\frac{1}{2}}(\partial\Omega)$ ein Hilbertraum. Den zu $W_2^{s-\frac{1}{2}}(\partial\Omega)$ dualen Raum $W_2^{-s+\frac{1}{2}}(\partial\Omega)$ erhält man durch Vervollständigung von $C(\partial\Omega)$ bezüglich der Norm

$$\|\psi\|_{-s+\frac{1}{2}} := \sup_{\varphi \in W_2^{s-\frac{1}{2}}(\partial\Omega)} \frac{|(\psi,\varphi)|}{\|\varphi\|_{s-\frac{1}{2}}}.$$

Dabei ist $(\psi,\varphi) = \int_{\partial\Omega} \psi(y)\overline{\varphi(y)}\, d\sigma(y)$. Im Sinne der Dualität kann auch hier das Skalarprodukt (ψ,φ) für beliebige $\varphi \in W_2^{s-\frac{1}{2}}(\partial\Omega)$, $\psi \in W_2^{-s+\frac{1}{2}}(\partial\Omega)$ erklärt werden, so daß in diesem Falle die Ungleichung

$$|(\psi,\varphi)| \le \|\psi\|_{-s+\frac{1}{2}} \|\varphi\|_{s-\frac{1}{2}} \tag{4}$$

gilt. Die folgende Raumklasse wurde von J. M. Berezanskij und J. A. Rojtberg untersucht (vgl. /3/, /4/, /7/ u. a.).
Es sei s eine beliebige ganze, $r > 0$ eine natürliche Zahl. Unter $\tilde{W}^s_{2,r}(\Omega)$ verstehen wir dann die Vervollständigung von $C^\infty(\overline{\Omega})$ bezüglich der Norm

$$|||u|||_s^2 = \|u\|_s^2 + \sum_{j=1}^{r} \left\|\frac{\partial^{j-1}u}{\partial\nu^{j-1}}\right\|_{s-j+\frac{1}{2}}.$$

Dabei bezeichnet $\nu = \nu(x)$ die Außennormale an Ω im Punkte $x \in \partial\Omega$. Zur Motivierung der Definition beachte man $s - (j-1) - \frac{1}{2} = s - j + \frac{1}{2}$. Es ist leicht zu sehen, daß für $s \geq r$

der Raum $\tilde{W}^s_{2,r}(\Omega)$ mit $W^s_2(\Omega)$ zusammenfällt. Für $s < r$ gilt $\tilde{W}^s_{2,r}(\Omega) \subset W^s_2(\Omega)$. Wir betrachten jetzt einen Differentialausdruck $L = L(x,D)$ der Ordnung $l \leq r$ in Ω mit beliebig oft differenzierbaren Koeffizienten. $B = B(y,D)$ sei ein auf dem Rand $\partial\Omega$ erklärter Differentialausdruck der Ordnung $t \leq r - 1$, von dem wir ebenfalls voraussetzen, daß die Koeffizienten (bezüglich lokaler Koordinaten) beliebig oft differenzierbar sind. Dann gilt das folgende wichtige Lemma, das z. B. in /3/ bewiesen ist.

Lemma 1: Für beliebige ganzzahlige s existieren Konstanten $C_s > 0$, die nicht von u abhängen, so daß für alle $u \in C^\infty(\overline{\Omega})$ die Ungleichungen

$$\|Lu\|_{s-l} \leq C_s \,|||u|||_s, \quad \|Bu\|_{s-t-\frac{1}{2}} \leq C_s \,|||u|||_s$$

gelten. Die Abschließungen der Abbildungen $u \to Lu$, $u \to Bu|_{\partial\Omega}$ $(u \in C^\infty(\overline{\Omega}))$ bilden also $\tilde{W}^s_{2,r}(\Omega)$ stetig in $W^{s-l}_2(\Omega)$ bzw. $W_2^{s-t-\frac{1}{2}}(\partial\Omega)$ ab.

3. Für das Weitere setzen wir jetzt voraus, daß

$$L(x,D) = \sum_{|\alpha| \leq 2m} a_\alpha(x) D^\alpha$$

ein in $\overline{\Omega}$ eigentlich elliptischer Differentialoperator der Ordnung 2m mit beliebig oft differenzierbaren Koeffizienten ist. Ferner sei

$$B_j(y,D) = \sum_{|\gamma| \leq m_j} b_{j\gamma}(y) D^\gamma \quad (j=1,\ldots,m)$$

ein System von Randoperatoren auf $\partial\Omega$, welches den folgenden Bedingungen genügt:

(i) Die Koeffizienten $b_{j\gamma}$ sind beliebig oft differenzierbar.
(ii) Das System ist normal.
(iii) Das System überdeckt den Operator L auf $\partial\Omega$.

Bezüglich der Definition der im vorangehenden verwendeten Begriffe verweisen wir auf /6/. Mit dem Randwertproblem

$$Lu = f \text{ in } \Omega, \quad B_j u|_{\partial\Omega} = \varphi_j \quad (j=1,\dots,m) \tag{5}$$

kann man in bekannter Weise eine Greensche Formel in Verbindung bringen (vgl. /6/). Dazu setzen wir zusätzlich $m_j \leq 2m-1$ für $j = 1, \dots, m$ voraus. Dann existiert ein (nicht eindeutig bestimmtes) normales System $\{C_j\}_{j=1,\dots,m}$ von Randoperatoren mit beliebig oft differenzierbaren Koeffizienten und $\operatorname{ord} C_j = l_i \leq 2m-1$ derart, daß $\{B_1,\dots,B_m, C_1,\dots,C_m\}$ auf $\partial\Omega$ ein Dirichlet-System der Ordnung 2m bildet. Dabei heißt allgemein ein System $\{A_j\}_{j=1,\dots,k}$ von Randoperatoren ein Dirichlet-System der Ordnung k, wenn es auf $\partial\Omega$ normal ist und die Ordnungen m_j der Operatoren A_j für $j = 1, \dots, k$ alle Zahlen $0,1, \dots, k-1$ durchlaufen. Ist das System $\{C_j\}_{j=1,\dots,m}$ fixiert, so findet man in eindeutiger Weise weitere 2m Randoperatoren $\{B'_j\}_{j=1,\dots,m}$, $\{C'_j\}_{j=1,\dots,m}$ mit beliebig oft differenzierbaren Koeffizienten auf $\partial\Omega$, so daß die folgenden Eigenschaften gelten:

1. $\operatorname{ord} B'_j = m'_j = 2m - 1 - l_j$,
   $\operatorname{ord} C'_j = l'_j = 2m - 1 - m_j$.

2. Das System $\{B'_1,\dots,B'_m, C'_1,\dots,C'_m\}$ ist ein Dirichlet-System der Ordnung 2m auf $\partial\Omega$ und für $u,v \in C^\infty(\overline{\Omega})$ gilt die Greensche Formel

$$\int_\Omega Lu\,\overline{v}\,dx - \int_\Omega u\overline{L^* v}\,dx = \sum_{j=1}^m \int_{\partial\Omega} C_j u \overline{B'_j v}\,d\sigma - \sum_{j=1}^m \int_{\partial\Omega} B_j u \overline{C'_j v}\,d\sigma. \tag{6}$$

$L^*(x,D)v = \sum_{|\alpha|\leq 2m} (-1)^{|\alpha|} D^\alpha(\overline{a}_\alpha v)$ ist der zu $L(x,D)$ adjungierte Differentialausdruck. Das Randwertproblem

$$L^* v = g \text{ in } \Omega, \quad B'_j v|_{\partial\Omega} = \psi_j \quad (j=1,\dots,m) \tag{7}$$

heißt zu (5) adjungiert.

Wir setzen nun (für beliebige ganzzahlige s) zur Abkürzung $\widetilde{W}^s_{2,2m}(\Omega) =: \widetilde{W}^s_2(\Omega)$ und vereinbaren, daß die Anwendung von L bzw. eines Randoperators der Ordnung $\leq 2m-1$ auf eine Funktion aus $\widetilde{W}^s_2(\Omega)$ stets im Sinne von Lemma 1 zu verstehen ist. Entsprechend dieser Vereinbarung wollen wir eine wichtige Erweiterung der Formel (6) vornehmen.

<u>Lemma 2:</u> Im Sinne der Dualität bezüglich des L^2-Skalarproduktes in $L^2(\Omega)$ bzw. $L^2(\partial\Omega)$ (vgl. Punkt 2) gilt für beliebige ganzzahlige s und $u \in \widetilde{W}^s_2(\Omega)$, $v \in \widetilde{W}^{2m-s}_2(\Omega)$ die Formel

$$(u,L^*v) + \sum_{j=1}^{m} (C_j u,\ B_j' v) = (Lu,v) + \sum_{j=1}^{m} (B_j u,\ C_j' v). \tag{8}$$

<u>Beweis:</u> Wir wählen Folgen (u_n), (v_n) mit $u_n \in C^\infty(\overline{\Omega})$, $v_n \in C^\infty(\overline{\Omega})$ und $|||u - u_n|||_s \to 0$, $|||v - v_n|||_{2m-s} \to 0$ und führen in der für u_n, v_n geltenden Formel (6) den Grenzübergang $n \to \infty$ durch. Daraus folgt unmittelbar die Behauptung; denn betrachten wir etwa den Summanden $(C_j u_n,\ B_j' v_n)$, so gilt nach Lemma 1

$$\| C_j u_n - C_j u \|_{s-l_j-\frac{1}{2}} \leq C_1\ |||u_n - u|||_s \to 0,$$

$$\| B_j' v_n - B_j' v \|_{2m-s-m_j'-\frac{1}{2}} \leq C_2\ |||v_n - v|||_s \to 0.$$

Wegen der Stetigkeit des Skalarprodukts bezüglich beider Faktoren (Formel (4)) erhalten wir

$$(C_j u_n,\ B_j' v_n) \to (C_j u,\ B_j' v).$$

Der rechte Ausdruck hat wegen $C_j u \in W_2^{s-l_j-\frac{1}{2}}(\partial\Omega)$, $B_j' v \in W_2^{2m-s-m_j'-\frac{1}{2}}(\partial\Omega)$ und $2m - s - m_j' - \frac{1}{2} = 2m - s - (2m - l_j - 1) - \frac{1}{2} = -s + l_j + \frac{1}{2}$ einen Sinn.

4. Wir betrachten jetzt das Randwertproblem (5) für $f \in W_2^s(\Omega)$, $\varphi_j \in W_2^{2m+s-m_j-\frac{1}{2}}(\partial\Omega)$ und setzen (für beliebige ganzzahlige s)

$$K_{(s,2m+s-m_j-\frac{1}{2})} = W_2^s(\Omega) \times \prod_{j=1}^{m} W_2^{2m+s-m_j-\frac{1}{2}}(\partial\Omega).$$

Offenbar definiert der Operator

$$\mathcal{L}_s = (L, B_1,\ldots,B_m)$$

für $s \geq 0$ in natürlicher Weise eine Abbildung von $W_2^{2m+s}(\Omega)$ in $K_{(s,2m+s-m_j-\frac{1}{2})}$. Entsprechendes gilt für das adjungierte Problem (7) bezüglich des Operators

$$\mathcal{L}_s^* = (L^*, B_1',\ldots,B_m'),$$

der $W_2^{2m+s}(\Omega)$ in $K_{(s,2m+s-m_j'-\frac{1}{2})}$ abbildet.

Die zugehörigen Nullräume N bzw. N^* sind endlichdimensional und auf Grund der von uns gemachten Annahmen in $C^\infty(\bar{\Omega})$ enthalten (vgl. /3/).

Für $s \geq 0$ sei $H_s = W_2^s(\Omega) \ominus N$, $H_s^* = W_2^s(\Omega) \ominus N^*$.

Für $F = (f, \varphi_1,\ldots,\varphi_m) \in K_{(s,2m+s-m_j-\frac{1}{2})}$,

$G = (g, \psi_1,\ldots,\psi_m) \in K_{(s,2m+s-m_j'-\frac{1}{2})}$, $u \in N$, $v \in N^*$

benutzen wir die Bezeichnung

$$[F,C'v] = (f,v) + \sum_{j=1}^{m} (\varphi_j, C_j'v).$$

Man kann zeigen, daß jedes Element

$$F = (f, \varphi_1,\ldots,\varphi_m) \in K_{(s,2m+s-m_j-\frac{1}{2})}$$

in eindeutiger Weise in die direkte Summe

$F = F' + F''$ mit $F' = (v',0,\ldots,0)$ $(v' \in N^*)$,

$F'' \in K_{(s,2m+s-m_j-\frac{1}{2})}$ $([F'',C'v] = 0$ für alle $v \in N^*)$

zerlegt werden kann. Die zugehörigen, durch $Q_{N^*}F = F'$, $Q_{N^*_\perp}F = F''$ definierten Projektionsoperatoren sind in $K_{(s,2m+s-m_j-\frac{1}{2})}$ stetig. Ganz analog erklärt man die Projektoren Q_N und $Q_{N_\perp}$ im Raum $K_{(s,2m+s-m_j'-\frac{1}{2})}$. Wir setzen

$$K_s = Q_{N_\perp}K_{(s,2m+s-m_j'-\frac{1}{2})}, \quad K_s^* = Q_{N^*_\perp}K_{(s,2m+s-m_j-\frac{1}{2})}.$$

Zum Beispiel ist dann in /3/ bewiesen, daß der Operator $\mathcal{L}_s$ für ganzzahlige $s \geq 0$ einen Homöomorphismus von H_{2m+s} (bezüglich der Norm $\|\cdot\|_{2m+s}$) auf den Raum K_s^* (bezüglich der üblichen Produktnorm in $K_{(s,2m+s-m_j-\frac{1}{2})}$) darstellt. Entsprechend liefert $\mathcal{L}_s^*$ einen Homöomorphismus von H^*_{2m+s} auf K_s. Entscheidend für unsere Ziele ist nun der Umstand, daß sich dieser Homöomorphiesatz für beliebige ganzzahlige s verallgemeinern läßt (vgl. /3/, /4/). Dazu führen wir die Räume

$\tilde{H}_s = \tilde{W}_2^s(\Omega) \ominus N$, $\tilde{H}_s^* = \tilde{W}_2^s(\Omega) \ominus N^*$

ein und erklären K_s, K_s^* wie oben, was möglich ist, da die Projektoren $Q_{N_\perp}$ bzw. $Q_{N^*_\perp}$ in $K_{(0,2m-m_j'-\frac{1}{2})}$ bzw. $K_{(0,2m-m_j-\frac{1}{2})}$ erklärt sind und diese Räume für $s < 0$ in $K_{(s,2m+s-m_j'-\frac{1}{2})}$ bzw. $K_{(s,2m+s-m_j-\frac{1}{2})}$ dicht liegen.

<u>Satz 1:</u> **Für beliebige ganzzahlige s existiert ein Homöomorphismus Λ_s von $\tilde{H}_{2m+s}$ auf K_s^*. Für $s \geq 0$ ist Λ_s durch die Einschränkung von $\mathcal{L}_s$ auf $\tilde{H}_{2m+s}$ gegeben. Für $s < 0$ erhält man Λ_s durch stetige Abschließung des Operators $\mathcal{L}_0$, wenn man diesen als Abbildung von $\tilde{H}_{2m+s}$ in K_s^* auffaßt. Entsprechend existiert ein Homöomorphismus Λ_s^* von $\tilde{H}^*_{2m+s}$ auf K_s, der für $s \geq 0$ durch die Einschränkung von $\mathcal{L}_s^*$ auf $\tilde{H}^*_{2m+s}$ und für $s < 0$ durch stetige**

Abschließung von $\mathcal{L}_0^*$ (aufgefaßt als Abbildung von $\tilde{H}^*_{2m+s}$ in K_s) gegeben ist. Insbesondere gilt also für die Lösung von (5)

$$c_1 \| u \|_{2m+s} \leq \| f \|_s^2 + \sum_{j=1}^{m} \| \varphi_j \|^2_{2m+s-m_j-\frac{1}{2}} \leq c_2 \| u \|^2_{2m+s}$$

und für die Lösung von (7)

$$c_3 \| v \|_{2m+s} \leq \| g \|_s^2 + \sum_{j=1}^{m} \| \varphi_j \|_{2m+s-m_j'-\frac{1}{2}} \leq c_4 \| v \|^2_{2m+s}. \qquad (9)$$

5. Man kann den Satz 1 benutzen, um für die Lösungen von (5) einen Darstellungssatz zu beweisen, der im wesentlichen auch bereits auf Berezanskij und Rojtberg /4/ zurückgeht. Außer den bisherigen Annahmen machen wir dazu für das Folgende die zusätzliche Voraussetzung $N = N^* = \{0\}$.

Satz 2: Es sei $q > \frac{n}{2}$, $\alpha \geq 0$ ein fixierter Multiindex und

$$(f, \varphi_1, \ldots, \varphi_m) \in W_2^{q+|\alpha|-2m} \times \prod_{j=1}^{m} W_2^{q+|\alpha|-m_j-\frac{1}{2}}(\partial\Omega).$$

Dann existieren für $x \in \Omega$ Funktionen $\Gamma^{(B)}_{x,\alpha} \in \tilde{W}_2^{2m-q-|\alpha|}(\Omega)$ derart, daß die Ableitung $D^\alpha u$ der Lösung $u \in \tilde{W}_2^{q+|\alpha|}(\Omega)$ von (5) die Darstellung

$$(-1)^{|\alpha|} D^\alpha u(x) = (f, \Gamma^{(B)}_{x,\alpha}) + \sum_{j=1}^{m} (\varphi_j, C_j' \Gamma^{(B)}_{x,\alpha}) \qquad (10)$$

besitzt.

Beweis: Es sei $v \in W_2^{q+|\alpha|}(\Omega)$. Wegen $q > \frac{n}{2}$ folgt aus den Einbettungssätzen $D^\alpha v \in C(\bar{\Omega})$ und

$$|(D_x^\alpha \delta_x, v)| = |D_x^\alpha v(x)| \leq \tilde{c}_\alpha \| v \|_{q+|\alpha|} \qquad (11)$$

für alle $v \in W_2^{q+|\alpha|}(\Omega)$, d. h. $D_x^\alpha \delta_x \in W_2^{-q-|\alpha|}(\Omega)$.

Folglich besitzt das Problem

$$L^*v = D_x^\alpha \delta_x,\ B_j'v|_{\partial\Omega} = 0 \qquad (j=1,\dots,m) \tag{12}$$

nach Satz 1 (mit $s = -q - |\alpha|$) eine Lösung im Raum $\tilde{W}_2^{2m-q-|\alpha|}(\Omega)$, die wir mit $\Gamma_{x,\alpha}^{(B)}$ bezeichnen wollen. Aus Satz 1 folgt weiter, daß die Lösung u von

$$Lu = f,\ B_ju|_{\partial\Omega} = \varphi_j \qquad (j=1,\dots,m)$$

auf Grund der gemachten Annahmen in $\tilde{W}_2^{q+|\alpha|}(\Omega)$ liegt. Nach Lemma 2 können wir u und $v = \Gamma_{x,\alpha}^{(B)}$ in die Formel (8) einsetzen. Wir erhalten

$$(u, D_x^\alpha \delta_x) + \sum_{j=1}^m (C_ju, 0) = (f, \Gamma_{x,\alpha}^{(B)}) + \sum_{j=1}^m (\varphi_j, C_j' \Gamma_{x,\alpha}^{(B)}),$$

d. h.

$$(-1)^{|\alpha|} D^\alpha u(x) = (f, \Gamma_{x,\alpha}^{(B)}) + \sum_{j=1}^m (\varphi_j, C_j' \Gamma_{x,\alpha}^{(B)}).$$

<u>Lemma 3:</u> Für einen beliebigen Multiindex α gilt die Ungleichung

$$|||\Gamma_{x,\alpha}^{(B)}|||_{2m-q-|\alpha|} \leq c_\alpha$$

mit einer von $x \in \Omega$ unabhängigen Konstanten.

<u>Beweis:</u> Wegen (11) gilt

$$\|D_x^\alpha \delta_x\|_{-q-|\alpha|} = \sup_{v \in W_2^{q+|\alpha|}(\Omega)} \frac{|(D_x^\alpha \delta_x, v)|}{\|v\|_{q+|\alpha|}} \leq \tilde{c}_\alpha$$

unabhängig von x. Wendet man den Satz 1 (Formel (9)) auf das Problem (12) an, so ergibt sich

$$|||\Gamma_{x,\alpha}^{(B)}|||_{2m-q-|\alpha|} \leq \frac{1}{c_3} \|D_x^\alpha \delta_x\|_{-q-|\alpha|} \leq \frac{1}{c_3} \tilde{c}_\alpha =: c_\alpha.$$

<u>Satz 3:</u> Unter den in Satz 2 gemachten Voraussetzungen gilt für die Lösung u des Problems (5) die Abschätzung

$$|D^{\alpha}u(x)| \leq \varrho \left\{ \|f\|_{q+|\alpha|-2m} + \sum_{j=1}^{m} \|\varphi_j\|_{q+|\alpha|-m_j-\frac{1}{2}} \right\}$$

mit einer von x unabhängigen Konstanten.

Beweis: Aus (10), (3) und (4) folgt

$$|D^{\alpha}u(x)| \leq \|f\|_{q+|\alpha|-2m} \, \|\Gamma^{(B)}_{x,\alpha}\|_{2m-q-|\alpha|}$$

$$+ \sum_{j=1}^{m} \|\varphi_j\|_{q+|\alpha|-m_j-\frac{1}{2}} \|C'_j \, \Gamma^{(B)}_{x,\alpha}\|_{-q-|\alpha|+m_j+\frac{1}{2}}.$$

Ferner ist $\|\Gamma^{(B)}_{x,\alpha}\|_{2m-q-|\alpha|} \leq |||\Gamma^{(B)}_{x,\alpha}|||_{2m-q-|\alpha|}$ und wegen ord $C'_j = 2m - 1 - m_j$ folgert man aus Lemma 1 und Lemma 3

$$\|C'_j \, \Gamma^{(B)}_{x,\alpha}\|_{-q-|\alpha|+m_j+\frac{1}{2}} \leq c_j \, |||\Gamma^{(B)}_{x,\alpha}|||_{2m-q-|\alpha|} \leq c_j C_{\alpha}$$

unabhängig von x. Mit $\varrho = \max(1, c_1 C_{\alpha}, \ldots, c_m C_{\alpha})$ erhalten wir dann die Behauptung.

Wir werden uns jetzt auf den Fall $f = 0$ und auf das Dirichlet-Problem beschränken, d. h.

$$Lu = 0, \; \left.\frac{\partial^{j-1}u}{\partial\nu^{j-1}}\right|_{\partial\Omega} = \varphi_j \qquad (j=1,\ldots,m). \tag{13}$$

Satz 4: Es gelte $q > \frac{n}{2}$, und $u \in C^{\infty}(\overline{\Omega})$ sei in Ω Lösung der Gleichung $Lu = 0$. Dann gilt für beliebige ganzzahlige $l \geq 0$ mit einer von u unabhängigen Konstanten $C > 0$ die Ungleichung

$$\|u\|_{C^{l}(\overline{\Omega})} \leq C \, \|u\|_{C^{q+l}(\partial\Omega)} := C \sum_{|\alpha| \leq q+l} \; \sup_{y \in \partial\Omega} |D^{\alpha}u(y)|. \tag{14}$$

Beweis: Wir fixieren einen beliebigen Multiindex α mit $|\alpha| = l$. Wegen $u \in C^{\infty}(\overline{\Omega})$ können wir die Randeinschränkungen

$$\left.\frac{\partial^{j-1}u}{\partial\nu^{j-1}}\right|_{\partial\Omega} =: \varphi_j$$

als Elemente von $W_2^{q+|\alpha|-j+\frac{1}{2}}(\partial\Omega)$ und u als Lösung des Problems (13) auffassen. Aus Satz 3 (mit $B_j = \frac{\partial^{j-1}}{\partial\nu^{j-1}}$, d. h. $m_j = j - 1$) erhalten wir somit

$$|D^\alpha u(x)| \leq \varrho \sum_{j=1}^{m} \|\varphi_j\|_{q+|\alpha|-j+\frac{1}{2}} \leq \varrho \sum_{j=1}^{m} \|\varphi_j\|_{q+|\alpha|-j+1}$$

$$\leq \varrho_1 \|u|_{\partial\Omega}\|_{q+|\alpha|} = \varrho_1 \sum_{|\beta|\leq q+|\alpha|} \int_{\partial\Omega} |D^\beta u(y)|^2 d\sigma(y)$$

$$\leq \varrho_2 \|u\|_{C^{q+|\alpha|}(\partial\Omega)}.$$

Da die entsprechende Abschätzung für alle γ mit $|\gamma| \leq 1$ in dieser Weise gewonnen werden kann und offenbar

$$\|u\|_{C^{q+|\gamma|}(\partial\Omega)} \leq \|u\|_{C^{q+1}(\partial\Omega)}$$

für $|\gamma| \leq 1$ gilt, folgt die Behauptung.

Für $l = 0$ ergibt sich aus der Ungleichung (14)

$$\|u\|_{C(\bar{\Omega})} \leq C\|u\|_{C^q(\partial\Omega)}, \tag{15}$$

während die Agmon-Miranda-Abschätzung (1)

$$\|u\|_{C(\bar{\Omega})} \leq k \|u\|_{C^{m-1}(\partial\Omega)} \tag{16}$$

liefert. Für $\frac{n}{2} < q < m - 1$ ist also (15) schärfer als (16). Dies trifft zum Beispiel zu bei $n = 2,3$ für die Fälle $m \geq 4$, d. h. für Operatoren der Ordnung ≥ 8, und bei $n = 4,5$ für $m \geq 5$, d. h. für Operatoren der Ordnung ≥ 10. Satz 4 gilt auch für Lösungen der Gleichung $Lu = 0$ mit geringeren Glattheitsvoraussetzungen, wovon man sich durch Grenzübergang überzeugt. Ferner ist aus dem Vorangehenden klar (und deshalb wurden die Betrachtungen bis einschließlich Satz 3 in dieser Allgemeinheit durchgeführt) wie man (auch für den inhomogenen Fall) allgemeineren Randbedingungen angepaßte, zu Satz 4 analoge Aussagen formulieren kann.

Literatur

/1/ Agmon, S.: Maximum theorems for solutions of higher order elliptic equations.
Bull. Amer. Math. Soc. 66, 77 - 80 (1960)

/2/ Agmon, S., Douglis, A., and Nirenberg, L.: Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions I. Comm. Pure Appl. Math. 12, 623 - 727 (1959)

/3/ Березанский, Ю. М.: Разложение по собственным функциям самосопряженных операторов. Киев 1965

/4/ Березанский, Ю.М., и Ройтберг, Я.А.: Теорема о гомеоморфизмах и функция Грина для общих эллиптических граничных задач.
Укр. матем. ж. I9, Но.5, 3-32 (1967)

/5/ Красовский, Ю.П.: Оценки роста производных решений однородных эллиптических уравнений вблизи границы.
Докл. АН СССР I84, Но.3, 534-537 (1969)

/6/ Lions, J.-L., et Magenes, E.: Problèmes aux limites non homogènes et applications. Paris 1968

/7/ Ройтберг, Я.А.: О значениях на границе области обобщенных решений.
Матем. сб. 86 (I28), 248-267 (1971)

/8/ Schulze, B. W., und Wildenhain, G.: Methoden der Potentialtheorie für elliptische Differentialgleichungen beliebiger Ordnung. Berlin 1977, Basel 1977

eingegangen: 05. 06. 1980

Anschrift des Verfassers:

Prof. Dr. rer. nat. habil. Günther Wildenhain
Wilhelm-Pieck-Universität Rostock
Sektion Mathematik
Universitätsplatz 1
DDR-2500 Rostock

Rostock. Math. Kolloq. 16, 59 - 74 (1981)

Günther Wildenhain

Über die Darstellbarkeit polyharmonischer Funktionen in der Kugel

1. Bezeichnet $K_z^R \subset \mathbb{R}^n$ eine (abgeschlossene) Kugel mit dem Mittelpunkt z und dem Radius R, S_z^R den Rand der Kugel und g(y) eine auf S_z^R vorgegebene stetige Randfunktion, so liefert bekanntlich das sogenannte Poisson-Integral

$$u(x) = \frac{1}{R\omega_n} \int_{S_z^R} \frac{R^2 - |x-z|^2}{|x-y|^n} g(y) d\sigma(y) \tag{1}$$

(ω_n Inhalt der n-dimensionalen Einheitskugeloberfläche, $d\sigma(y)$ gewöhnliches Oberflächemaß) für $|x| < R$ die Lösung des Dirichlet-Problems

$$\Delta u = 0 \text{ in } K_z^R \setminus S_z^R, \quad u|_{S_z^R} = g$$

(vgl. /7/). Man kann allgemeiner die Frage stellen, unter welchen Bedingungen eine in $K_z^R \setminus S_z^R$ harmonische Funktion durch ein allgemeines Poisson-Integral in der Gestalt

$$u(x) = \int_{S_z^R} \frac{R^2 - |x-z|^2}{|x-y|^n} d\mu(y) \tag{2}$$

mit Hilfe eines Maßes μ mit dem Träger auf S_z^R darstellbar ist. Eine notwendige und hinreichende Bedingung dafür ist

$$\sup_{0<r<R} \left(\int_{S_z^r} |u(y)| d\sigma(y) \right) < \infty,$$

d. h. eine gewisse Einschränkung an das Verhalten in Randnähe.

Dies läßt sich weiter präzisieren: Genau dann besitzt u eine Darstellung gemäß (2) mit

$$d\mu(y) = f(y)d\sigma(y),$$

$$f \in L^p(S_z^R) \quad (1 < p \leq \infty),$$

wenn

$$\sup_{0<r<R} \Big(\int_{S_z^r} |u(y)|^p d\sigma(y)\Big)^{\frac{1}{p}} < \infty$$

gilt (vgl. /11/). In der vorliegenden Arbeit werden wir eine Verallgemeinerung des hinreichenden Teils dieser Aussage sowohl für Lösungen der homogenen als auch der inhomogenen polyharmonischen Gleichung beweisen.

2. Die Lösung des Dirichlet-Problems

$$\Delta^m u = 0 \text{ in } K_z^R \setminus S_z^R,$$

$$\frac{\partial^j u}{\partial n^j}\Big|_{S_z^R} = g_j \qquad (j=0,\dots,m-1)$$

($m \geq 1$ ganz, n Innennormale an S_z^R, $g_j \in C^{m-j-1}(S_z^R)$) im Punkt $x \in K_z^R \setminus S_z^R$ ist gegeben durch

$$u(x) = \frac{(-1)^{m-1}(R^2-|x-z|^2)^m}{(m-1)!R^{n-1}\omega_n} \sum_{j=0}^{m-1} \int_{S_z^R} \binom{m-1}{j} g_j(y) \frac{\partial^{m-j-1}}{\partial(R^2)^{m-j-1}} \frac{R^{n-2}}{|x-y|^n} d\sigma(y) \tag{3}$$

(vgl. /5/, /10/). Im folgenden betrachten wir o.B.d.A. $z = 0$ und schreiben kurz ($S^R=S_o^R$, $K^R=K_o^R$)

$$u(x) = \sum_{j=0}^{m-1} \int_{S^R} P_j(x,y)g_j(y)d\sigma(y).$$

Die Funktionen $P_j(x,y)$ werden als Poisson-Kerne bezeichnet. Die uns interessierende Frage lautet dann: Unter welchen Bedingungen läßt sich eine in $K^R \setminus S^R$ m-polyharmonische Funktion als verallgemeinertes Poisson-Integral

$$P(\mu_0, \mu_1, \dots, \mu_{m-1}; S^R)(x) := \sum_{j=0}^{m-1} \int P_j(x,y) d\mu_j(y)$$

durch ein Maßtupel $(\mu_0, \mu_1, \dots, \mu_{m-1})$ mit dem Träger auf S^R darstellen? In /6/ wurde gezeigt, daß die Bedingung

$$\sum_{j=0}^{m-1} \int_{S^{R-\delta}} \left|\frac{\partial^j u}{\partial r^j}(y)\right| d\sigma(y) \leq C < \infty \qquad \text{für } 0 < \delta \leq \delta_0$$

dafür hinreichend ist, wobei $\frac{\partial}{\partial r}$ die Ableitung in radialer Richtung bedeutet. Im folgenden wollen wir ein präziseres Resultat beweisen. Auf S^R sowie auf S^ϱ $(\varrho \leq R)$ führen wir sphärische Koordinaten ein. Den zugehörigen $(n-1)$-dimensionalen Raum der Koordinaten $(\varphi_1, \dots, \varphi_{n-1})$ bezeichnen wir mit S. Ist f eine auf S^ϱ erklärte Funktion, so können wir sie gemäß

$$f(y) = f(\varphi_1, \dots, \varphi_{n-1}, \varrho) = f(\varphi, \varrho)$$

als Funktion auf S auffassen, wobei ϱ die Rolle eines Parameters spielt. Unter $W_p^l(S)$ $(l \geq 0$ ganz, $1 < p < \infty)$ verstehen wir dann den gewöhnlichen Sobolev-Raum aller Funktionen $u(\varphi) \in L^p(S)$, für welche alle distributionentheoretischen Ableitungen (bezüglich $\varphi_1, \dots, \varphi_{n-1}$) bis zur Ordnung $\leq l$ ebenfalls in $L^p(S)$ liegen. Die Norm in $W_p^l(S)$ ist definiert durch

$$\|u\|_{p,l} := \left\{ \sum_{|\alpha| \leq l} \int_S |D^\alpha u(\varphi)|^p d\varphi_1 \dots d\varphi_{n-1} \right\}^{\frac{1}{p}}$$

$$\left(D^\alpha = \frac{\partial^{|\alpha|}}{\partial \varphi_1^{\alpha_1} \dots \partial \varphi_{n-1}^{\alpha_{n-1}}},\ |\alpha| = \alpha_1 + \dots + \alpha_{n-1},\ \alpha_i \geq 0 \text{ ganz}\right).$$

Gehen wir speziell von Funktionen $u(y) = u(\varphi, \varrho)$ aus, die auf

S^{ϱ} erklärt sind, so schreiben wir $W_p^l(S^{\varrho})$. Wir können $W_p^l(S^{\varrho})$ in $W_p^l(S)$ eingebettet denken.

Satz 1: In $K^R \setminus S^R$ sei eine m-polyharmonische Funktion, d. h. eine Lösung der Gleichung $\Delta^m u = 0$, gegeben. Für $j=0,1,\ldots,m-1$ und eine nichtnegative ganze Zahl l gelte

$$\left\| \frac{\partial^j u(\varphi,\varrho)}{\partial \varrho^j} \right\|_{p,l} \leq C < \infty \text{ für } R - \varepsilon_o \leq \varrho < R. \tag{4}$$

Dann existieren Funktionen $f_j \in W_p^l(S^R)$ $(j=0,1,\ldots,m-1)$ derart, daß u für $|x| < R$ in der Gestalt

$$u(x) = \sum_{j=o}^{m-1} \int_{S^R} P_j(x,y) f_j(y) d\sigma(y) \tag{5}$$

darstellbar ist.

Beweis: Für $y \in S^{\varrho}$ setzen wir

$$\frac{\partial^j u}{\partial \varrho^j}(y) = \frac{\partial^j u}{\partial \varrho^j}(\varphi_1,\ldots,\varphi_{n-1},\varrho) = u^{(j)}(\varphi,\varrho)$$

$(j=0,1,\ldots,m-1)$.

Nach Voraussetzung gilt

$$\|u^{(o)}(\varphi,\varrho)\|_{p,l} \leq C \quad \text{für} \quad R - \varepsilon_o \leq \varrho < R.$$

Da eine bezüglich der Norm beschränkte Funktionenmenge in $W_p^l(S)$ schwach kompakt ist, existiert eine Folge $(\varrho_i^{(o)})$ mit $\lim\limits_{i\to\infty} \varrho_i^{(o)} = R$ derart, daß die Folge $u^{(o)}(\varphi,\varrho_i^{(o)})$ für $i \to \infty$ in $W_p^l(S)$ schwach gegen ein Element $f_o \in W_p^l(S)$ konvergiert $(u^{(o)}(\varphi,\varrho_i^{(o)}) \rightharpoonup f_o)$. Es gilt also

$$\lim_{i\to\infty} F(u^{(o)}(\varphi,\varrho_i^{(o)}) - f_o(\varphi)) = 0 \tag{6}$$

für alle stetigen linearen Funktionale $F \in (W_p^1(S))'$. Für feste $x \in K^R$ mit $|x| < R - \varepsilon_o$ und $y \in S^R$ setzen wir

$$v_j(\varphi) := P_j(x,y) = \tilde{P}_j(\varphi,R) \quad (o \leq j \leq m-1).$$

Da auf Grund der konkreten Gestalt der Poisson-Kerne (vgl. (3)) offenbar $v_j \in L^q(S) \quad (\frac{1}{p}+\frac{1}{q} = 1)$ ist, gilt für

$$F_{v_j}(h) := \int_S h(\varphi)v_j(\varphi)d\varphi$$

die Abschätzung

$$|F_{v_j}(h)| \leq \|h\|_{L^p(S)} \|v_j\|_{L^q(S)} \leq \|v_j\|_{L^q(S)} \|h\|_{p,1}$$

für alle $h \in W_p^1(S)$, d. h., $F_{v_j} \in (W_p^1(S))'$ für alle j. Setzen wir speziell $F = F_{v_o}$ in (6) ein, so folgt

$$\lim_{i\to\infty} F_{v_o}(u^{(o)}(\varphi,\varrho_i^{(o)}) - f_o(\varphi)) = 0, \text{ d. h.,}$$

$$\lim_{i\to\infty} \int_S \tilde{P}_o(\varphi,R)u^{(o)}(\varphi,\varrho_i^{(o)})d\varphi$$
$$= \int_S \tilde{P}_o(\varphi,R)f_o(\varphi)d\varphi = \int_{S^R} P_o(x,y)f_o(y)d\sigma(y).$$

Wegen

$$\|u^{(1)}(\varphi,\varrho_i^{(o)})\|_{p,1} \leq C \quad \text{für alle } i$$

existiert weiter eine Teilfolge $(\varrho_i^{(1)})$ von $(\varrho_i^{(o)})$ mit

$$\lim_{i\to\infty} \varrho_i^{(1)} = R \text{ und } u^{(1)}(\varphi,\varrho_i^{(1)}) \rightharpoonup f_1 \in W_p^1(S) \text{ für } i\to\infty.$$

Wendet man mit $F = F_{v_1}$ wieder die obige Überlegung an, so erhält man anlog

$$\lim_{i\to\infty} \int_S \tilde{P}_1(\varphi,R)u^{(1)}(\varphi,\varrho_i^{(1)})d\varphi$$
$$= \int_S \tilde{P}_1(\varphi,R)f_1(\varphi)d\varphi = \int_{S^R} P_1(x,y)f_1(y)d\sigma(y).$$

Setzt man das Verfahren fort, so ergibt sich für $j = 0,1,\dots,m-1$ die Existenz von Funktionen $f_j \in W_p^1(S^R)$ mit

$$\lim_{i\to\infty} \int_S \tilde{P}_j(\varphi,R)u^{(j)}(\varphi,\varrho_i^{(j)})d\varphi$$

$$= \int_S \tilde{P}_j(\varphi,R)f_j(\varphi)d\varphi = \int_{S^R} P_j(x,y)f_j(y)d\sigma(y).$$

Dabei ist jeweils $(\varrho_i^{(j)})$ eine Teilfolge von $(\varrho_i^{(j-1)})$.

Für $\varrho_i := \varrho_i^{(m-1)}$ gilt also

$$\lim_{i\to\infty} \sum_{j=0}^{m-1} \int_S \tilde{P}_j(\varphi,R)u^{(j)}(\varphi,\varrho_i)d\varphi$$

$$= \sum_{j=0}^{m-1} \int_{S^R} P_j(x,y)f_j(y)d\sigma(y). \tag{7}$$

Für feste x mit $|x| < \varrho$ und $y \in S^\varrho$ setzen wir nun

$$\tilde{P}_{j,\varrho}(\varphi,\varrho) = P_{j,\varrho}(x,y) = P_j(x,y).$$

Da u als m-polyharmonische Funktion in $K^R \setminus S^R$ beliebig oft differenzierbar ist, gilt $u^{(j)}(\varphi,\varrho_i) \in C^\infty(S^{\varrho_i})$ für $j = 0, 1, \dots, m-1$ und alle i. Nach dem allgemeinen Darstellungssatz für die Lösung des Dirichlet-Problems mit glatten Randwerten (Formel (3)) läßt sich daher $u(x)$ für $|x| < R - \varepsilon_0$ und alle i in der Gestalt

$$u(x) = \sum_{j=0}^{m-1} \int_S \tilde{P}_{j,\varrho_i}(\varphi,\varrho_i)u^{(j)}(\varphi,\varrho_i)d\varphi$$

$$= \sum_{j=0}^{m-1} \int_{S^{\varrho_i}} P_j(x,y)\,\frac{\partial^j u}{\partial\varrho^j}(y)d\sigma(y) \tag{8}$$

schreiben. Dies ermöglicht die Abschätzung

$$|u(x) - \sum_{j=o}^{m-1} \int_S \tilde{P}_j(\varphi,R)u^{(j)}(\varphi,\varrho_i)d\varphi|$$

$$=|\sum_{j=o}^{m-1} \int_S \tilde{P}_{j,\varrho_i}(\varphi,\varrho_i)u^{(j)}(\varphi,\varrho_i)d\varphi - \sum_{j=o}^{m-1} \int_S \tilde{P}_j(\varphi,R)u^{(j)}(\varphi,\varrho_i)d\varphi|$$

$$=|\sum_{j=o}^{m-1} \int_S \left[\tilde{P}_{j,\varrho_i}(\varphi,\varrho_i) - \tilde{P}_j(\varphi,R)\right] u^{(j)}(\varphi,\varrho_i)d\varphi| \tag{9}$$

$$\leq \sum_{j=o}^{m-1} \|\tilde{P}_{j,\varrho_i}(\varphi,\varrho_i) - \tilde{P}_j(\varphi,R)\|_{L^q(S)} \|u^{(j)}(\varphi,\varrho_i)\|_{p,1}$$

$$\leq C \sum_{j=o}^{m-1} \|\tilde{P}_{j,\varrho_i}(\varphi,\varrho_i)-\tilde{P}_j(\varphi,R)\|_{L^q(S)} \leq C_1 \sup_{\varphi,j} |\tilde{P}_{j,\varrho_i}(\varphi,\varrho_i)-\tilde{P}_j(\varphi,R)| .$$

Dabei wurde die Hölder-Ungleichung und die Voraussetzung (4) des Satzes benutzt. Da auf Grund der Struktur der Poisson-Kerne die rechte Seite der obigen Ungleichung für $i \to \infty$ offenbar gegen Null strebt, erhalten wir

$$u(x) = \lim_{i \to \infty} \sum_{j=o}^{m-1} \int_S \tilde{P}_j(\varphi,R)u^{(j)}(\varphi,\varrho_i)d\varphi. \tag{10}$$

Vergleicht man (7) und (10), so ergibt sich schließlich

$$u(x) = \sum_{j=o}^{m-1} \int_{S^R} P_j(x,y)f_j(y)d\sigma(y).$$

Es sei noch vermerkt, daß die Annahme $|x| < R - \varepsilon_o$ keine Einschränkung bedeutet, denn ist x mit $|x| < R$ gegeben, so können wir anschließend $\varepsilon_o > 0$ so klein wählen, daß $|x| < R - \varepsilon_o$ erfüllt ist.

Bemerkung: Aus dem Beweis von Satz 1 ist ersichtlich, daß die Aussage im wesentlichen bestehen bleibt, falls man an Stelle von (4)

$$\left\| \frac{\partial^j u(\varphi,\varrho)}{\partial \varrho^j} \right\|_{p,l_j} \le C_j < \infty \quad \text{für} \quad R - \varepsilon_0 \le \varrho < R$$

($0 \le j \le m-1$) voraussetzt, wobei jetzt die $l_j \ge 0$ verschiedene nichtnegative ganze Zahlen sein dürfen. Es gilt dann die Darstellung (5) mit $f_j \in W_p^{l_j}(S^R)$ $\quad (j=0,1,\dots,m-1)$.
Im Beweis hat man lediglich bei der Abschätzung (9) eine geringfügige Modifikation vorzunehmen.

3. Die Aussage von Satz 1 kann noch weiter modifiziert und in Zusammenhang mit allgemeineren Resultaten von J. M. Berezanskij und J. A. Rojtberg (vgl. /3/, /4/, /9/) gebracht werden. Wir wollen diese Aspekte hier nur andeuten und daran anschließende, tiefergehende Untersuchungen einer folgenden Arbeit überlassen.
Es sei $\Omega \subset \mathbb{R}^n$ ein beschränktes Gebiet mit beliebig oft differenzierbarem Rand $\partial\Omega$. Für ganzzahlige $l \ge 0$ bezeichnen wir mit $W_p^{l-\frac{1}{p}}(\partial\Omega)$ $(1 < p < \infty)$ den Vektorraum aller Funktionen φ auf $\partial\Omega$, die (im Sinne der Einbettungssätze, vgl. /3/, S. 86) Randwerte einer Funktion aus $W_p^l(\Omega)$ sind.
Durch

$$\|\varphi\|_{p,l-\frac{1}{p}} := \inf \|u\|_{p,l}$$

(Infimum erstreckt über alle $u \in W_p^l(\Omega)$ mit $u|_{\partial\Omega} = \varphi$) wird $W_p^{l-\frac{1}{p}}(\partial\Omega)$ zu einem Banachraum. Den zu $W_p^{l-\frac{1}{p}}(\partial\Omega)$ dualen Raum $W_q^{-l+\frac{1}{p}}(\partial\Omega)$ $(\frac{1}{p}+\frac{1}{q}=1)$ erhält man durch Vervollständigung des Raumes $C(\partial\Omega)$ bezüglich der Norm

$$\|\psi\|_{q,-l+\frac{1}{p}} := \sup_{\varphi \in W_p^{l-\frac{1}{p}}(\partial\Omega)} \frac{|(\psi,\varphi)|}{\|\varphi\|_{p,l-\frac{1}{p}}}$$

$((\psi,\varphi) = \int_{\partial\Omega} \psi(y)\overline{\varphi(y)}d\sigma(y))$. Im Sinne dieser Dualität läßt sich in der üblichen Weise das Skalarprodukt (ψ,φ) für beliebige $\varphi \in W_p^{l-\frac{1}{p}}(\partial\Omega)$, $\psi \in W_q^{-l+\frac{1}{p}}(\partial\Omega)$ erklären (vgl. /3/). Der Beweis von Satz 1 kann nun weiter so modifiziert werden, daß sich die folgende Aussage ergibt: Unter der Voraussetzung

$$\left\| \frac{\partial^j u(\varphi,\varrho)}{\partial\varrho^j} \right\|_{p,l-j-\frac{1}{p}} \leq C_j < \infty \quad \text{für } R-\varepsilon_0 \leq \varrho < R \tag{11}$$

($0 \leq j \leq m-1$, $1 < p < \infty$, $l > \frac{n}{2}$ ganz) existieren Funktionen $f_j \in W_p^{l-j-\frac{1}{p}}(S^R)$, so daß für $|x| < R$ die Darstellung

$$u(x) = \sum_{j=0}^{m-1} \int_{S^R} P_j(x,y)f_j(y)d\sigma(y) \tag{12}$$

besteht. Für $l - j - \frac{1}{2} < 0$ ist das Integral hierbei im Sinne der Dualität zwischen $W_p^{l-j-\frac{1}{p}}(S^R)$ und $W_q^{-l+j+\frac{1}{p}}(S^R)$ zu verstehen. Aus Resultaten von J. M. Berezanskij und J. A. Rojtberg /4/ kann man folgern:

Gilt $l > \frac{n}{2}$ und $f_j \in W_p^{l-j-\frac{1}{p}}(S^R)$, so existiert eine eindeutige Lösung des verallgemeinerten Dirichlet-Problems der Gleichung $\Delta^m u = 0$ in K^R zu den Randwerten f_j. Die Lösung liegt im Raum $W_p^l(K^R)$ und läßt sich für $|x| < R$ in der Gestalt

$$u(x) = \sum_{j=0}^{m-1} (f_j, P_j(x,\cdot)) \tag{13}$$

darstellen, wobei die Skalarprodukte im Sinne der Dualität zwischen $W_p^{l-j-\frac{1}{p}}(S^R)$ und $W_p^{-l+j+\frac{1}{p}}(S^R)$ zu verstehen sind. Der Zusatz "verallgemeinert" bezieht sich auf eine in gewissem Sinne verallgemeinerte Randwertannahme, die wir hier nicht näher beschreiben wollen (vgl. /3/, /4/). Da (13) mit (12) übereinstimmt, können wir weiter schließen, daß die m-polyharmonische Funktion u unter der Voraussetzung (11) im Raum $W_p^l(K^R)$ liegt. Aus dem klassischen Sobolevschen Einbettungssatz folgt insbesondere $u \in C^k(\bar{K}^R)$ für $0 \leq k < l - \frac{n}{2}$, d. h. Glattheit mit Einschluß des Randes aller Ableitungen bis zur Ordnung k.

Wir bemerken noch, daß für $l > \frac{n}{2}$ die Darstellbarkeit der m-polyharmonischen Funktion u in der Gestalt (13) durch Funktionen $f_j \in W_p^{l-j-\frac{1}{p}}(S^R)$ zur Eigenschaft $u \in W_p^l(K^R)$ äquivalent ist. Auch dies folgt aus Resultaten von J. A. Rojtberg /9/.

4. Um den Satz 1 auf Lösungen der inhomogenen Gleichung $\Delta^m u = g$ zu verallgemeinern, benötigen wir einige vorbereitende Betrachtungen.

Aus der sogenannten Agmon-Miranda-Ungleichung (vgl. /2/, /10/) folgt die Existenz einer von u unabhängigen Konstanten $C_R > 0$, so daß für jede in K^R m-polyharmonische Funktion u mit $u \in C^{m-1}(\bar{K}^R)$ die Abschätzung

$$\sum_{|\alpha| \leq m-1} \sup_{x \in \bar{K}^R} |D^\alpha u(x)| \leq C_R \sum_{|\alpha| \leq m-1} \sup_{x \in S^R} |D^\alpha u(x)| \tag{14}$$

besteht. In der analogen Ungleichung für die Kugel K^ϱ $(\varrho < R)$ soll die Konstante mit C_ϱ bezeichnet werden.

Lemma 1: Es existiert eine Konstante $C > 0$ mit $C_\varrho \leq C$ für $R - \varepsilon_0 \leq \varrho \leq R$.

Beweis: Wir führen den Beweis hier nur für $m = n = 2$, indem wir auf ein für diesen Fall bekanntes, allgemeineres Resultat von C. Miranda /8/ und G. Adler /1/ zurückgreifen. Es sei $\Omega \subset \mathbb{R}^2$ ein beschränktes Gebiet mit glattem Rand, $x = (x_1, x_2) \in \Omega$, $y = (y_1, y_2) \in \partial\Omega$, $\omega = \sphericalangle(x-y, n)$ mit der Orientierung von $x - y$ nach n (n Innennormale). Wir betrachten ein System kartesischer Koordinaten (ξ_1^y, ξ_2^y) mit dem Ursprung in y. Die positive ξ_2^y-Achse habe die Richtung von n. $Z_y(\tau)$ bezeichne den Teil der Kurve $\partial\Omega$ mit $-\tau \leq \xi_1^y \leq \tau$.

Wir nennen Ω zur Klasse $\mathfrak{A}(\tau)$ gehörig, wenn eine Zahl $\tau > 0$ derart gefunden werden kann, daß es in jedem Punkt $y \in \partial\Omega$ möglich ist

(i) einen Kreis K_i^j mit dem Radius τ zu konstruieren, dessen Rand y enthält und dessen Inneres in Ω liegt,

(ii) einen Kreis K_a^y mit dem Radius τ und folgenden Eigenschaften zu konstruieren:

- K_a enthält mindestens einen Punkt, welcher nicht zu Ω gehört.
- Der Rand von K_a^y enthält den Punkt y.
- K_a^y enthält keinen weiteren Punkt von $Z_y(\tau)$.

Unter der Voraussetzung $\Omega \in \mathfrak{A}(\tau)$ wurde in /1/, /8/ bewiesen, daß für jede in $\Omega \subset \mathbb{R}^2$ biharmonische Funktion $u \in C^1(\overline{\Omega})$ die folgenden Ungleichungen gelten:

$$\max_{x \in \overline{\Omega}} |\operatorname{grad} u(x)| \leq K_3 \max_{\partial\Omega} |\varphi_1| + K_1 \max_{\partial\Omega} |\varphi_2| + K_2 \max_{\partial\Omega} |\varphi_1'|,$$

$$|u(x)| \leq \delta(x)(K_1 \max_{\partial\Omega} |\varphi_2| + K_2 \max_{\partial\Omega} |\varphi_1'|) + (K_3\delta(x)+1) \max_{\partial\Omega} |\varphi_1|.$$

Dabei ist $\delta(x)$ der Abstand des Punktes $x \in \overline{\Omega}$ vom Rand $\partial\Omega$,

$K_1 = \sqrt{56}$, $K_2 = 26400 + (41+25\,\Phi)\,\frac{L}{\tau}$,

$K_3 = \frac{78}{\tau}$, $\Phi = \sup\limits_{x \in \Omega} \max\limits_{\partial\Omega} |\omega|$, L die Länge von $\partial\Omega$,

$u|_{\partial\Omega} = \varphi_1$, $\frac{\partial u}{\partial n}|_{\partial\Omega} = \varphi_2$, und φ_1' bedeutet die Ableitung nach der Bogenlänge.

Offensichtlich liefern diese Ungleichungen für $\Omega = K^R \setminus S^R$ Abschätzungen für die Konstante C_R. Aus der Gestalt der Konstanten K_1, K_2, K_3 ist ersichtlich, daß sich diese beim Übergang von K^R zu K^ϱ ($\varrho < R$) stetig ändern, da sich offenbar Φ und L stetig ändern und da dieser Übergang dem Übergang zu einer anderen Gebietsklasse $\mathfrak{A}(\tau)$ mit stetiger Änderung des Parameters τ entspricht. Daraus folgt die Behauptung.

Wir betrachten jetzt die zum homogenen Dirichlet-Problem der Gleichung $\Delta^m u = g$ gehörige Greensche Funktion $G_R(x,y)$ der Kugel $K^R \subset \mathbb{R}^n$. $G_\varrho(x,y)$ sei die entsprechende Greensche Funktion zu K^ϱ ($\varrho < R$).

<u>Lemma 2:</u> Es sei x ein fester Punkt mit $|x| < R$. Dann gilt gleichmäßig bezüglich $y \in K^R$, $y \neq x$, $\lim\limits_{\varrho \to R} |G_R(x,y) - G_\varrho(x,y)| = 0$.

<u>Beweis:</u> Bezeichnet $E(x,y)$ die Fundamentallösung des m-polyharmonischen Operators mit der Singularität in x, so gilt bekanntlich (vgl. /10/)

$$G_R(x,y) = E(x,y) + u_x(y),$$

wobei $u_x(y)$ die eindeutig bestimmte Lösung des Dirichlet-Problems der Gleichung $\Delta^m u = 0$ in K^R mit

$$\left.\frac{\partial^j u_x}{\partial n^j}\right|_{S^R} = -\left.\frac{\partial^j E(x,y)}{\partial n^j}\right|_{S^R} \qquad (j=0,1,\dots,m-1)$$

ist. Entsprechend gilt

$$G_\varrho(x,y) = E(x,y) + u_x^{(\varrho)}(y),$$

$$\left.\frac{\partial^j u_x^{(\varrho)}}{\partial n^j}\right|_{S^\varrho} = -\left.\frac{\partial^j E(x,y)}{\partial n^j}\right|_{S^\varrho} \qquad (j=0,1,\dots,m-1).$$

Betrachten wir in K^ϱ die Differenz $G_R(x,y) - G_\varrho(x,y)$, so folgt aus (14)

$$|G_R(x,y) - G_\varrho(x,y)| = |u_x(y) - u_x^{(\varrho)}(y)|$$

$$\leq C_\varrho \sum_{j=o}^{m-1} \sup_{y \in S^\varrho} \left| \frac{\partial^j (u_x(y) - u_x^{(\varrho)}(y))}{\partial n^j} \right|$$

$$= C_\varrho \sum_{j=o}^{m-1} \sup_{y \in S^\varrho} \left| \frac{\partial^j}{\partial n^j}(u_x(y) + E(x,y)) \right| = C_\varrho \sum_{j=o}^{m-1} \sup_{y \in S^\varrho} \left| \frac{\partial^j G_R(x,y)}{\partial n^j} \right|.$$

Wegen $\left.\dfrac{\partial^j G_R(x,y)}{\partial n^j}\right|_{S^R} = 0$ $(j=0,1,\ldots,m-1)$, der Stetigkeit dieser Ableitungen in Randnähe und $C_\varrho \leq C$ (Lemma 1) strebt die rechte Seite für $\varrho \to R$ gleichmäßig gegen Null, womit die Behauptung bewiesen ist.

5. Nunmehr kann Satz 1 auf Lösungen der inhomogenen Gleichung verallgemeinert werden.

Satz 2: Es sei $g \in C^{\varkappa}(K^R)$ $(0 < \varkappa < 1)$, und u sei innerhalb K^R (klassische) Lösung der Gleichung $\Delta^m u = g$. Für $j = 0,1,\ldots,m-1$ und eine nichtnegative ganze Zahl l gelte

$$\left\| \frac{\partial^j u(\varphi,\varrho)}{\partial \varrho^j} \right\|_{p,l} \leq C < \infty \text{ für } R - \varepsilon_0 \leq \varrho < R.$$

Dann existieren Funktionen $f_j \in W_p^l(S^R)$ $(j=0,1,\ldots,m-1)$ derart, daß u für $|x| < R$ in der Gestalt

$$u(x) = \int_{K^R} G_R(x,y) g(y) dy + \sum_{j=o}^{m-1} \int_{S^R} P_j(x,y) f_j(y) d\sigma(y)$$

darstellbar ist.

Beweis: Wir setzen (mit den Bezeichnungen des Beweises von Satz 1)

$$v(x) := \int_{K^R} G_R(x,y)g(y)dy + \sum_{j=0}^{m-1} \int_S \tilde{P}_j(\varphi, R)u^{(j)}(\varphi, \varrho_i)d\varphi. \quad (15)$$

Ferner läßt sich u(x) innerhalb S^{ϱ_i} analog zu (8) in der Gestalt

$$u(x) = \int_{K^{\varrho_i}} G_{\varrho_i}(x,y)g(y)dy + \sum_{j=0}^{m-1} \int_S \tilde{P}_j(\varphi, \varrho_i)u^{(j)}(\varphi, \varrho_i)d\varphi$$

schreiben. Somit ergibt sich die Abschätzung

$$|u(x) - v(x)| \le \left| \int_{K^{\varrho_i}} G_{\varrho_i}(x,y)g(y)dy - \int_{K^R} G_R(x,y)g(y)dy \right| + |w_i(x)|, \quad (16)$$

wobei $w_i(x)$ mit der linken Seite der Ungleichung (9) identisch ist. Wegen (9) gilt also $w_i(x) \to 0$ für $i \to \infty$. Der erste Summand auf der rechten Seite von (16) läßt sich wie folgt abschätzen:

$$\left| \int_{K^{\varrho_i}} G_{\varrho_i}(x,y)g(y)dy - \int_{K^R} G_R(x,y)g(y)dy \right|$$

$$= \left| \int_{K^{\varrho_i}} G_R(x,y)g(y)dy + \int_{K^R \setminus K^{\varrho_i}} G_R(x,y)g(y)dy - \int_{K^{\varrho_i}} G_{\varrho_i}(x,y)g(y)dy \right|$$

$$\le \int_{K^{\varrho_i}} |G_R(x,y) - G_{\varrho_i}(x,y)| \, |g(y)| dy + \int_{K^R \setminus K^{\varrho_i}} |G_R(x,y)| \, |g(y)| dy.$$

Wegen Lemma 2 sowie wegen der Stetigkeit von $G_R(x,y)$ in Randnähe strebt die rechte Seite für $i \to \infty$ gegen Null. Unter Berücksichtigung von (15) folgt daraus die Behauptung. Es ist klar, daß Satz 2 analog wie Satz 1 modifiziert werden kann.

Literatur

/1/ Adler, G.: Maggiorazione del gradiente delle funzioni biarmoniche di due variabili.
Rend. Accad. Sci. Fis. Mat. Napoli IV, Ser. 28, 225 - 239 (1961)

/2/ Agmon, S.: Maximum theorems for solutions of higher order elliptic equations.
Bull. Amer. Math. Soc. 66, 77 - 80 (1960)

/3/ Березанский, Ю.М.: Разложение по собственным функциям самосопряженных операторов. Киев 1965

/4/ Березанский, Ю.М., и Ройтберг, Я.А.: Теорема о гомеоморфизмах и функция Грина для общих эллиптических граничных задач.
Укр. матем. ж. 19, Но. 5, 3 - 32 (1967)

/5/ Edenhofer, J.: Eine Integraldarstellung der Lösung der Dirichletschen Aufgabe bei der Polypotentialgleichung im Falle einer Hyperkugel.
Math. Nachr. 69, 149 - 162 (1975)

/6/ Gonzáles, L., y Wildenhain, G.: Sobre la representación de Funciones Poliarmónicos en una bola en $\mathbb{R}^n$.
Revista Sociedad Cubana de Matematicos, Habana, Cuba (en publicación)

/7/ Helms, L. L.: Introduction to potential theory.
New York 1969

/8/ Miranda, C.: Formule di maggiorazione e teorema di esistenza per le funzioni biarmoniche in due variabili.
Giorn. Mat. Battaglini 78, 97 - 118 (1949)

/9/ Ройтберг, Я.А.: О значениях на границе области обобщенных решений.
Матем. сб. 86 (128), 248-267 (1971)

/10/ Schulze, B. W., und Wildenhain, G.: Methoden der Potentialtheorie für elliptische Differentialgleichungen beliebiger Ordnung.
Berlin 1977, Basel 1977

/11/ Stein, E. M.: Singular integrals and differentiability properties of functions.
Princeton, New Jersey 1970

eingegangen: 05. 06. 1980

Anschrift der Verfassers:

Prof. Dr. rer. nat. habil. Günther Wildenhain
Wilhelm-Pieck-Universität Rostock
Sektion Mathematik
Universitätsplatz 1
DDR-2500 Rostock

Rostock. Math. Kolloq. 16, 75 - 80 (1981)

Lothar Berg

Ein Grenzfall beim Goursatschen Anfangswertproblem

Betrachtet wird die lineare partielle Differentialgleichung zweiter Ordnung

$$a_2 z_{xx} + a_1 z_{xy} + a_0 z_{yy} = f(x,y) \tag{1}$$

mit nicht verschwindenden konstanten Koeffizienten unter den Goursatschen Anfangsbedingungen

$$z(x,0) = g(x), \quad z(0,y) = h(y), \tag{2}$$

wobei die Verträglichkeitsbedingung $g(0) = h(0)$ erfüllt sei. Die gegebenen Funktionen $f(x,y)$, $g(x)$, $h(y)$ seien für $x = y = 0$ holomorph, und die gesuchte Funktion $z(x,y)$ soll die gleiche Holomorphieeigenschaft besitzen, wobei x und y als komplexe Veränderliche aufzufassen sind. Nach einem Satz von L. Hörmander /4/, S. 116 - 118, wird die eindeutige Lösbarkeit dieses Anfangswertproblems garantiert, wenn

$$|a_0| + |a_2| < \frac{1}{4} e^{-2} |a_1|$$

gilt. Dies bedeutet für die Wurzeln λ_1, λ_2 mit $|\lambda_1| \geq |\lambda_2|$ der zu (1) gehörenden charakteristischen Gleichung

$$a_2 \lambda^2 + a_1 \lambda + a_0 = 0, \tag{3}$$

daß $|\lambda_1|$ hinreichend groß und $|\lambda_2|$ hinreichend klein sein muß. In /1/ wurde die eindeutige Lösbarkeit für beliebige Wurzeln von (3) mit $|\lambda_1| > |\lambda_2|$ bewiesen, so daß wir uns jetzt dem Grenzfall $|\lambda_1| = |\lambda_2|$ zuwenden.

Spezielle Lösungen: Zunächst beweisen wir für die allgemeinere lineare Gleichung

$$\sum_{i=0}^{n} a_i \frac{\partial^n z(x,y)}{\partial x^i \partial y^{n-i}} = f(x,y) \tag{4}$$

mit konstanten Koeffizienten, die nicht alle verschwinden sollen, ohne Berücksichtigung von Anfangsbedingungen den folgenden

<u>Satz 1:</u> Ist $f(x,y)$ für $|x| < r$, $|y| < s$ holomorph, so besitzt (4) stets eine spezielle Lösung, die für $|x| < r/2$, $|y| < s/2$ holomorph ist.

<u>Beweis:</u> Mit Hilfe der Differentiationsoperatoren $D_x = \partial/\partial x$, $D_y = \partial/\partial y$ läßt sich (4) bei passender Normierung der Koeffizienten in der Form

$$\prod_{i=1}^{p} (D_x - \alpha_i D_y) \prod_{j=1}^{q} (D_y - \beta_j D_x) \, z(x,y) = f(x,y) \tag{5}$$

mit $p + q = n$ schreiben, wobei α_i, β_j^{-1} die Wurzeln der zu (4) gehörenden charakteristischen Gleichung sind und $\alpha_i = 0$, $\beta_j = 0$ für gewisse i,j zugelassen ist. Die Einteilung der Wurzeln sei so vorgenommen, daß

$$|\alpha| \le \frac{s}{r}, \quad |\beta| \le \frac{r}{s} \tag{6}$$

gilt für alle $\alpha = \alpha_i$ und alle $\beta = \beta_j$.

Die Gleichungen erster Ordnung

$$u_x - \alpha u_y = w, \quad v_y - \beta v_x = w$$

besitzen bekanntlich (vgl. etwa /2/) spezielle Lösungen der Form

$$\begin{aligned} u(x,y) &= x \int_0^1 w(x(1-\xi),\, y + \alpha x \xi)\, d\xi, \\ v(x,y) &= y \int_0^1 w(x + \beta y \eta,\, y(1-\eta))\, d\eta. \end{aligned} \tag{7}$$

Wenden wir jetzt auf die Funktion

$$w(x,y) = f(ax + by, cx + dy) \tag{8}$$

mit

$$|a| \leq 1,\ |b| \leq \frac{r}{s},\ |c| \leq \frac{s}{r},\ |d| \leq 1 \tag{9}$$

eine der Integrationen (7) an, so erhalten wir bei dem ersten Integral den Integranden

$$f(ax(1-\xi) + b(y+\alpha x\xi),\ cx(1-\xi) + d(y+\alpha x\xi))$$

$$= f((a(1-\xi) + b\alpha\xi)x + by,\ (c(1-\xi) + d\alpha\xi)x + dy),$$

und da wegen (6) und (9)

$$|a(1-\xi) + b\alpha\xi| \leq 1,\ |c(1-\xi) + d\alpha\xi| \leq \frac{s}{r}$$

gilt, hat er wieder dieselbe Bauart wie (8) mit (9). Analog besitzt das zweite der Integrale (7) im Fall (8) mit (9) den Integranden

$$f(a(x+\beta y\eta) + by(1-\eta),\ c(x+\beta y\eta) + dy(1-\eta))$$

$$= f(ax + (a\beta\eta + b(1-\eta))y,\ cx + (c\beta\eta + d(1-\eta))y),$$

der wegen

$$|a\beta\eta + b(1-\eta)| \leq \frac{r}{s},\ |c\beta\eta + d(1-\eta)| \leq 1$$

ebenfalls wieder die Bauart (8) mit (9) besitzt.

Wählen wir jetzt $|x| < r/2$, $|y| < s/2$, so folgt aus (9)

$$|ax + by| < r,\ |cx + dy| < s,$$

d. h., die Argumente von (8) liegen im Holomorphiegebiet der Funktion $f(x,y)$, und die zugehörigen Integrale (7) sind ebenfalls holomorph. Da insbesondere $f(x,y)$ selbst von der Bauart (8) mit (9) ist und die Gleichung (5) sich durch schrittweise Anwendung von (7) auflösen läßt, ist Satz 1 bewiesen.

<u>Homogene Gleichungen:</u> Auf Grund von Satz 1 brauchen wir uns bei dem Anfangswertproblem (1), (2) nur noch mit dem Fall zu befassen, daß (1) eine homogene Gleichung ist. Aus (8) ist wegen (9) nämlich unmittelbar ersichtlich, daß die soeben konstruierte spezielle Lösung im Fall $x = 0$ sogar für $|y| < s$ und im Fall $y = 0$ sogar für $|x| < r$ holomorph ist.

Wie bereits angekündigt, soll jetzt für die Nullstellen der Gleichung (3) $|\lambda_1| = |\lambda_2|$ gelten, wobei wir den Fall der Doppelwurzel zunächst ausschließen.

Satz 2: Ist $g(x)$ für $|x| < r$ und $h(y)$ für $|y| < s$ holomorph und gilt für die Wurzeln der Gleichung (3)

$$\lambda_2 = \lambda_1 e^{2\pi i \vartheta}, \tag{10}$$

wobei ϑ eine algebraische Zahl ist, dann besitzt das Anfangswertproblem (1), (2) mit $f(x,y) \equiv 0$ genau eine Lösung $z(x,y)$, die für

$$|x| < \frac{1}{2} \operatorname{Min}\left(r, \frac{s}{t}\right), \; |y| < \frac{1}{2} \operatorname{Min}(tr, s) \tag{11}$$

mit $t = |\lambda_1|$ holomorph ist.

Beweis: Die allgemeine Lösung von (1) hat bekanntlich die Gestalt

$$z(x,y) = \varphi(\lambda_1 x + y) + \psi(\lambda_2 x + y). \tag{12}$$

Machen wir die Ansätze

$$g(x) = \sum_{n=0}^{\infty} g_n x^n, \quad h(y) = \sum_{n=0}^{\infty} h_n y^n$$

mit $g_0 = h_0$, so erhalten wir wie in /1/

$$\varphi(y) = \sum_{n=1}^{\infty} \frac{g_n - \lambda_2^n h_n}{\lambda_1^n - \lambda_2^n} y^n, \quad \psi(y) = h(y) - \varphi(y). \tag{13}$$

Wegen (10) gilt

$$|1-(\lambda_2/\lambda_1)^n| = |1-e^{2\pi i \vartheta n}| = 2|\sin \pi\vartheta n| \geq 4|\vartheta n - m|,$$

wenn wir m so wählen, daß $|\vartheta n - m| \leq 1/2$ ist, und $|\sin x| \geq \frac{2}{\pi}|x|$ für $|x| \leq \pi/2$ berücksichtigen. Nach einem Satz von Liouville (vgl. /3/) gibt es zu jeder algebraischen Zahl k-ten Grades eine positive Zahl c, so daß

$$|\vartheta - \frac{m}{n}| \geq cn^{-k}$$

ist für alle rationalen Zahlen m/n. Somit erhalten wir zusammen mit der vorhergehenden Abschätzung

$$|\lambda_1^n - \lambda_2^n| \geq 4ct^n n^{1-k},$$

und aus (13) ist ersichtlich, daß die Funktionen $\varphi(y)$ und $\psi(y)$ für $|y| < \mathrm{Min}\,(tr,s)$ holomorph sind. Da unter den Ungleichungen (11) für $j = 1,2$

$$|\lambda_j x + y| \leq t\,|x| + |y| < \mathrm{Min}\,(tr,s)$$

gilt, folgt aus (12) unsere Behauptung.

Bemerkungen: 1°. Im Fall der Doppelwurzel $\lambda_1 = \lambda_2 = \lambda \neq 0$ lautet die allgemeine Lösung von (1)

$$z(x,y) = \varphi(\lambda x + y) + x\,\psi(\lambda x + y),$$

wobei die Anfangsbedingungen (2)

$$\varphi(y) = h(y), \quad \psi(y) = \frac{\lambda}{y}\,(g(\tfrac{y}{\lambda}) - h(y))$$

liefern. Hieraus ist ersichtlich, daß die Aussage von Satz 2 auch in diesem Fall gültig bleibt.

2°. Ist ϑ in (10) eine rationale, nicht ganze Zahl, so erkennt man aus (13), daß das betrachtete Anfangswertproblem im allgemeinen unlösbar ist. Ist jedoch ϑ eine transzendente Zahl, so bleibt die Aussage von Satz 2 ebenfalls erhalten, wenn ϑ einer Ungleichung der Form

$$|\vartheta - \frac{m}{n}| \geq e^{-\varepsilon_n}$$

mit $\varepsilon_n > 0$ und $\varepsilon_n = o(n)$ genügt. Ist nur $\varepsilon_n = O(n)$, so besitzt (1), (2) auch noch stets eine holomorphe Lösung, aber das Holomorphiegebiet (11) verkleinert sich dann. Bei noch stärkerem Anwachsen von ε_n gibt es im allgemeinen keine holomorphe Lösung mehr. Alle existierenden Lösungen mit $|\lambda_1| = |\lambda_2|$ sind natürlich in bezug auf Störungen der Koeffizienten instabil.

3°. Eine Übertragung der Ergebnisse auf die allgemeinere Gleichung (4) dürfte zumindest im Fall der mehrfachen Wurzeln keine Schwierigkeiten bereiten. Die Ergebnisse von /1/ und /4/ beziehen sich sogar auf den Fall, daß in (4) auch Ableitungen von kleinerer Ordnung als n auftreten.

Literatur

/1/ Berg, L.: Asymptotische Abschätzung inverser Matrizen mit einer Anwendung auf partielle Differentialgleichungen. Z. angew. Math. Mech. (im Druck)

/2/ Berg, L.: Operatorenrechnung I. Algebraische Methoden. Berlin 1972

/3/ Chintschin, A. J.: Die Elemente der Zahlentheorie. In: Alexandroff, P. S., Markuschewitsch, A. I., Chintschin, A. J. (Herausg.): Enzyklopädie der Elementarmathematik I, Arithmetik (Übers. a. d. Russ.). Berlin 1977

/4/ Hörmander, L.: Linear Partial Differential Operators. Berlin - Heidelberg - New York 1976

eingegangen: 18. 07. 1980

Anschrift des Verfassers:

Prof. Dr. Lothar Berg
Wilhelm-Pieck-Universität Rostock
Sektion Mathematik
Universitätsplatz 1
DDR-2500 Rostock

Rostock. Math. Kolloq. 16, 81 - 86 (1981)

Mirko Polonijo

Associativity and n-ary cyclic bisymmetry

Studying the connection between associativity and cyclic bisymmetry H. Kiesewetter in his paper /1/ proved the following propositions:

Proposition 1: From the assumptions

1. for all $x_1, x_2 \in X$, there is $\varphi(x_1,x_2) \in X$,
2. for all $x_1, x_2 \in X$ the relation $\varphi(x_2,\varphi(x_1,x_2)) = x_1$ is valid (cyclicity),
3. for all $x_1, x_2, x_3, x_4 \in X$ $\varphi(\varphi(x_1,x_2), \varphi(x_3,x_4))$ is invariant for any permutation of the elements x_1, x_2, x_3, x_4 (bisymmetry)

it follows that

a. for an arbitraty fixed element e of X the operation

$$\psi(x_1,x_2) := \varphi(e,\varphi(x_1,x_2)) \tag{1}$$

is an abelian group operation on X,

b. e is the identity,

c. $x^{-1} := \varphi(\varphi(e,e),x)$ is the inverse of the element $x \in X$.

Proposition 2: From the assumptions

a. ψ is an abelian group operation on X,

b. e is the identity,

c. x^{-1} is the inverse of an element $x \in X$

it follows that

1. for all $x_1, x_2 \in X$, there is

$$\varphi(x_1,x_2) := \psi(d,\psi(x_1,x_2)^{-1}) \in X, \tag{2}$$

where $d \in X$ is an arbitrary fixed element,

2. for all x_1, $x_2 \in X$ the relation $\varphi(x_2, \varphi(x_1,x_2)) = x_1$ is valid,

3. for all x_1, x_2, x_3, $x_4 \in X$, $\varphi(\varphi(x_1,x_2), \varphi(x_3,x_4))$ is invariant for any permutation of the elements x_1, x_2, x_3, x_4.

As Kiesewetter underlined, it means that to any cyclic bisymmetric operation corresponds a family of abelian group operations (e is a parameter of the family) and conversely to any abelian group operation corresponds a family of cyclic bisymmetric operations (d is a parameter of the family).
The purpose of the present note is to generalize Kiesewetter's results for the case of an n-ary operation, i. e. the following theorems will be proved:

<u>Theorem 1:</u> From the assumptions

1. for all x_1, x_2, ..., $x_n \in X$ there is $\varphi(x_1,x_2,\ldots,x_n) \in X$,

2. for all x_1, x_2, ..., $x_n \in X$ the relation
$\varphi(x_2,x_3,\ldots,x_n, \varphi(x_1,x_2,\ldots,x_n)) = x_1$ is valid
(i. e. φ is cyclic),

3. for all $x_{ij} \in X$, $i,j \in \{1,2,\ldots,n\}$, the expression
$$\varphi(\varphi(x_{11},x_{12},\ldots,x_{1n}), \varphi(x_{21},x_{22},\ldots,x_{2n}), \ldots, \varphi(x_{n1},x_{n2},\ldots,x_{nn}))$$
is invariant for any permutation of elements x_{ij}, $i,j \in \{1,2,\ldots,n\}$ (i. e. φ is bisymmetric)

it follows that

a. for an arbitratry fixed $e \in X$, the binary operation

$$\psi(x_1,x_2) := \varphi(e,e,\ldots,e, \varphi(x_1,x_2,e,\ldots,e)) \qquad (3)$$

is an abelian group operation on X,

b. e is the identity,

c. $x^{-1} := \varphi(\varphi(e,e,\ldots,e), e,\ldots,e,x)$ is the inverse of the element $x \in X$.

Theorem 2: From the assumptions

a. ψ is an abelian group operation on X,

b. e is the identity,

c. x^{-1} is the inverse of the element $x \in X$
it follows that

1. for all $x_1, x_2, \ldots, x_n \in X$ there is

$$\varphi(x_1,x_2,\ldots,x_n) := \psi(d, \psi(\psi(\psi(\ldots\psi(\psi(x_1,x_2),x_3), \ldots),x_{n-1}),x_n)^{-1}) \in X, \tag{4}$$

where $d \in X$ is an arbitrary fixed element,

2. for all $x_1,x_2,\ldots,x_n \in X$ the relation
$\varphi(x_2,x_3,\ldots,x_n, \varphi(x_1,x_2,\ldots,x_n)) = x_1$ is valid,

3. for all $x_{ij} \in X$, $i,j \in \{1,2,\ldots,n\}$, the expression

$$\varphi(\varphi(x_{11},x_{12},\ldots,x_{1n}), \varphi(x_{21},x_{22},\ldots,x_{2n}), \ldots, \varphi(x_{n1},x_{n2},\ldots,x_{nn}))$$

is invariant for any permutation of elements
$x_{ij} \in X$, $i,j \in \{1,2,\ldots,n\}$.

By the group computation, it is easy to show that the statement of theorem 2 is true and by the way we have $\varphi(e,e,\ldots,e) = d$.

Now, we are going to give the

Proof of Theorem 1: Let e be an arbitrary fixed element of X.
Since $y_i = \varphi(x_i,e,\ldots,e)$ for $x_i \in X$, $i \in \{1,2,\ldots,n\}$,
implies (by cyclicity) $x_i = \varphi(e,e,\ldots,y_i)$,
it follows that

$$\varphi(x_1,x_2,\ldots,x_n) = \varphi(\varphi(e,\ldots,e,y_1), \varphi(e,\ldots,e,y_2), \ldots, \varphi(e,\ldots,e,y_n))$$

is invariant for any permutation of $x_1, x_2, \ldots, x_n$
(because of bisymmetry), i. e. φ is symmetric.

In particular we get

$$\psi(x_1,x_2) = \psi(x_2,x_1)$$

for all x_1, $x_2 \in X$.

Let us put

$$d = \varphi(e,e,\dots,e).$$

Then the equation

$$\varphi(e,e,\dots,e,d) = e$$

holds and

$$\varphi(e,e,\dots,e,\varphi(x_1,x_2,\dots,x_n),\ \varphi(x_{n+1},x_{n+2},\dots,x_{2n}))$$

$$= \varphi(\varphi(e,\dots,e,d),\dots,\varphi(e,\dots e,d),\varphi(x_1,\dots,x_n),\ \varphi(x_{n+1},\dots,x_{2n}))$$

is invariant for any permutation of x_1, x_2, ..., x_{2n}.

Therefore, for all x, y, $z \in X$ the relation

$$\varphi(e,\dots,e,\varphi(x,y,e,\dots,e),\ \varphi(z,e,\dots,e))$$
$$=\varphi(e,\dots,e,\varphi(y,z,e,\dots,e),\ \varphi(x,e,\dots,e))$$

is valid. Putting

$$\varphi(x,e,\dots,e) = u,\quad \varphi(z,e,\dots,e) = v,$$

the last equality can be written in the form

$$\varphi(e,\dots,e,\varphi(x,y,e,\dots,e),v) = \varphi(e,\dots,e,\varphi(y,z,e,\dots,e),u).$$

Hence

$$\varphi(e,\dots,e,\varphi(e,\dots,e,\varphi(x,y,e,\dots,e),v),\ \varphi(e,\dots,e))$$
$$=\varphi(e,\dots,e,\varphi(e,\dots,e,\varphi(y,z,e,\dots,e),u),\ \varphi(e,\dots,e)),$$

i. e.

$$\varphi(e,\dots,e,\varphi(e,\dots,e,e,\varphi(x,y,e,\dots,e)),\ \varphi(e,\dots,e,v))$$
$$=\varphi(e,\dots,e,\varphi(e,\dots,e,e,\varphi(y,z,e,\dots,e)),\ \varphi(e,\dots,e,u))$$

and we obtain

$$\varphi(e,\dots,e,\psi(x,y),z) = \varphi(e,\dots,e,\psi(y,z),x)$$

from which it follows that

$$\psi(\psi(x,y),z) = \psi(\psi(y,z),x),$$

and we have proved the associativity of ψ, because ψ is a commutative operation.

Finally we obtain

$$\begin{aligned}\psi(e,x) &= \varphi(e,e,\dots,e,\varphi(e,x,\dots,e))\\ &= \varphi(e,e,\dots,e,\varphi(x,e,\dots,e)) = x\end{aligned}$$

and

$$\begin{aligned}\psi(x^{-1},x) &= \varphi(e,\dots,e,e,\varphi(x^{-1},x,e,\dots,e))\\ &= \varphi(e,\dots,e,\varphi(d,e,\dots,e),\ \varphi(x^{-1},x,e,\dots,e))\\ &= \varphi(e,\dots,e,\varphi(d,e,\dots,x),\ \varphi(x^{-1},e,e,\dots,e))\\ &= \varphi(e,\dots,e,x^{-1},\varphi(x^{-1},e,\dots,e)) = e,\end{aligned}$$

this completes the proof of the theorem.

In his article Kiesewetter noted the reciprocity between the formulas (1) and (2) and the same is the case in the generalized situation as will be shown in the following theorems.

Theorem 3: If φ is an n-ary cyclic-bisymmetric operation on X and there exists an abelian group operation ψ on X with the identity e and $d = \varphi(e,e,\dots,e)$ such that (4) is valid, then ψ has the representation (3).

Proof: Because of (4) we have

$$\varphi(e,e,\dots,e,x) = \psi(d,x^{-1})$$

and

$$\varphi(x_1,x_2,e,\dots,e) = \psi(d,\psi(x_1,x_2)^{-1}) = \psi(d,\psi(x_1^{-1},x_2^{-1}))$$

for all $x, x_1, x_2 \in X$. Therefore

$$\begin{aligned}&\varphi(e,e,\dots,e,\varphi(x_1,x_2,e,\dots,e))\\ &= \varphi(e,e,\dots,e,\psi(d,\psi(x_1^{-1},x_2^{-1})))\\ &= \psi(d,\psi(d,\psi(x_1^{-1},x_2^{-1}))^{-1}) = \psi(x_1,x_2),\end{aligned}$$

which was to be proved.

Theorem 4: If ψ is an abelian group operation on X with the identity e and if there exists an n-ary cyclic-bisymmetric operation φ on X such that (3) is valid, then φ has the representation (4), where $d = \varphi(e,e,\dots,e)$.

Proof: Recall that from the proof of Theorem 1 we know that $\varphi(x_1,x_2,\dots,x_n)$ and

$\varphi(e,e,\dots,e,\varphi(x_1,x_2,\dots,x_n),\ \varphi(x_{n+1},x_{n+2},\dots,x_{2n}))$ are

invariant for any permutation of their elements. Now, for $x \in X$, let be $y = \varphi(d,e,\dots,e,x)$ and by cyclicity of φ we get with similar arguments as above $x^{-1} = \varphi(d,e,\dots,e,x)$. Further, for all $x_1,x_2,\dots,x_{k+1} \in X$, $1 \le k \le n-1$, we have

$$\psi(\varphi(x_1,\dots,x_k,e,\dots,e),x_{k+1}^{-1}) = \varphi(x_1,\dots,x_{k+1},e,\dots,e)$$

which gives, taking in mind

$$\psi(d,x^{-1}) = \varphi(e,\dots,e,\varphi(d,\varphi(d,e,\dots,e,x),e,\dots,e)) = \varphi(x,e,\dots,e),$$

the result

$$\psi(d,\psi(\psi(\dots\psi(\psi(x_1,x_2),x_3),\dots,x_{n-1}),x_n)^{-1})$$
$$= \psi(\varphi(x_1,\dots,x_{n-1},e)x_n^{-1}) = \varphi(x_1,\dots,x_n),$$

and the theorem is proved.

Reference

/1/ Kiesewetter, H.: Ein Zusammenhang zwischen assoziativen und zyklisch-bisymmetrischen Verknüpfungen. Aequationes Math. 1, 83 - 88 (1970)

received: June 25, 1980

Author's address:

Prof. Mirko Polonijo
Department of Mathematics
University of Zagreb
Yu-41001 Zagreb, p.p. 187
Marulićev trg. 19
Yugoslavia

Rostock. Math. Kolloq. 16, 87 - 93 (1981)

Helmut Thielcke

Rechentechnische Aspekte bei der numerischen Auswertung von mehrfachen Integralen

Bei der numerischen Behandlung von mehrfachen Integralen ergeben sich sowohl Rechenzeit- als auch Genauigkeitsprobleme, und zwar je nach Vielfachheit der auftretenden Integrale. Im allgemeinen kann man daher nicht mit allgemeingültigen, von der Form des Integranden und der Integrationsgrenzen unabhängigen Verfahren arbeiten. Wie im Falle einfacher Integrale so erst recht bei mehrfachen Integralen empfiehlt es sich, durch Anwendung Gaußscher Quadraturformeln, die in Abhängigkeit von der Form des Integranden und der Integrationsgrenzen speziell ausgewählt wurden (und zwar für jede Integrationsvariable gesondert), mit einem Minimum an Stützstellen eine gewünschte Genauigkeit der numerischen Quadratur zu erreichen.
Im folgenden wird hierzu ein Algorithmus angegeben. Anschließend werden einige numerische Beispiele betrachtet, um unterschiedliche Verfahren miteinander zu vergleichen. Der Vergleich erscheint dabei möglich, obwohl aus Gründen des Rechenaufwandes nur einfachste Integranden und mit einer Ausnahme nur zweifache Integrale betrachtet wurden. Alle Beispiele wurden mit dem programmierbaren Taschenrechner SR 52 von TI auf der Grundlage des angegebenen Algorithmus mit der Vereinfachung $u(i,\underline{s}) = 0$ gerechnet.

Beschreibung des Algorithmus

Als Basis für den Algorithmus dient eine Datenstruktur. In ihr sind Quadraturvektoren Q_i von im allgemeinen unterschiedlicher Länge enthalten, deren Komponenten folgende Bedeutung haben:

m_i - Anzahl der Stützstellen einer Quadraturformel,

$x_1^i \dots x_{m_i}^i$ - Stützstellen und

$w_1^i \ldots w_{m_i}^i$ - Gewichte der Quadraturformel sowie

(a_i, b_i) - Quadraturintervall, auf das Bezug genommen wird.

Zur Berechnung eines gegebenen n-fachen Integrals ist für jede Integrationsvariable s_i die Auswahl eines zugeordneten Quadraturvektors Q_i vorzunehmen. Ferner sind die Grenzen eines Bezugsintervalls $(p(i,\underline{s}), q(i,\underline{s}))$, das im Hinblick auf die Lage des Quadraturintervalls (a_i, b_i) normiert ist (im allgemeinen wird $p = a_i$ und $q = b_i$ sinnvoll sein, man vergleiche aber die Beispiele (5) und (7)), und die Integrationsgrenzen $u(i,\underline{s})$, $v(i,\underline{s})$ als Funktionen der Argumente $i, s_1, \ldots, s_{i-1}$ zu definieren.

Der nun angegebene Algorithmus ermöglicht auf einfache Weise die flexible Verknüpfung der unterschiedlichen Quadraturvektoren Q_i zu einer der gegebenen Form des Integranden $f(\underline{s})$ und der Grenzen $u(i,\underline{s})$, $v(i,\underline{s})$ angepaßten Berechnung eines n-fachen Integrals:

```
...
     summe := 0;
     for j := 1 step 1 until n do
     begin s_j := 0 ; k_j := 1 end;
marke1 : j := n ;
marke2 : faktor := 1 ;
     for i := 1 step 1 until n do
     if x^i_{k_i} < p(i,s) v x^i_{k_i} > q(i,s) then goto marke3 else
     begin nenner := p(i,s) - q(i,s) ;
           alpha := ( u(i,s) - v(i,s) ) / nenner ;
           beta := ( p(i,s)v(i,s) - q(i,s)u(i,s) ) / nenner;
           s_i := alpha · x^i_{k_i} + beta ;
           faktor := faktor · alpha · w^i_{k_i}
     end ;
     summe := summe + faktor · f(s) ;
marke3 : k_j := k_j + 1 ; if m_j ≥ k_j then goto marke2;
marke4 : k_j := 1 ; j := j - 1 ; if j > 0 then
     begin k_j := k_j + 1 ; if m_j ≥ k_j then goto marke1;
           goto marke4
     end ;
...
```

Beispiele

In den Beispielen (1) bis (6) werden die Integranden

$$f_1(x,y) = \begin{Bmatrix} (1-x^2-y^2)^{1/2}, & 1 > x^2 + y^2 \\ 0 & , \text{sonst} \end{Bmatrix},$$

$$f_2(x,y) = 1 - x^2 - y^2,$$

$$f_3(x,y) = \begin{Bmatrix} 1, & 1 > x^2 + y^2 \\ 0, & \text{sonst} \end{Bmatrix}$$ sowie $(1-x^2)^{1/2}$ und Potenzen

von xy betrachtet, als Integrationsgrenzen treten mit wenigen, gesondert gekennzeichneten Ausnahmen 0 und 1 auf.

(1) Nach der "Methode der optimalen Koeffizienten" (vgl. /1/, S. 261 ff), einem einfachen Verfahren ohne Berücksichtigung der Form des Integranden, das die Stützstellen $x_k = \frac{k}{u_m}$,

$y_k = x_k u_{m-1} - [x_k u_{m-1}]$ mit Hilfe der Fibonacci-Folge $\{u_o = u_1 = 1,\ u_{m+1} = u_m + u_{m-1}\}$ bestimmt, ergeben sich folgende relative Fehler:

Anzahl der Punkte / Integrand	5	8	13	21	34	55	89	144
$f_1(x,y)$	-.1917	-.0938	-.0572	-.0332	-.0248	-.0045	-.0054	-.0039
$f_2(x,y)$	-.0400	-.0156	-.0059	-.0023	-.0009	-.0003	-.0001	-.00005
$f_3(x,y)$	.2732	.1141	.0774	.0307	.0111	.0186	.0157	.0080
$(1-x^2)^{1/2}$	-.1606	-.0961	-.0569	-.0342	-.0206	-.0125	-.0076	-.0046
$(xy)^{1/2}$	-.1505	-.0987	-.0516	-.0349	-.0183	-.0127	-.0067	-.0047
$(xy)^{1/4}$	-.1547	-.0929	-.0517	-.0313	-.0175	-.0108	-.0060	-.0038
$(xy)^{1/8}$	-.1708	-.1026	-.0597	-.0356	-.0208	-.0125	-.0073	-.0044
$(xy)^{3/2}$	-.2778	-.2327	-.1139	-.0919	-.0449	-.0356	-.0174	-.0137
xy	-.2000	-.1563	-.0769	-.0597	-.0294	-.0228	-.0112	-.0087
$(xy)^5$	-.7922	-.7079	-.4666	-.3637	-.2202	-.1592	-.0929	-.0647
$(xy)^{30}$	-1.000	-1.000	-.9998	-.9987	-.9836	-.9390	-.8504	-.7205 .

(2) Mit äquidistanten Stützstellen $x_k = y_k = (k-.5)/N$ sind die relativen Fehler von ähnlicher Größenordnung:

Anzahl der Punkte / Integrand	2^2	3^2	4^2	5^2	6^2	7^2	8^2	9^2
$f_1(x,y)$	.0314	.0407	.0260	.0082	.0047	.0026	.0079	.0049
$f_2(x,y)$	.1250	.0556	.0313	.0200	.0139	.0102	.0078	.0062
$f_3(x,y)$	-.0451	.1318	.0345	.0186	-.0097	.0134	.0345	.0060
$(1-x^2)^{1/2}$	.0375	.0206	.0135	.0097	.0074	.0059	.0048	.0040
$(xy)^{1/2}$	.0496	.0284	.0190	.0139	.0107	.0086	.0071	.0060
$(xy)^{1/4}$	.0477	.0296	.0210	.0161	.0129	.0107	.0091	.0079
$(xy)^{1/8}$	.0315	.0205	.0150	.0118	.0096	.0081	.0070	.0062
$(xy)^{3/2}$	-.0627	-.0292	-.0169	-.0110	-.0077	-.0057	-.0044	-.0035
xy	0	0	0	0	0	0	0	0
$(xy)^5$	-.4890	-.2492	-.1470	-.0962	-.0676	-.0500	-.0385	-.0305
$(xy)^{30}$	-1.000	-.9981	-.9801	-.9309	-.8551	-.7671	-.6789	-.5971 .

(3) Die Gaußschen Quadraturformeln (Gewichtsfunktion gleich Eins, Bezugsintervall gleich dem Quadraturintervall (-1,1)) erbringen deutlich genauere Ergebnisse (angegeben wird der relative Fehler):

	2^2	3^2	4^2	5^2	6^2	7^2	8^2	9^2
$f_1(x,y)$	.0136	.0046	.0021	.0011	.00069	.00045	.00031	.00022
$f_2(x,y)$	0	0	0	0	0	0	0	0
$f_3(x,y)$	.0136	.0046	.0021	.0011	.00069	.00045	.00031	.00022
$(1-x^2)^{1/2}$	.0136	.0046	.0021	.0011	.00069	.00045	.00031	.00022
$(xy)^{1/2}$	.0218	.0076	.0035	.0019	.00140	.00074	.00051	.00036
$(xy)^{1/4}$	.0256	.0107	.0056	.0034	.0022	.0016	.0011	.00086
$(xy)^{1/8}$	.0185	.0085	.0078	.0030	.0021	.0015	.0011	.00088
$(xy)^{3/2}$	-.0061	-.0009	-.0003	-.0001	-.00004	-.00002	-.00001	-.000005
xy	0	0	0	0	0	0	0	0
$(xy)^5$	-.1597	0	0	0	0	0	0	0
$(xy)^{30}$	-.9998	-.9432	-.6124	-.2401	-.0611	-.0107	-.0013	-.00011.

Im Falle der Integranden $f_1(x,y)$ und $f_3(x,y)$ wurden die Integrale $\int_0^1 \int_0^{\sqrt{1-x^2}} f_i(x,y)\, dy\, dx$ ausgewertet.

(4) Extrapoliert man die jeweils letzten vier Näherungswerte nach der Formel $a + b/n^2 + c/n^3 + d/n^4$ (das lineare Glied kann entfallen, da die Anzahl der Stützpunkte quadratisch wächst), so berechnet sich der verbesserte Integralwert zu

$$I = \det(y_k,\ 1/k^2,\ 1/k^3,\ 1/k^4)/\det(1,\ 1/k^2,\ 1/k^3,\ 1/k^4),$$

$k = 6, \ldots, 9$, und gewinnt damit erheblich an Genauigkeit:

Funktion	extrapolierter Wert	relativer Fehler
$f_1(x,y)$	.5235976	-.0000023
$(1-x^2)^{1/2}$	.7853964	-.0000023
$(xy)^{1/2}$	.4444428	-.0000037
$(xy)^{1/4}$	.6399812	-.0000294
$(xy)^{1/8}$	.7901056	-.0000226
$(xy)^{3/2}$	.1600002	.0000015 .

(5) Bei der Berechnung des Integrals $\int_0^1 \int_0^{\sqrt{1-x^2}} f_1(x,y)\, dy\, dx$

mit Hilfe der (einseitigen) Gauß-Tschebyscheff-Quadraturformeln (Gewichtsfunktion gleich $1/\sqrt{1-x^2}$, Bezugsintervall (0, 1), Quadraturintervall (-1, 1)) wird im Vergleich zum Beispiel (3) der Vorzug einer Anpassung der Quadraturformel an den Integranden und die Integrationsgrenzen deutlich:

Stützstellen	Näherungswert	relativer Fehler
2^2	.521006	-.00495
3^2	.523135	-.00089
4^2	.523457	-.00027
5^2	.523542	-.00011
6^2	.523571	-.000052
7^2	.523584	-.000028
8^2	.523590	-.000016
9^2	.523593	-.000010

Auch in diesem Fall kann mit Vorteil extrapoliert werden. Man erhält den Näherungswert .5235989 mit einem relativen Fehler von .00000028.

(6) Einen ähnlichen Effekt erzielt man z. B. für die obigen Wurzelfunktionen durch Auswahl von entsprechend angepaßten Stützstellen und Gewichten (Anzahl = 4, Intervall (0,1)):

i	x^i	w^i	Funktion	relativer Fehler
1	.0195328197	.0623619419	$(xy)^{1/2}$	0
2	.1733969280	.2596950951	$(xy)^{3/2}$	0
3	.5229560269	.4069291363	$(xy)^{1/4}$	.00026
4	.8890524971	.2710138267	$(xy)^{1/8}$	.00038 .

Die hier gegebene Genauigkeit wird mit Hilfe einfacher Gaußscher Quadraturformeln nicht einmal mit 9^2 Punkten, also mehr als 5-facher Stützstellenanzahl, erreicht (vgl. Beispiel (3)).

(7) Abschließend betrachten wir das vierfache Integral

$$\frac{\pi^2}{60} = \int_0^1 \int_0^{\sqrt{1-s_1^2}} \int_0^{\sqrt{1-s_1^2-s_2^2}} \int_0^{\sqrt{1-s_1^2-s_2^2-s_3^2}} \sqrt{1-s_1^2-s_2^2-s_3^2-s_4^2}\, ds_4 ds_3 ds_2 ds_1$$

und berechnen seinen Wert wie in Beispiel (5) mittels einseitiger Gauß-Tschebyscheff-Quadratur:

Anzahl	Näherungswert	relativer Fehler
2^4	.164219	-.00167
3^4	.164380	-.00069
4^4	.164454	-.00024
5^4	.164477	-.00010 .

Eine Extrapolation nach dem Ansatz $a + b/n^4 + c/n^5 + d/n^6$ (Glieder bis zur dritten Ordnung können weggelassen werden, da die Anzahl der Stützstellen mit der vierten Ordnung wächst) erbringt den Wert .1644922 mit einem relativen Fehler von -.0000076.

Literatur

/1/ Beresin, I. S., und Shidkow, N. P.: Numerische Methoden I. Berlin 1970

eingegangen: 20. 08. 1980

Anschrift des Verfassers:

Dr. rer. nat. Helmut Thielcke
Kombinat Schiffbau Rostock
Stammbetrieb
Wismarsche Straße 6/7
DDR-2500 Rostock

Rostock. Math. Kolloq. 16, 95 - 101 (1981)

Carsten Stade

Ein Verfahren mit quadratischer Konvergenzgeschwindigkeit zur simultanen Berechnung zweier Nullstellen beliebiger Ordnung[1]

In der Arbeit /1/ von L. Berg wird ein Iterationsverfahren vorgestellt, das jeweils zwei Nullstellen einer differenzierbaren Funktion durch Approximation mittels Parabeln gleichzeitig berechnet, wobei die Nullstellen der Parabeln Näherungen für die der Funktion sind. Dabei werden Nullstellen aber nur einfach abgespalten. Im folgenden wird ein modifiziertes Verfahren betrachtet, bei dem Nullstellen höherer Ordnung mehrfach abgespalten werden.

Die Hauptergebnisse der vorliegenden Arbeit lauten:

Das Verfahren aus /1/ konvergiert für einfache Nullstellen quadratisch und für mehrfache linear. Die Konvergenzfaktoren entsprechen im zweiten Fall denen des Newtonverfahrens. Das modifizierte Verfahren konvergiert quadratisch, wenn man die Wurzeln entsprechend ihrer Vielfachheit abspaltet.

Abschließend werden Vorschläge für eine günstige rechentechnische Realisierung angegeben.

Bestimmung der Konvergenzgeschwindigkeit

Es sei $f(x)$ eine gegebene, dreimal stetig differenzierbare Funktion, x_1 und x_2 seien verschiedene Näherungswerte für zwei verschiedene Nullstellen von $f(x)$. Mit den Abkürzungen $f_i = f(x_i)$ und $f_i' = f'(x_i)$ für $i = 1, 2$ ergibt sich aus dem Verfahren von /1/ das erwähnte modifizierte Verfahren, indem man f_i durch $m_i f_i$ ersetzt, aber f_i' unverändert läßt, so daß die Gleichungen

1 Geänderte Fassung eines Teils der Jahresarbeit 1979/80

$$(x_i f_i' - m_i f_i)p + f_i' q = 2x_i m_i f_i - x_i^2 f_i' \tag{1}$$

entstehen, die für $m_1 = m_2 = 1$ wieder in das ursprüngliche Verfahren übergehen. Entsprechend der Vorgehensweise in /1/ hat man aus den beiden Gleichungen (1) die Werte p und q zu berechnen, wobei die Lösungen der quadratischen Gleichung

$$x^2 + px + q = 0$$

dann die verbesserten Näherungen für die gesuchten Nullstellen von $f(x)$ sind. Dies bedeutet, daß die beiden Nullstellen von $f(x)$ als Lösung der Fixpunktgleichung

$$y = g(y) \tag{2}$$

mit

$$y = (x_1, x_2)^T, \; g(y) = (g_1(y), g_2(y))^T$$

und

$$g_1(y) = \tfrac{1}{2}(-p + \sqrt{p^2-4q}), \; g_2(y) = \tfrac{1}{2}(-p - \sqrt{p^2-4q}) \tag{3}$$

iterativ bestimmt werden. Dabei ergeben sich p und q aus (1) nach der Cramerschen Regel zu

$$p = D_p/D, \; q = D_q/D$$

mit

$$\begin{aligned} D &= (x_1-x_2)f_1'f_2' + m_2 f_2 f_1' - m_1 f_1 f_2', \\ D_p &= (x_2^2-x_1^2)f_1'f_2' + 2x_1 m_1 f_1 f_2' - 2x_2 m_2 f_2 f_1', \\ D_q &= (x_1-x_2)x_1 x_2 f_1' f_2' + 2(x_1-x_2)m_1 f_1 m_2 f_2 \\ &\quad + 2x_1 x_2(m_2 f_2 f_1' - m_1 f_1 f_2') + m_1 f_1 f_2' x_2^2 - m_2 f_2 f_1' x_1^2. \end{aligned} \tag{4}$$

Zur Untersuchung der Konvergenzgeschwindigkeit bezeichnen wir den Fixpunkt von (2) mit

$$y^* = (x_1^*, x_2^*)^T$$

und ziehen die Taylorsche Formel

$$g(y) = g(y^*) + g'(y^*)(y-y^*) + O(y-y^*)^2 \tag{5}$$

heran, wobei $g'(y) = (\partial g_i(y)/\partial x_j)$ $(i,j=1,2)$ die Jacobi-Matrix bezeichnet und das Restglied für $y \to y^*$ im Sinne von

$$\|O(y-y^*)^2\| \le M \,\|y-y^*\|^2$$

mit der euklidischen Norm und einer passenden Konstanten M zu verstehen ist. Aus (3) ergibt sich

$$\frac{\partial g_{1/2}}{\partial x_i} = -\frac{1}{2}\left(\frac{\partial p}{\partial x_i} \mp \frac{p\partial p/\partial x_i - 2\partial q/\partial x_i}{(p^2-4q)^{1/2}}\right), \quad i = 1, 2, \tag{6}$$

wobei für g_1 das obere und für g_2 das untere Vorzeichen gilt. Mit den Abkürzungen $F_i = m_i f_i/f_i'$, $i = 1, 2$, erhält man aus den Gleichungen (4)

$$p = \frac{D_p}{D} = \frac{(x_2^2-x_1^2) + 2x_1F_1 - 2x_2F_2}{x_1-x_2+F_2-F_1} = -(x_1+x_2) + \frac{(x_1-x_2)(F_1+F_2)}{x_1-x_2+F_2-F_1} \tag{7}$$

und

$$q = \frac{D_q}{D} = x_1x_2 + \frac{2(x_1-x_2)F_1F_2 - x_1x_2(F_1-F_2) + F_1x_2^2 - F_2x_1^2}{x_1-x_2+F_2-F_1}. \tag{8}$$

Die Nullstellen x_i von $f(x)$ mögen die Vielfachheiten k_i besitzen, so daß für $i = 1, 2$

$$f_i = (x_i-x_i^*)^{k_i}\, r_i(x_i)$$

mit $r_i(x_i^*) \neq 0$ gilt und die logarithmische Ableitung

$$f_i'/f_i = k_i(x_i-x_i^*)^{-1} + r_i'(x_i)/r_i(x_i)$$

lautet. Mit den weiteren Abkürzungen $m_i/k_i = a_i$ ergibt sich

$$F_i = m_i f_i/f_i' = a_i(x_i-x_i^*) + O(x_i-x_i^*)^2, \quad i = 1, 2.$$

Wenn wir dieses Ergebnis in (7) und (8) einsetzen, $x_1 - x_2$ kürzen und die Nenner durch Potenzreihenentwicklung beseitigen, so erhalten wir

$$p = -(x_1^*+x_2^*) + (a_1-1)(x_1-x_1^*) + (a_2-1)(x_2-x_2^*) + O(y-y^*)^2, \quad (9)$$

$$q = x_1x_2 - a_1x_2(x_1-x_1^*) - a_2x_1(x_2-x_2^*) + O(y-y^*)^2. \quad (10)$$

Offensichtlich gilt $\lim\limits_{y\to y^*} p = -(x_1^*+x_2^*)$, $\lim\limits_{y\to y^*} q = x_1^*x_2^*$ und folglich bei geeigneter Wahl des Vorzeichens

$$\lim_{y\to y^*}\sqrt{p^2 - 4q} = x_1^* - x_2^* \neq 0.$$

Nebenbei erkennen wir hieraus wegen (3), daß

$$\lim_{y\to y^*} g(y) = y^*$$

und y^* somit tatsächlich ein Fixpunkt der Gleichung (2) ist. Da (9) und (10) endliche Taylorentwicklungen sind, können wir dort die partiellen Ableitungen von p und q an der Stelle $y = y^*$ unmittelbar ablesen:

$$\partial p/\partial x_i\big|_{y=y^*} = a_i - 1, \; \partial q/\partial x_i\big|_{y=y^*} = (1-a_i)x_j^*, \quad i, j = 1,2; \; i\neq j.$$

Wegen (6) ergibt sich daraus

$$\partial g_i/\partial x_i\big|_{y=y^*} = 1 - a_i, \; \partial g_i/\partial x_j\big|_{y=y^*} = 0. \quad (11)$$

Folglich konvergiert das in /1/ vorgestellte Verfahren, bei dem $a_i = k_i^{-1}$ ist, in der Umgebung jeder einfachen Nullstelle quadratisch und in der Umgebung jeder mehrfachen Nullstelle linear mit dem Konvergenzfaktor $1 - k_i^{-1}$. Dieser Konvergenzfaktor tritt bekanntlich auch beim gewöhnlichen Newtonverfahren auf (vgl. etwa /2/).
Aus (5) und (11) ersieht man weiter, daß das modifizierte Verfahren mit $a_i = m_i/k_i$ quadratisch konvergiert, wenn $a_1 = a_2 = 1$

ist, also wenn die Nullstellen entsprechend ihren Vielfachheiten abgespalten werden. Es weist damit die gleichen Eigenschaften bez. der Konvergenzgeschwindigkeit auf wie das modifizierte Newtonverfahren (vgl. /2/).
Sind die Vielfachheiten der Nullstellen unbekannt, so kann man bekanntlich quadratische Konvergenz erreichen, indem man das ursprüngliche Verfahren für f/f' an Stelle von f benutzt, denn f/f' hat nur einfache Nullstellen.

Ein Spezialfall bei der rechentechnischen Realisierung

Bei Testrechnungen trat mehrfach der Fall auf, daß im Verlaufe der Rechnung eine Nullstelle im Rahmen der Rechengenauigkeit bereits exakt berechnet wurde, aber die andere wegen eines schlechteren Konvergenzfaktors oder eines schlechteren Startwertes noch nicht erreicht war. Da die erreichte Genauigkeit bei der weiteren Rechnung durch Rundungsfehler wieder verlorenging, wurde von zwei aufeinanderfolgenden Näherungen für dieselbe Nullstelle diejenige mit dem betragsmäßig kleineren Funktionswert dann beim nächsten Iterationsschritt verwendet, sofern die Iteration noch nicht abgebrochen werden konnte.
Beispielsweise ergaben sich bei dem Polynom

$$f(x) = \sum_{v=0}^{5} (5x)^v/v!$$

auf dem Rechner ES 1040 bei vier Iterationsschritten die Werte:

Schrittnummer	$Re(x_1)$	$Im(x_1)$	$f(x_1)$
0	-0.5	0.0	1.65
1	-0.44695181	0.00249535	$2.36\ 10^{-2}$
2	-0.43635452	0.00024286	$6.91\ 10^{-4}$
3	-0.43612152	0.00000018	$3.70\ 10^{-7}$
4	-0.43612128	-0.00000018	$1.65\ 10^{-6}$

Schrittnummer	$Re(x_2)$	$Im(x_2)$	$f(x_2)$
0	0.0	-0.6	$5.30\ 10^{-1}$
1	0.06198123	-0.62390041	$1.85\ 10^{-1}$
2	0.04823351	-0.62503147	$8.74\ 10^{-3}$
3	0.04795930	-0.62566710	$2.71\ 10^{-5}$
4	0.04796115	-0.62566662	$7.53\ 10^{-6}$

Abschließend soll dieser Fall noch mit dem Newtonverfahren verglichen werden. Es sei $x_1 = x_1^*$ eine bereits exakt erreichte Nullstelle. Für p und q ergibt sich dann aus (7) und (8)

$$p = -(x_1^* + x_2) + F_2(x_1^* - x_2)(F_2 + x_1^* - x_2)^{-1},$$

$$q = x_1^* x_2 - x_1^* F_2(x_1^* - x_2)(F_2 + x_1^* - x_2)^{-1},$$

und hieraus folgt bei geeigneter Wahl des Vorzeichens

$$\sqrt{p^2 - 4q} = x_1^* - x_2 + F_2(x_1^* - x_2)(F_2 + x_1^* - x_2)^{-1}.$$

Damit nehmen die Komponenten (3) von g(y) die Gestalt

$$g_1(y) = x_1^*,\ g_2(y) = x_2 + f_2 m_2(x_1^* - x_2)(f_2 m_2 - f_2'(x_1^* - x_2))^{-1} \quad (12)$$

an. Für eine einfache Nullstelle x_2^* lautet der zugehörige Konvergenzfaktor

$$\frac{1}{2}\,\partial^2 g_2/\partial x_2^2\Big|_{y=y^*} = \frac{1}{2}\,f''(x_2^*)/f'(x_2^*) + (x_1^* - x_2^*)^{-1},$$

während beim Newtonverfahren bekanntlich (vgl. /2/)

$$\frac{1}{2}\,\partial^2 g_2/\partial x_2^2\Big|_{y=y^*} = \frac{1}{2}\,f''(x_2^*)/f'(x_2^*)$$

ist. Das Iterationsverfahren mit (12) ist daher dem Newtonverfahren genau dann überlegen, wenn

1. $(x_1^* - x_2^*)f''(x_2^*)/f'(x_2^*) < 0$ und
2. $|(x_1^* - x_2^*)^{-1}| < |f''(x_2^*)/f'(x_2^*)|$

gilt (vgl. auch /1/).

Literatur

/1/ Berg, L.: Über die gleichzeitige Berechnung zweier Nullstellen. Computing 24, 87 - 91 (1980)

/2/ Kiesewetter, H., und Maeß, G.: Elementare Methoden der numerischen Mathematik. Berlin 1974

eingegangen: 16. 07. 1980

Anschrift des Verfassers:

stud. math. Carsten Stade
Wilhelm-Pieck-Universität Rostock
Sektion Mathematik
Universitätsplatz 1
DDR-2500 Rostock

Rostock. Math. Kolloq. 16, 103 - 128 (1981)

Dieter Schott

Endlich erzeugte Projektionsverfahren zur Lösung linearer Gleichungen im Hilbertraum

1. Einführung

Gegeben seien zwei Hilberträume X und Y. Der Raum der linearen und stetigen Operatoren von X in Y wird mit $L(X,Y)$ bezeichnet ($L(X) := L(X,X)$). Die Mengen $\mathbb{R}(T)$ und $\mathbb{N}(T)$ bedeuten Wertebereich bzw. Nullraum eines Operators $T \in L(X,Y)$. T^* ist der zu T adjungierte Operator. Der Ausdruck $T|M$ symbolisiert die Einschränkung von T auf die Menge $M \subseteq X$. Unter lin M verstehen wir die lineare Hülle, unter $\overline{M}$ die Abschließung von M. Der Einheitsoperator (auf X bzw. Y) wird durch I dargestellt. R ist die Menge der reellen Zahlen.

Zur Gewinnung von Lösungen bzw. verallgemeinerten Lösungen der linearen Gleichung

$$Ax = b \quad (A \in L(X,Y),\ b \in Y) \tag{1.1}$$

eignen sich Iterationsverfahren der Form

$$x_{n+1} = (I-D_nA)x_n + D_nb, \tag{1.2}$$

die nach Vorgabe einer Operatorfolge (D_n) mit $D_n \in L(Y,X)$ und eines Startelementes $x_0 \in X$ eindeutig bestimmt sind. Die Restiteration erfolgt dabei nach der Vorschrift

$$r_{n+1} = (I-AD_n)r_n \quad (r_n := b-Ax_n). \tag{1.3}$$

Für eine Reihe spezieller, vorwiegend zyklischer Verfahren dieser Gestalt gibt es bereits Untersuchungen zur Konvergenz der Folge (x_n) und zu den Lösungseigenschaften des Grenzelementes $x_\infty := \lim\limits_{n\to\infty} x_n$ (siehe z. B. /9/, /1/, /6/, /5/, /4/, /2/).

Hier wollen wir vor allem die von W. Peters und G. Maeß bez. des PSH- und SPA-Verfahrens erzielten Ergebnisse verallgemeinern (siehe /6/, /5/), indem wir

a) auf die dort gegebene spezielle Gestalt der Operatoren D_n verzichten,

b) uns von den endlichdimensionalen Räumen lösen,

c) statt der zyklischen Verfahren die allgemeinere Klasse der endlich erzeugten Verfahren zulassen (siehe Abschnitt 4).

2. Verallgemeinerte Lösungen von linearen Gleichungen

2.1 Der Begriff der verallgemeinerten Lösung

Wir definieren zunächst in Anlehnung an /4, S. 32/ den Begriff der verallgemeinerten Lösung einer Gleichung (1.1). Schließlich führen wir Bedingungen an, unter denen die Iterationsfolge (x_n) des Verfahrens (1.2) gegen eine solche verallgemeinerte Lösung strebt.

Gegeben seien zwei Hilberträume X, Y und eine Folge (Z_n) von weiteren Hilberträumen. Wir betrachten die Gleichung

$$Ax = b \qquad (A \in L(X,Y),\ b \in Y) \tag{2.1}$$

und das daraus von einer Operatorfolge (D_n) mit $D_n \in L(Y,Z_n)$ erzeugte Ersatz(gleichungs)system

$$D_n Ax = D_n b \qquad (n = 0,1,2,\ldots). \tag{2.2}$$

Sind

$$X(A,b) := \{x \in X \mid Ax = b\}, \tag{2.3}$$

$$\tilde{X}(A,b) := \bigcap_n X(D_nA, D_nb) = \{x \in X \mid D_n(b-Ax)=0\ (\forall n)\} \tag{2.4}$$

die Lösungsmengen von (2.1) bzw. (2.2), so gilt offenbar $\tilde{X}(A,b) \supseteq X(A,b)$. Daher ist (2.2) für alle $b \in Y$ konsistent zur Ausgangsgleichung (2.1). Insbesondere ergibt sich auch $\bigcap_n \mathbb{N}(D_nA) \supseteq \mathbb{N}(A)$.

Definition 2.1: a) Die Lösungen des Ersatzsystems (2.2), d. h. die Elemente von $\tilde{X}(A,b)$, heißen bezüglich (D_n) verallgemeinerte Lösungen der Gleichung (2.1).

b) Zwei Ersatzsysteme von (2.1) mit den erzeugenden Folgen (D_n) bzw. (D_n') heißen äquivalent, wenn sie die gleichen Lösungsmengen besitzen.

Der Begriff der verallgemeinerten Lösung wird noch abgeschwächt, wenn man zuläßt, daß das Ersatzsystem (2.2) für endlich viele n verletzt sein darf. In diesem Falle ist die Menge aller bezüglich (D_n) verallgemeinerten Lösungen von (2.1) durch

$$X^+(A,b) := \lim_{N\to\infty} \bigcap_{n=N}^{\infty} X(D_nA, D_nb)$$

gegeben. Wir wollen hier aber nicht näher darauf eingehen.

Lemma 2.2: Die Mengen $\tilde{X}(A,b)$ und $X(A,b)$ fallen unter folgenden Bedingungen zusammen:

a) Es gilt $\bigcap_n \mathbb{N}(D_n) = \{0\}$.

b) Es gilt $X(A,b) \neq \emptyset$ und $\bigcap_n \mathbb{N}(D_nA) = \mathbb{N}(A)$.

Beweis: Es ist jeweils nur $\tilde{X}(A,b) \subseteq X(A,b)$ zu zeigen. Dazu sei $x \in \tilde{X}(A,b)$, d. h. $D_n(b-Ax) = 0 \quad (\forall n)$.

a) Aus $\bigcap_n \mathbb{N}(D_n) = \{0\}$ erhält man sofort $b-Ax = 0$ bzw.

$x \in X(A,b)$.

b) Wegen $X(A,b) \neq \emptyset$ existiert ein $x^* \in X$ mit $Ax^* = b$. Daraus entsteht

$$D_n(b-Ax) = D_n(Ax^*-Ax) = D_nA(x^*-x) = 0 \quad (\forall n).$$

Unter Beachtung von $\bigcap_n \mathbb{N}(D_nA) = \mathbb{N}(A)$ gewinnt man schließlich

$$A(x^*-x) = Ax^* - Ax = b - Ax = 0.$$

Daher ist $x \in X(A,b)$.

Ist eine der Voraussetzungen des Lemmas 2.2 erfüllt, so ergibt jede bezüglich (D_n) verallgemeinerte Lösung von (2.1) auch eine Lösung von (2.1).

Lemma 2.3: Zwei Ersatzsysteme von (2.1) mit den erzeugenden Folgen (D_n) und (D_n') sind für beliebige $b \in Y$ äquivalent, wenn die Gleichung $\bigcap_n \mathbb{N}(D_n) = \bigcap_n \mathbb{N}(D_n')$ besteht.

Beweis: Aus $D_n(b-Ax) = 0 \; (\forall n)$ folgt wegen $b-Ax \in \bigcap_n \mathbb{N}(D_n) = \bigcap_n \mathbb{N}(D_n')$ auch $D_n'(b-Ax) = 0 \; (\forall n)$ und umgekehrt.

Sind (D_n) und (D'_n) die erzeugenden Folgen zweier äquivalenter Ersatzsysteme, so ist jede bezüglich (D_n) verallgemeinerte Lösung zugleich eine bezüglich (D'_n) verallgemeinerte Lösung.

2.2 Verallgemeinerte Lösungen und Iterationsverfahren

Wir setzen nun $Z_n = X$ $(\forall n)$ voraus. Dann ist das System (2.2) gleichbedeutend mit

$$x = (I-D_nA)x + D_nb = x + D_n(b-Ax) \quad (D_n \in L(Y,X)).$$

Dazu gehört in natürlicher Weise das Iterationsverfahren

$$x_{n+1} = T_nx_n + D_nb = x_n + D_n(b-Ax_n) \quad (T_n := I-D_nA) \tag{2.5}$$

mit einem beliebigen Startelement $x_o \in X$. Dieses Verfahren kann mit den Abkürzungen

$$P_n := T_nT_{n-1}\dots T_o, \quad B_n := \sum_{i=o}^{n} T_nT_{n-1}\dots T_{i+1}D_i \tag{2.6}$$

explizit auch in der Gestalt

$$x_{n+1} = P_nx_o + B_nb \tag{2.5'}$$

geschrieben werden. Die Reste $r_n := b - Ax_n$ transformieren sich dabei nach der Vorschrift

$$r_{n+1} = S_nr_n = Q_nr_o \quad (S_n := I-AD_n,\ Q_n := S_nS_{n-1}\dots S_o). \tag{2.7}$$

Die bezüglich (D_n) verallgemeinerten Lösungen von (2.1) sind gerade die Fixpunkte des Iterationsverfahrens (2.5), d. h. die Elemente x_o mit $x_n = x_o$ $(\forall n)$.

Wichtige Eigenschaften des Iterationsverfahrens beinhaltet

<u>Lemma 2.4:</u> Es gilt

<u>a</u>) $S_nA = AT_n$, $Q_nA = AP_n$, <u>b</u>) $B_n = \sum_{i=o}^{n} D_iQ_{i-1}$ $(Q_{-1}:=I)$,

<u>c</u>) $B_nA = I - P_n$, <u>d</u>) $AB_n = I - Q_n$.

<u>Beweis:</u> Die Gleichungen <u>a</u>) sind offensichtlich. Die Beziehung <u>b</u>) zeigt man leicht durch Induktion. Man findet sie in versteckter Form auch schon in /2, S. 377/. Die Beziehungen <u>c</u>) und <u>d</u>) sind in /4, S. 33/ und /2, S. 376 - 377/ abgeleitet.

Die im weiteren auftretenden Grenzwerte von Operatorfolgen sind entweder im Sinne der punktweisen (bzw. starken) oder der gleichmäßigen Operatortopologie zu verstehen.
Aus Lemma 2.4 ergibt sich dann unmittelbar

<u>Lemma 2.5:</u> Mit den Bezeichnungen $B_\infty := \lim\limits_{n\to\infty} B_n$, $P_\infty := \lim\limits_{n\to\infty} P_n$ und $Q_\infty := \lim\limits_{n\to\infty} Q_n$ gilt

<u>a</u>) $\exists P_\infty \Longleftrightarrow \exists B_\infty | \mathbb{R}(A)$, <u>b</u>) $\exists Q_\infty \Longleftrightarrow \exists \lim\limits_{n\to\infty} (AB_n)$,

<u>c</u>) $\exists B_\infty \Longrightarrow \exists P_\infty, \exists Q_\infty$, <u>d</u>) $\exists A^{-1} : \exists P_\infty \Longleftrightarrow \exists Q_\infty \Longleftrightarrow \exists B_\infty$.

Existiert B_∞, so erhält man

<u>e</u>) $P_\infty = I - B_\infty A$, $Q_\infty = I - AB_\infty$, $AP_\infty = Q_\infty A$.

<u>f</u>) B_∞ ist eine innere Inverse von A
$(AB_\infty A = A) \Longleftrightarrow AP_\infty = 0 \Longleftrightarrow Q_\infty A = 0$.

<u>g</u>) B_∞ ist eine äußere Inverse von A
$(B_\infty AB_\infty = B_\infty) \Longleftrightarrow P_\infty B_\infty = 0 \Longleftrightarrow B_\infty Q_\infty = 0$.

<u>h</u>) Ist P_∞ ein Projektor, so ist B_∞ eine äußere Inverse von A auf $\mathbb{R}(A)$.

<u>i</u>) Ist Q_∞ ein Projektor, so ist B_∞ eine innere Inverse von A auf $\mathbb{R}(B_\infty)$.

<u>j</u>) Sind P_∞ und Q_∞ Orthoprojektoren, so ist B_∞ die Moore-Penrosesche verallgemeinerte Inverse
$(AB_\infty A = A,\ B_\infty AB_\infty = B_\infty,\ (AB_\infty)^* = AB_\infty,\ (B_\infty A)^* = B_\infty A)$.

<u>Folgerung 2.6:</u> Existiert B_∞, so konvergiert das Iterationsverfahren (2.5) für alle Startelemente x_o und alle b. Dabei gilt

$$x_\infty := \lim_{n\to\infty} x_n = P_\infty x_o + B_\infty b \qquad (2.8)$$
$$= (I - \sum_{i=o}^{\infty} D_i Q_{i-1} A) x_o + \sum_{i=o}^{\infty} D_i Q_{i-1} b.$$

Beweis: Die Behauptung gewinnt man sofort aus (2.5'), Lemma 2.4b) und Lemma 2.5e).

Wir kommen nun zu der Frage, unter welchen Bedingungen der Iterationsprozeß (2.5) (verallgemeinerte) Lösungen von (2.1) liefert.

Lemma 2.7: Es existiere B_∞. In diesem Falle ist der Grenzwert $x_\infty := \lim_{n\to\infty} x_n$ genau dann eine Lösung (eine bezüglich (D_n) verallgemeinerte Lösung) von (2.1), wenn $Q_\infty r_o = 0$ $(\forall n : D_n Q_\infty r_o = 0)$ gilt.

Beweis: Zunächst liefert Folgerung 2.6 die Gleichung $x_\infty = P_\infty x_o + B_\infty b$. Daraus entsteht unter Beachtung von Lemma 2.5e)

$$Ax_\infty = AP_\infty x_o + AB_\infty b = Q_\infty Ax_o + (I - Q_\infty)b = b - Q_\infty(b - Ax_o) = b - Q_\infty r_o$$

bzw.

$$D_n Ax_\infty = D_n b - D_n Q_\infty r_o.$$

Damit sind die Behauptungen gezeigt.

Ist B_∞ eine innere Inverse von A, so gilt $Q_\infty A = 0$. Unter dieser Voraussetzung sind die Bedingungen $D_n Q_\infty r_o = 0$ $(\forall n)$, $Q_\infty r_o = 0$ für beliebige Startelemente x_o durch $D_n Q_\infty b = 0$ $(\forall n)$, $Q_\infty b = 0$ ersetzbar.

Soll x_∞ für beliebige b (und beliebige x_o) stets eine Lösung bzw. eine bezüglich (D_n) verallgemeinerte Lösung von (2.1) sein, so ist dafür die Bedingung $Q_\infty = 0$ bzw. $D_n Q_\infty = 0$ $(\forall n)$ notwendig und hinreichend.

Satz 2.8: Es seien folgende Voraussetzungen erfüllt:

a) Es gibt Lösungen (bezüglich (D_n) verallgemeinerte Lösungen) von (2.1).

b) Der Operator $P_\infty := \lim_{n\to\infty} P_n$ existiert. Außerdem gilt

$$\mathbb{R}(P_\infty) \subseteq \mathbb{N}(A)$$

$$\left(\mathbb{R}(P_\infty) \subseteq \bigcap_n \mathbb{N}(D_n A)\right), \text{ d. h. } AP_\infty = 0 \ (\forall n : D_n AP_\infty = 0).$$

Dann ist der Grenzwert $x_\infty := \lim_{n \to \infty} x_n$ für beliebige x_o vorhanden und stellt eine Lösung (eine bezüglich (D_n) verallgemeinerte Lösung) von (2.1) dar.

Beweis: Aufgrund von a) gibt es ein $x' \in X$ mit $Ax' = b$ bzw. $D_nAx' = D_nb$ $(\forall n)$. Daher erhält man

$$\begin{aligned} x_{n+1} &= T_nx_n + D_nb = T_nx_n + D_nAx' = T_nx_n + (I-T_n)x' , \\ &= P_nx_o + \sum_{i=o}^{n} T_nT_{n-1}\dots T_{i+1}(I-T_i)x' = P_nx_o + (I-P_n)x' \\ &= P_n(x_o-x') + x'. \end{aligned}$$

Nach Voraussetzung b) liefert der Grenzübergang $n \to \infty$

$$x_\infty = P_\infty(x_o-x') + x'.$$

Schließlich ist für $\mathbb{R}(P_\infty) \subseteq \mathbb{N}(A)$ und $Ax' = b$

$$Ax_\infty = AP_\infty(x_o-x') + Ax' = Ax' = b$$

bzw. für $\mathbb{R}(P_\infty) \subseteq \bigcap_n \mathbb{N}(D_nA)$ und $D_nAx' = D_nb$

$$D_nAx_\infty = D_nAP_\infty(x_o-x') + D_nAx' = D_nAx' = D_nb.$$

2.3 Verallgemeinerte Lösungen und zyklische Iterationsverfahren

Ist (D_n) eine zyklische Folge mit der Zyklenlänge k $(\forall n : D_n = D_{n+k})$, ergibt sich aus (2.5) durch Zusammenfassung von jeweils k Iterationsschritten bekanntlich ein stationäres Iterationsverfahen der Form

$$x_{(n+1)} := Tx_{(n)} + Db = T^{n+1}x_{(o)} + \sum_{i=o}^{n} T^iDb$$

$$(T:=P_{k-1},\ D:=B_{k-1}) \qquad (2.9)$$

mit der Restiteration

$$r_{(n+1)} = Sr_{(n)} = S^{n+1}r_{(o)} \qquad (S:=Q_{k-1},\ r_{(n)}:=b-Ax_{(n)}). \quad (2.10)$$

Dabei gilt

$$I - DA = I - \sum_{i=0}^{k-1} T_{k-1}T_{k-2}\cdots T_{i+1}D_iA = I - \sum_{i=0}^{k-1} T_{k-1}T_{k-2}\cdots T_{i+1}(I-T_i)$$
$$= P_{k-1} = T,$$

$$I - AD = I - \sum_{i=0}^{k-1} AD_iQ_{i-1} = I - \sum_{i=0}^{k-1}(I-S_i)S_{i-1}S_{i-2}\cdots S_o = Q_{k-1} = S$$

sowie

$$T^n = P_{nk-1},\ \sum_{i=0}^{n-1} T^iD = B_{nk-1},\ S^n = Q_{nk-1}.$$

Setzt man $x_{(o)} = x_o$, erhält man außerdem

$$x_{(n)} = x_{nk},\ r_{(n)} = r_{nk}.$$

Existiert x_∞, so für $x_{(o)} = x_o$ auch $x_{(\infty)} = x_\infty$. Umgekehrt muß aber aus der Existenz von $x_{(\infty)}$ nicht die von x_∞ folgen.

Das Verfahren (2.9) besitzt wieder die Gestalt (2.5), wobei anstelle von (D_n) jetzt die konstante Folge (D) tritt. Es ist in natürlicher Weise der Ersatzgleichung

$$DAx = Db \tag{2.11}$$

zugeordnet. Ist x eine bezüglich (D_n) verallgemeinerte Lösung von (2.1), dann ist x wegen

$$DAx=B_{k-1}Ax=\sum_{i=0}^{k-1} T_{k-1}T_{k-2}\cdots T_{i+1}D_iAx=\sum_{i=0}^{k-1} T_{k-1}T_{k-2}\cdots T_{i+1}D_ib=B_{k-1}b=Db$$

gleichzeitig eine bezüglich (D) verallgemeinerte Lösung von (2.1). Die Umkehrung gilt jedoch i. allg. nicht.

3. Endlich erzeugte kompakte Projektionsfolgen

Dieser Abschnitt enthält Ergebnisse der Arbeiten /7/ und /8/, die für die Konvergenzuntersuchung einer Klasse von Iterationsverfahren wesentlich sind (siehe Abschnitt 4).

Ist bei einem Iterationsverfahren (2.5) die Menge $\{D_n\}$ endlich und zugleich die Reihenfolge der Operatoren D_n in bestimmten Grenzen frei wählbar, wird man zweckmäßigerweise eine zyklische Strategie bevorzugen. Bei einer Reihe von Verfahren hängt dagegen die Wahl des Operators D_n von den Iterierten $x_n, x_{n-1}, \ldots, x_o$ ab. Hier ergibt sich i. allg. keine zyklische Folge (D_n). Daher ist die nachstehende Begriffsbildung auch von praktischem Interesse.

<u>Definition 3.1:</u> Eine Operatorfolge (U_n) heißt <u>endlich erzeugt</u>, falls es eine endliche Anzahl von Operatoren $U_{n_1}, U_{n_2}, \ldots, U_{n_m}$ aus der Menge $\{U_n\}$ mit folgenden Eigenschaften gibt:

<u>a</u>) Es ist $\{U_{n_1}, U_{n_2}, \ldots, U_{n_m}\} = \{U_n\}$.

<u>b</u>) Die konstanten Folgen $(U_{n_1}), (U_{n_2}), \ldots, (U_{n_m})$ sind Teilfolgen von (U_n).

<u>c</u>) Es gilt $\sup\limits_n \min \{k \mid k > n,\ U_k = U_n\} < \infty$.

Eine solche Folge (U_n) enthält also nur endlich viele verschiedene Operatoren, die in bestimmten, nicht über alle Grenzen wachsenden Abständen in der Folge immer wieder auftreten. Zyklische Operatorfolgen sind offenbar endlich erzeugt.

Es seien X und Y Hilberträume.

<u>Definition 3.2:</u> Eine Operatorfolge (U_n) mit $U_n \in L(X)$ (bzw. $U_n \in L(Y)$) heißt <u>kompakte Projektionsfolge</u>, falls mit den Abkürzungen

$$\mathbb{R} := \bigcap_n \mathbb{R}(U_n), \quad \mathbb{N} := \overline{\operatorname{lin} \bigcup_n \mathbb{N}(U_n)} \tag{3.1}$$

gilt:

<u>a</u>) Es ist $X = \mathbb{R} \oplus \mathbb{N}$ (bzw. $Y = \mathbb{R} \oplus \mathbb{N}$).

<u>b</u>) Die Operatoren U_n sind Projektoren, auf $\mathbb{N}$ bezüglich einer geeigneten (äquivalenten) energetischen Norm sogar Orthoprojektoren. (Der Nulloperator zählt hier als Orthoprojektor.)

<u>c</u>) Mindestens einer der Operatoren U_n ist auf $\mathbb{N}$ kompakt (vollstetig).

<u>Bemerkungen:</u> 1) Die Zerlegung $X = \mathbb{R} \oplus \mathbb{N}$ (bzw. $Y = \mathbb{R} \oplus \mathbb{N}$) definiert den Projektor

$$U = U^2,\ \mathbb{R}(U) = \mathbb{R},\ \mathbb{N}(U) = \mathbb{N}. \tag{3.2}$$

2) Für Orthoprojektoren U_n ist die Bedingung a) automatisch erfüllt. Dabei erhält man $\mathbb{R} = \mathbb{N}^{\perp}$, d. h., U ist wieder ein Orthoprojektor (siehe /7/).

3) Ein Projektor U_n ist genau dann auf $\mathbb{N}$ kompakt, wenn $\mathbb{R}(U_n|\mathbb{N})$ eine endliche Dimension hat. Einerseits gilt nämlich $U_n M = M$ für alle $M \subseteq \mathbb{R}(U_n|\mathbb{N})$. Andererseits ist die Dimension von $\mathbb{R}(U_n|\mathbb{N})$ genau in dem Falle endlich, wenn jede abgeschlossene und beschränkte Teilmenge von $\mathbb{R}(U_n|\mathbb{N})$ kompakt ist (siehe /10, S. 96 - 97/).

4) Für eine endlich erzeugte kompakte Projektionsfolge ergibt sich notwendigerweise:

- Es ist $\mathbb{R} = \bigcap_{i=1}^{m} \mathbb{R}(U_{n_i})$ und $\mathbb{N} = \overline{\operatorname{lin} \bigcup_{i=1}^{m} \mathbb{N}(U_{n_i})}$.

- Einer der Operatoren U_{n_i} $(i=1,2,\dots,m)$ ist kompakt auf $\mathbb{N}$.

Die folgenden Ergebnisse findet man in /7/ und /8/.

Satz 3.3: Die Folge (U_n) sei eine kompakte Projektionsfolge. Dann gilt:

a) Die Operatoren U_n lassen die Teilräume $\mathbb{R}$ und $\mathbb{N}$ invariant.

b) Der Projektor U (siehe (3.2)) ist ein Projektionskern von (U_n), d. h., es besteht die Beziehung $U = U_n U = U U_n$ $(\forall n)$.

Satz 3.4: Die Folge (U_n) $(U_n \in L(X))$ sei eine endlich erzeugte kompakte Projektionsfolge. Dann gilt

$$\lim_{n \to \infty} \|U_n U_{n-1} \dots U_o - U\| = 0.$$

Sind außerdem (E_n) $(E_n \in L(Y,X))$ und (F_n) $(F_n \in L(X,Y))$ (gleichmäßig) beschränkte Operatorfolgen mit $\mathbb{R}(E_n) \subseteq \mathbb{N}$ $(\forall n)$ und $\mathbb{N}(F_n) \supseteq \mathbb{R}$ $(\forall n)$, dann ist $\left(\sum_{i=o}^{n} U_n U_{n-1} \dots U_{i+1} E_i\right)$ beschränkt und $\left(\sum_{i=o}^{n} F_i U_{i-1} U_{i-2} \dots U_o\right)$ konvergent in der gleichmäßigen Operatortopologie.

Satz 3.5: Die Folge (U_n).$(U_n \in L(X))$ sei eine kompakte Projektionsfolge mit einem Zyklus der Länge k ($\forall$ n: $U_n=U_{n+k}$). Die Folge (E_n) $(E_n \in L(Y,X))$ besitze ebenfalls einen Zyklus der Länge k ($\forall$ n: $E_n=E_{n+k}$). Außerdem sei die Bedingung

$\mathbb{R}(\sum_{i=0}^{k-1} U_{k-1}U_{k-2}\dots U_{i+1}E_i) \subseteq \mathbb{N}$ erfüllt. Dann konvergiert die Teilfolge

$$\left(\sum_{i=0}^{n} (U_{k-1}U_{k-2}\dots U_o)^i \sum_{j=0}^{k-1} U_{k-1}U_{k-2}\dots U_{j+1}E_j\right)$$

von $(\sum_{i=0}^{n} U_nU_{n-1}\dots U_{i+1}E_i)$ gleichmäßig.

Für die Folge $(\sum_{i=0}^{n} F_iU_{i-1}U_{i-2}\dots U_o)$ bekommt man also bei entsprechenden Voraussetzungen über (U_n) stärkere Aussagen als für die Folge $(\sum_{i=0}^{n} U_nU_{n-1}\dots U_{i+1}E_i)$. Die Ursache liegt vor allem darin, daß die erste Folge eine Partialsummenfolge einer Reihe ist, die zweite dagegen i. allg. nicht. Die bestehenden Unterschiede illustriert das nachstehende einfache

Beispiel 3.6: Es sei $X = Y = R^2$ und

$$U_n = U_n^2 = U = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \quad (\forall n).$$

Man berechnet

$$\mathbb{R}(U) = \mathbb{R} = \left\{\alpha\begin{pmatrix} 1 \\ -1 \end{pmatrix} \middle| \alpha \in R\right\},\ \mathbb{N}(U) = \mathbb{N} = \left\{\beta\begin{pmatrix} 1 \\ 1 \end{pmatrix} \middle| \beta \in R\right\}.$$

Wegen

$$U|\mathbb{N} = 0|\mathbb{N},\ R^2 = \mathbb{R} \oplus \mathbb{N}$$

ist $(U_n) = (U)$ eine endlich erzeugte kompakte Projektionsfolge. Für die beschränkte Operatorfolge (E_n) mit

$$E_{2n} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix},\ E_{2n+1} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \quad (\forall n)$$

erhält man

$$\mathbb{R}(E_n) = \left\{\beta\begin{pmatrix} 1 \\ 1 \end{pmatrix} \middle| \beta \in R\right\} = \mathbb{N} \quad (\forall n).$$

Aus $UE_n = 0$ ($\forall$ n) folgt

$$\sum_{i=0}^{n} U_n U_{n-1} \dots U_{i+1} E_i = E_n + \sum_{i=0}^{n-1} UE_i = E_n \quad (\forall n) \tag{3.3}$$

und damit im Einklang mit Satz 3.4 die Beschränktheit, nicht aber die Konvergenz der Folge $(\sum_{i=0}^{n} U_n U_{n-1} \dots U_{i+1} E_i)$. Zum anderen besitzen (U_n) und (E_n) einen gemeinsamen Zyklus der Länge 2. Schließlich ist auch die Bedingung $\mathbb{R}(U_1 E_0 + E_1) = \mathbb{R}(E_1) = \mathbb{N}$ erfüllt. Die Teilfolge

$$(\sum_{i=0}^{n} (U_1 U_0)^i (U_1 E_0 + E_1)) = (\sum_{i=0}^{n} U^i E_1) = (E_1)$$

konvergiert offensichtlich und bestätigt so die Aussage des Satzes 3.5. Setzt man allerdings auch

$$F_{2n} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \quad F_{2n+1} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \quad (\forall n),$$

divergiert die Folge

$$(\sum_{i=0}^{n} F_i U_{i-1} U_{i-2} \dots U_0) = (F_0 + \sum_{i=0}^{n} F_i U) \tag{3.4}$$

wegen

$$F_{2n} U = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}, \quad F_{2n+1} U = \frac{1}{2} \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix} \quad (\forall n).$$

Dieses Resultat steht aber nicht im Widerspruch zu Satz 3.4, da hier die Bedingung $\mathbb{N}(F_n) \supseteq \mathbb{R}$ ($\forall$ n) verletzt ist. Für

$$F_{2n} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad F_{2n+1} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \quad (\forall n)$$

dagegen gilt

$$\mathbb{N}(F_n) = \left\{\alpha \begin{pmatrix} 1 \\ -1 \end{pmatrix} \middle| \alpha \in R\right\} = \mathbb{R} \quad (\forall n).$$

Hier konvergiert die Folge (3.4) im Einklang mit Satz 3.4 wegen $F_n U = 0$ ($\forall$ n).

4. Iterationsverfahren mit endlich erzeugten kompakten Projektionsfolgen

In diesem Abschnitt werden die bisherigen Resultate benutzt, um Aussagen zur Konvergenz des Iterationsverfahrens (2.5) und zur (verallgemeinerten) Lösbarkeit der Gleichung (2.1) herzuleiten. Eine Darstellung von $x_\infty = \lim_{n\to\infty} x_n$ wird nicht explizit angegeben. Hierzu sei auf Abschnitt 2 verwiesen.

<u>Definition 4.1:</u> Die Iterationsfolge (x_n) des Verfahrens (2.5) heißt <u>P- bzw. G-konvergent</u>, falls sowohl (P_n) als auch (B_n) in der punktweisen bzw. gleichmäßigen Operatortopologie konvergieren.

P-konvergente und erst recht G-konvergente Iterationsfolgen konvergieren also global, d. h. für beliebige x_0 und b.

<u>Satz 4.2:</u> Es seien folgende Voraussetzungen erfüllt:

a) Es existieren bezüglich (D_n) verallgemeinerte Lösungen von (2.1).

b) Die Folge $(T_n) = (I-D_nA)$ ist eine endlich erzeugte kompakte Projektionsfolge.

Dann ist (x_n) für beliebige x_0 konvergent. Weiterhin stellt der Grenzwert x_∞ stets eine bezüglich (D_n) verallgemeinerte Lösung von (2.1) dar.

<u>Beweis:</u> Nach Satz 3.4 konvergiert (P_n) gleichmäßig gegen den Projektionskern

$$P_\infty = P_\infty^2,\ \mathbb{R}(P_\infty) = \bigcap_n \mathbb{R}(T_n) = \bigcap_n \mathbb{N}(D_nA), \tag{4.1}$$
$$\mathbb{N}(P_\infty) = \overline{\operatorname{lin} \bigcup_n \mathbb{N}(T_n)} = \overline{\operatorname{lin} \bigcup_n \mathbb{R}(D_nA)}$$

von (T_n). Satz 2.8 liefert die Behauptung.

<u>Bemerkung:</u> Gilt a) für alle rechten Seiten b von (2.1), so ist (x_n) P-konvergent.

Satz 4.3: Es seien folgende Voraussetzungen erfüllt:

a) Die Folge (D_n) besitzt einen Zyklus der Länge k
$(\forall n: D_n = D_{n+k})$.

b) Die Folge $(T_n) = (I-D_nA)$ ist eine kompakte Projektionsfolge.

c) Es gilt $\overline{\mathrm{lin}\bigcup_n \mathbb{R}(D_n)} = \overline{\mathrm{lin}\bigcup_n \mathbb{R}(D_nA)}$.

Dann ist die Teilfolge $(x_{(n)})$ von (x_n) G-konvergent (siehe (2.9)). Weiterhin stellt der Grenzwert $x_{(\infty)}$ stets eine bezüglich $(D) = (B_{k-1})$ verallgemeinerte Lösung von (2.1) dar.

Beweis: Zunächst hat mit (D_n) auch $(I-D_nA)$ einen Zyklus der Länge k. Nach Satz 3.4 konvergiert (P_n) gleichmäßig gegen den Projektionskern (4.1). Außerdem findet man unter Beachtung von Lemma 2.4b)

$$\mathbb{R}(D) = \mathbb{R}(B_{k-1}) = \mathbb{R}\left(\sum_{i=0}^{k-1} D_i Q_{i-1}\right) \subseteq \overline{\mathrm{lin}\bigcup_n \mathbb{R}(D_n)} = \overline{\mathrm{lin}\bigcup_n \mathbb{R}(D_nA)}$$
$$= \mathbb{N}(P_\infty). \qquad (4.2)$$

Satz 3.5 sichert nun die gleichmäßige Konvergenz der Reihe $\sum_{i=0}^{\infty} T^i D$ $(T=P_{k-1})$. Dabei gilt $T^\infty := \lim_{n\to\infty} T^n = P_\infty$. Wegen (2.9) ist $(x_{(n)})$ somit G-konvergent. Schließlich erhält man mit der Abkürzung $S^\infty := \lim_{n\to\infty} S^n$

$$DS^\infty = D\left(I-A\sum_{i=0}^{\infty} T^i D\right) = D - \sum_{i=0}^{\infty}(I-T)T^i D = D-(I-T^\infty)D = T^\infty D = 0,$$

wenn man $T^\infty = P_\infty$ und (4.2) berücksichtigt. Nach Lemma 2.7 ist $x_{(\infty)}$ daher stets eine bezüglich (D) verallgemeinerte Lösung von (2.1).

Bemerkungen: 1) Die Voraussetzung c) des Satzes liegt offenbar vor, wenn $\mathbb{R}(D_n) = \mathbb{R}(D_nA)$ $(\forall n)$ gilt.

2) Die Folge (x_n) braucht unter den Voraussetzungen des Satzes

nicht konvergent zu sein. Für $X = Y = \mathbb{R}^2$ und

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad D_{2n} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \quad D_{2n+1} = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \quad (\forall n)$$

entsteht

$$T_n = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad B_{2n} = D_{2n}, \quad B_{2n+1} = D_{2n+1} \quad (\forall n)$$

(vgl. Beispiel 3.6). Nach Satz 3.4 ist (x_n) wegen $\mathbb{R}(D_n) \subseteq \mathbb{N}(P_\infty)$ und

$$\|x_{n+1}\| \leq \|P_n\| \, \|x_o\| + \|B_n\| \, \|b\|$$

aber zumindest beschränkt.

<u>Satz 4.4:</u> Es seien folgende Voraussetzungen erfüllt:

<u>a</u>) Die Folge (D_n) ist endlich erzeugt.

<u>b</u>) Die Folge $(S_n) = (I-AD_n)$ ist eine kompakte Projektionsfolge.

<u>c</u>) Es gilt $\bigcap_n \mathbb{N}(AD_n) = \bigcap_n \mathbb{N}(D_n)$.

Dann ist (x_n) G-konvergent. Weiterhin stellt der Grenzwert x_∞ stets eine bezüglich (D_n) verallgemeinerte Lösung von (2.1) dar.

<u>Beweis:</u> Nach Satz 3.4 konvergiert (Q_n) gleichmäßig gegen den Projektionskern

$$\begin{aligned} Q_\infty = Q_\infty^2, \quad \mathbb{R}(Q_\infty) &= \bigcap_n \mathbb{R}(S_n) = \bigcap_n \mathbb{N}(AD_n) = \bigcap_n \mathbb{N}(D_n), \\ \mathbb{N}(Q_\infty) &= \overline{\operatorname{lin} \bigcup_n \mathbb{N}(S_n)} = \overline{\operatorname{lin} \bigcup_n \mathbb{R}(AD_n)} \end{aligned} \tag{4.3}$$

von (S_n). Daher ergibt sich für die nach <u>a</u>) beschränkte Folge (D_n) die Beziehung $\mathbb{N}(D_n) \supseteq \mathbb{R}(Q_\infty)$ $(\forall n)$. Satz 3.4 sichert also auch die gleichmäßige Konvergenz von $(B_n) = (\sum_{i=0}^{n} D_i Q_{i-1})$ und $(P_n) = (I-B_nA)$ (siehe Lemma 2.4<u>c</u>)). Demnach ist (x_n) G-konvergent. Außerdem folgt aus $\mathbb{N}(D_n) \supseteq \mathbb{R}(Q_\infty)$ $(\forall n)$ unmittelbar $D_n Q_\infty r_o = 0$ $(\forall n)$. Lemma 2.7 liefert damit die letzte Behauptung.

5. Eine Klasse von Projektionsverfahren

5.1 Wir wollen jetzt die Ergebnisse des Abschnittes 4 auf eine Klasse von Projektionsverfahren anwenden.
Gegeben seien zwei Hilberträume X,Y und eine Folge (Z_n) von weiteren Hilberträumen. Wir betrachten die Gleichung

$$Ax = b \quad (A \in L(X,Y),\ b \in Y) \tag{5.1}$$

und das Ersatzgleichungssystem

$$G_nAx = G_nb \quad (G_n \in L(Y,Z_n);\ n=0,1,2,\ldots) \tag{5.2}$$

mit normal auflösbaren Operatoren A und $A_n := G_nA \in L(X,Z_n)$ $(\mathbb{R}(A) = \overline{\mathbb{R}(A)},\ \mathbb{R}(A_n) = \overline{\mathbb{R}(A_n)})$. Bevor wir uns den Projektionsverfahren zuwenden, beweisen wir zwei einfache Hilfssätze. Die dabei benutzten Beziehungen zwischen Nullräumen und Wertebereichen von Operatoren findet man etwa in /10, S. 250/.

<u>Lemma 5.1:</u> Es gilt

$$\mathbb{N}(A_n^*) = \{0\} \Leftrightarrow \mathbb{R}(A_n) = Z_n \Leftrightarrow \exists (A_nA_n^*)^{-1} \in L(Z_n).$$

<u>Beweis:</u> 1) Die erste Äquivalenz folgt aus $Z_n = \mathbb{R}(A_n) \oplus \mathbb{N}(A_n^*)$.
2) Es sei $\mathbb{N}(A_n^*) = \{0\}$. Wir geben uns ein beliebiges $z \in Z_n$, $z \neq 0$, vor. Damit ist $A_n^*z \neq 0$. Aufgrund der Zerlegung $X = \mathbb{N}(A_n) \oplus \mathbb{R}(A_n^*)$ gehört A_n^*z nicht zu $\mathbb{N}(A_n)$. Also erhält man $A_nA_n^*z \neq 0$ $(\forall z \neq 0)$ bzw. $\mathbb{N}(A_nA_n^*) = \{0\}$.
Wegen $Z_n = \mathbb{R}(A_n)$ und $X = \mathbb{N}(A_n) \oplus \mathbb{R}(A_n^*)$ existiert zu jedem $z \in Z_n$ ein $z' \in Z_n$ mit $z = A_nA_n^*z'$. Daher gilt $Z_n = \mathbb{R}(A_nA_n^*)$.
Aus $\mathbb{N}(A_nA_n^*) = \{0\}$ und $\mathbb{R}(A_nA_n^*) = Z_n$ ergibt sich nach einem bekannten Satz von Banach die Existenz der stetigen Inversen $(A_nA_n^*)^{-1}$.

3) Die stetige Inverse $(A_nA_n^*)^{-1}$ sei vorhanden. Nimmt man an, daß es ein $z \in \mathbb{N}(A_n^*)$ mit $z \neq 0$ gibt, dann bekommt man für dieses z im Widerspruch zur Voraussetzung $A_nA_n^*z = 0$. Demnach ist $\mathbb{N}(A_n^*) = \{0\}$.

Lemma 5.2: Existiert $(A_nA_n^*)^{-1} \in L(Z_n)$, dann sind $A_nA_n^*$ und $(A_nA_n^*)^{-1}$ positiv definite selbstadjungierte Operatoren.

Beweis: Es sei $(A_nA_n^*)^{-1} \in L(Z_n)$. Wegen $(A_nA_n^*)^* = A_nA_n^*$ und $((A_nA_n^*)^{-1})^* = ((A_nA_n^*)^*)^{-1}$ sind $A_nA_n^*$ und $(A_nA_n^*)^{-1}$ selbstadjungiert.
Mit A_n ist auch A_n^* normal auflösbar ($\mathbb{R}(A_n^*) = \overline{\mathbb{R}(A_n^*)}$). Nach Lemma 5.1 gilt außerdem $\mathbb{N}(A_n^*) = \{0\}$. Daher gibt es für alle $z \in Z_n$ ein $m > 0$ mit $\|A_n^*z\| \geq m\|z\|$ (siehe /10, S. 86/). Somit entsteht

$$(A_nA_n^*z,z) = (A_n^*z,A_n^*z) = \|A_n^*z\|^2 \geq m^2\|z\|^2.$$

Ebenso gelangt man mit $z = (A_nA_n^*)^{-1}z'$ zu

$$((A_nA_n^*)^{-1}z',\ z') = (z,A_nA_n^*z) = \|A_n^*z\|^2 \geq m^2\|z\|^2 \geq \frac{m^2}{\|A_nA_n^*\|^2}\|z'\|^2.$$

Demnach sind $A_nA_n^*$ und $(A_nA_n^*)^{-1}$ positiv definit.

5.2 Die Folgen (G_n) und (Z_n) seien nun so gewählt, daß die Operatoren A_n den Raum X auf die Räume Z_n abbilden ($\mathbb{R}(A_n) = Z_n$). Damit sind die A_n automatisch normal auflösbar. Weiterhin bedeutet das:

- Es ist auch $\mathbb{R}(G_n) = Z_n$.
- Es gilt $\mathbb{N}(A_n^*) = \{0\}$.
- Die Operatoren $A_nA_n^*$ und $(A_nA_n^*)^{-1}$ sind selbstadjungiert und positiv definit. Außerdem bilden sie Z_n auf sich ab.

Mit diesen Voraussetzungen untersuchen wir das Iterationsverfahren (2.5) in der speziellen Gestalt

$$x_{n+1} = T_nx_n + D_nb \quad (T_n := I-D_nA,\ D_n := A_n^*(A_nA_n^*)^{-1}G_n \in L(Y,X)). \qquad (5.3)$$

Daraus resultieren einige Besonderheiten, die wir zunächst angeben wollen.

Lemma 5.3: Es gilt $T_n^2 = T_n = T_n^*$ und $\mathbb{R}(T_n) = \mathbb{N}(D_nA) = \mathbb{N}(A_n)$, $\mathbb{N}(T_n) = \mathbb{R}(D_nA) = \mathbb{R}(D_n) = \mathbb{R}(A^*D_n^*) = \mathbb{R}(A_n^*)$.

Beweis: Mit $D_nA = A_n^*(A_nA_n^*)^{-1}A_n$ gewinnt man leicht

$$(D_nA)^2 = D_nA = (D_nA)^* = A^*D_n^*.$$

Daraus folgt wegen $T_n = I - D_nA$ unmittelbar $T_n^2 = T_n = T_n^*$.

Weiter erhält man aus $D_nAx = A_n^*(A_nA_n^*)^{-1}A_nx = 0$ sofort $(A_nA_n^*)^{-1}A_nx \in \mathbb{N}(A_n^*) = \{0\}$ und $A_nx = 0$. Umgekehrt zieht $A_nx = 0$ offensichtlich $D_nAx = 0$ nach sich. Daher ist

$\mathbb{R}(T_n) = \mathbb{N}(D_nA) = \mathbb{N}(A_n)$. Außerdem gilt

$\mathbb{N}(T_n) = \mathbb{R}(D_nA) = \mathbb{R}(A^*D_n^*)$ und $\mathbb{N}(T_n) = \mathbb{R}(T_n)^\perp = \mathbb{N}(A_n)^\perp = \mathbb{R}(A_n^*)$.

Aufgrund von $\mathbb{R}(G_n) = \mathbb{R}(G_nA) = Z_n$ ist schließlich

$\mathbb{R}(D_n) = \mathbb{R}(D_nA)$.

Lemma 5.4: Das Iterationsverfahren (5.3) projiziert nacheinander orthogonal auf die Lösungsmengen $X(A_n, G_nb)$ der Ersatzgleichungen (5.2).

Beweis: Die Mengen $X(A_n, G_nb)$ und $\mathbb{N}(A_n)$ verlaufen parallel. Nach Lemma 5.3 ergibt sich somit

$$P_nx_n = T_nx_n + (I-T_n)x_n',$$

wenn P_n der orthogonale Projektor auf $X(A_n, G_nb)$ und x_n' ein beliebiges Element von $X(A_n, G_nb)$ ist. Dabei garantiert die Voraussetzung $\mathbb{R}(A_n) = Z_n$ die Existenz eines solchen x_n'. Andererseits bekommt man auch

$$x_{n+1} = T_nx_n + A_n^*(A_nA_n^*)^{-1}G_nb = T_nx_n + A_n^*(A_nA_n^*)^{-1}A_nx_n' = T_nx_n + (I-T_n)x_n'.$$

Also ist $x_{n+1} = P_nx_n$.

Lemma 5.5: Für die Operatoren $S_n = I - AD_n$ gilt $S_n^2 = S_n$ und

$$\mathbb{R}(S_n) = \mathbb{N}(AD_n) = \mathbb{N}(D_n) = \mathbb{N}(G_n),\ \mathbb{N}(S_n) = \mathbb{R}(AD_n) = \mathbb{R}(AA_n^*).$$

Beweis: Zunächst findet man

$$(AD_n)^2 = AA_n^*(A_nA_n^*)^{-1}G_nAA_n^*(A_nA_n^*)^{-1}G_n = AD_n.$$

Wegen $S_n = I - AD_n$ folgt daraus $S_n^2 = S_n$. Die Gleichung

$$AD_ny = AA^*G_n^* (A_nA_n^*)^{-1}G_ny = 0$$

zieht unter Beachtung von $X = \mathbb{N}(A) \oplus \mathbb{R}(A^*)$ und $\mathbb{N}(A_n^*) = \{0\}$ sowohl $D_ny = A_n^*(A_nA_n^*)^{-1}G_ny = 0$ als auch $G_ny = 0$ nach sich. Umgekehrt ergibt $G_ny = 0$ natürlich $AD_ny = 0$. Das bedeutet aber $\mathbb{R}(S_n) = \mathbb{N}(AD_n) = \mathbb{N}(D_n) = \mathbb{N}(G_n)$. Aus $Z_n = \mathbb{R}(G_n) = \mathbb{R}((A_nA_n^*)^{-1}G_n)$ erhält man schließlich $\mathbb{N}(S_n) = \mathbb{R}(AD_n) = \mathbb{R}(AA_n^*)$.

Lemma 5.6: Ist $P \in L(X)$ ein Orthoprojektor, dann ist jeder Operator $Q \in L(Y)$ mit $AP = QA$ ein $(AA^*)^{-1}$-Orthoprojektor in $\mathbb{R}(A) = \overline{\mathbb{R}(A)}$.

Beweis: Aus $AP = QA$ folgt $Q^2A = QAP = AP^2 = AP = QA$, d. h. $Q^2|\mathbb{R}(A) = Q|\mathbb{R}(A)$. Außerdem läßt Q den Raum $\mathbb{R}(A)$ invariant. Weiterhin ist $(AA^*)^{-1}|\mathbb{R}(A)$ ein stetiger, selbstadjungierter und positiv definiter Operator (vgl. Lemmata 5.1, 5.2). Schließlich bekommt man auf $\mathbb{R}(A)$

$$(AA^*)Q^* = A(QA)^* = A(AP)^* = AP^*A^* = APA^* = QAA^*$$

und damit

$$((AA^*)^{-1}Q)^* = Q^*(AA^*)^{-1} = (AA^*)^{-1}Q.$$

Folgerung 5.7: Die Operatoren S_n und $AD_n = I - S_n$ sind $(AA^*)^{-1}$-Orthoprojektoren in $\mathbb{R}(A)$.

Lemma 5.8: Die Folgen (D_n) und (G_n) erzeugen äquivalente Ersatzsysteme (siehe Abschnitt 2.1).

Beweis: Nach Lemma 5.5 ist $\mathbb{N}(D_n) = \mathbb{N}(G_n)$ $(\forall n)$. Damit gewinnt man die Behauptung aus Lemma 2.2.

Lemma 5.9: Es gilt mit

$$\mathbb{R} = \bigcap_n \mathbb{R}(T_n) = \bigcap_n \mathbb{N}(A_n), \quad \mathbb{N} = \overline{\operatorname{lin} \bigcup_n \mathbb{N}(T_n)} = \overline{\operatorname{lin} \bigcup_n \mathbb{R}(A_n^*)}$$

die Beziehung $X = \mathbb{R} \oplus \mathbb{N}$.

__Beweis:__ Die Behauptungen ergeben sich aus Lemma 5.3. Insbesondere sind die T_n Orthoprojektoren, so daß die Gleichung $X = \mathbb{R} \oplus \mathbb{N}$ besteht (siehe Abschnitt 3).

Nach diesen Vorbereitungen können die Konvergenzsätze des Abschnittes 4 für das Verfahren (5.3) konkretisiert werden.

__Satz 5.10:__ Es seien folgende Voraussetzungen erfüllt:

a) Es existieren bezüglich (G_n) verallgemeinerte Lösungen von (5.1).

b) Die Folge (G_n) ist endlich erzeugt (siehe Definition 3.1).

c) Mindestens einer der Operatoren T_n ist kompakt auf

$$\mathbb{N} = \overline{\mathrm{lin} \bigcup_n \mathbb{R}(A_n^*)}.$$

Dann ist (x_n) für beliebige x_0 konvergent. Weiterhin stellt der Grenzwert x_∞ stets eine bezüglich (G_n) verallgemeinerte Lösung von (5.1) dar.

__Beweis:__ Mit (G_n) ist auch (T_n) endlich erzeugt. Außerdem erweist sich (T_n) in Verbindung mit Lemma 5.3, Lemma 5.9 und Voraussetzung c) als kompakte Projektionsfolge. Nun resultieren die Behauptungen aus Satz 4.2, wenn man Lemma 5.8 berücksichtigt.

__Bemerkung:__ Gilt a) für alle rechten Seiten b von (5.1), so ist (x_n) P-konvergent.

__Satz 5.11:__ Es seien folgende Voraussetzungen erfüllt:

a) Die Folge (G_n) besitzt eine Zyklus der Länge k

$(\forall n\colon G_n = G_{n+k})$.

b) Mindestens einer der Operatoren T_n ist kompakt auf

$$\mathbb{N} = \overline{\mathrm{lin} \bigcup_n \mathbb{R}(A_n^*)}.$$

Dann ist die Teilfolge $(x_{(n)})$ von (x_n) G-konvergent (siehe (2.9)). Weiterhin stellt der Grenzwert $x_{(\infty)}$ stets eine bezüglich $(D) = (B_{k-1})$ verallgemeinerte Lösung von (5.1) dar.

Beweis: Mit (G_n) hat auch (D_n) einen Zyklus der Länge k. (T_n) ist nach Lemma 5.3, Lemma 5.9 und Voraussetzung b) eine kompakte Projektionsfolge. Außerdem gilt aufgrund von Lemma 5.3 die Gleichung $\mathbb{R}(D_n) = \mathbb{R}(D_nA)$ $(\forall n)$. Damit besteht offenbar die Beziehung

$$\overline{\operatorname{lin} \bigcup_n \mathbb{R}(D_n)} = \overline{\operatorname{lin} \bigcup_n \mathbb{R}(D_nA)}.$$

Satz 4.3 liefert nun die Behauptung.

Satz 5.12: Es seien folgende Voraussetzungen erfüllt:

a) Die Folge (G_n) ist endlich erzeugt.

b) Es gilt $Y = \bigcap_n \mathbb{N}(G_n) \oplus \overline{\operatorname{lin} \bigcup_n \mathbb{R}(AA_n^*)}$.

c) Mindestens einer der Operatoren S_n ist kompakt auf $\overline{\operatorname{lin} \bigcup_n \mathbb{R}(AA_n^*)}$.

Dann ist (x_n) G-konvergent. Weiterhin stellt der Grenzwert x_∞ stets eine bezüglich (G_n) verallgemeinerte Lösung von (5.1) dar.

Beweis: Die Behauptungen ergeben sich sofort aus Satz 4.4, wenn man Lemma 5.5 und Folgerung 5.7 beachtet. Die S_n sind nämlich wegen $\mathbb{N} = \overline{\operatorname{lin} \bigcup_n \mathbb{R}(AA_n^*)} \subseteq \mathbb{R}(A)$ auch $(AA^*)^{-1}$-Orthoprojektoren in $\mathbb{N}$.

Ist $X = R^N$, $Y = R^M$ und (G_n) eine zyklische Folge bestimmter Zeilenauswahlmatrizen, so entsteht das in /6/ betrachtete PSH-Verfahren (Projektion auf Schnitträume von Hyperebenen). Dabei sind in Satz 5.10 die Voraussetzungen b), c), in Satz 5.11 die Voraussetzungen a), b) und in Satz 5.12 die Voraussetzungen a), c) von vornherein erfüllt. Das Hauptergebnis der Arbeit /6/ ist im wesentlichen in Satz 5.11 enthalten. Die Arbeiten /1/ und /9/ untersuchen spezielle Varianten des PSH-Verfahrens.

6. Eine Klasse von SPA-ähnlichen Verfahren

6.1 In diesem Abschnitt sollen Ergebnisse des Abschnitts 4 auf eine Klasse von SPA-ähnlichen Verfahren angewandt werden.

Gegeben seien zwei Hilberträume X,Y und eine Folge (Z_n) von weiteren Hilberträumen. Wir betrachten die Gleichung

$$Ax = b \quad (A \in L(X,Y),\ b \in Y) \tag{6.1}$$

und das Ersatzgleichungssystem

$$H_n^*A^*Ax = H_n^*A^*b = (AH_n)^*b \quad (H_n \in L(Z_n,X);\ n=0,1,2,\ldots) \tag{6.2}$$

mit normal auflösbaren Operatoren A und $C_n := AH_n \in L(Z_n,Y)$ $(\mathbb{R}(A) = \overline{\mathbb{R}(A)},\ \mathbb{R}(C_n) = \overline{\mathbb{R}(C_n)})$. Entsprechend wie in Abschnitt 5 beweist man:

<u>Lemma 6.1:</u> Es gilt

$$\mathbb{N}(C_n) = \{0\} \Leftrightarrow \mathbb{R}(C_n^*) = Z_n \Leftrightarrow \exists (C_n^*C_n)^{-1} \in L(Z_n).$$

<u>Lemma 6.2:</u> Existiert $(C_n^*C_n)^{-1} \in L(Z_n)$, dann sind $C_n^*C_n$ und $(C_n^*C_n)^{-1}$ selbstadjungierte positiv definite Operatoren.

6.2 Die Folgen (H_n) und (Z_n) seien nun so gewählt, daß die Operatoren C_n^* den Raum Y auf die Räume Z_n abbilden $(\mathbb{R}(C_n^*) = Z_n)$. Damit sind die C_n automatisch normal auflösbar. Weiterhin bedeutet das:

- Es ist auch $\mathbb{R}(H_n^*) = Z_n$.
- Es gilt $\mathbb{N}(C_n) = \{0\}$.
- Die Operatoren $C_n^*C_n$ und $(C_n^*C_n)^{-1}$ sind selbstadjungiert und positiv definit. Außerdem bilden sie Z_n auf sich ab.

Mit diesen Voraussetzungen untersuchen wir das Iterationsverfahren (2.5) in der speziellen Gestalt

$$x_{n+1} = T_nx_n + D_nb \quad (T_n := I - D_nA,\ D_n := H_n(C_n^*C_n)^{-1}C_n^* \in L(Y,X)). \tag{6.3}$$

Daraus resultieren einige Besonderheiten, die wir zunächst angeben wollen.

<u>Lemma 6.3:</u> Für die Operatoren $S_n = I - AD_n$ gilt $S_n^2 = S_n = S_n^*$ und

$$\mathbb{R}(S_n) = \mathbb{N}(AD_n) = \mathbb{N}(D_n) = \mathbb{N}(C_n^*),\quad \mathbb{N}(S_n) = \mathbb{R}(AD_n) = \mathbb{R}(D_n^*A^*) = \mathbb{R}(C_n).$$

Beweis: Mit $AD_n = C_n(C_n^*C_n)^{-1}C_n^*$ gewinnt man leicht

$$(AD_n)^2 = AD_n = (AD_n)^* = D_n^*A^*.$$

Daraus folgt wegen $S_n = I - AD_n$ unmittelbar $S_n^2 = S_n = S_n^*$. Weiter erhält man aus $D_n y = 0$ schrittweise $AD_n y = C_n(C_n^*C_n)^{-1}C_n^*y=0$, $(C_n^*C_n)^{-1}C_n^*y \in \mathbb{N}(C_n) = \{0\}$ und $C_n^*y = 0$. Umgekehrt zieht $C_n^*y = 0$ offenbar $D_n y = 0$ nach sich. Daher ist $\mathbb{R}(S_n) = \mathbb{N}(AD_n) = \mathbb{N}(D_n) = \mathbb{N}(C_n^*)$. Außerdem gilt $\mathbb{N}(S_n) = \mathbb{R}(AD_n) = \mathbb{R}(D_n^*A^*)$ und

$$\mathbb{N}(S_n) = \mathbb{R}(S_n)^{\perp} = \mathbb{N}(C_n^*)^{\perp} = \mathbb{R}(C_n).$$

Lemma 6.4: Das zum Iterationsverfahren (6.3) gehörende Restverfahren $r_{n+1} = S_n r_n$ projiziert nacheinander orthogonal auf die Lösungsmengen der Gleichungen $C_n^*r = H_n^*A^*r = 0$.

Beweis: Die Behauptung ergibt sich sofort aus Lemma 6.3.

Lemma 6.5: Es gilt $T_n^2 = T_n$ und

$$\mathbb{R}(T_n) = \mathbb{N}(D_nA) = \mathbb{N}(C_n^*A),\quad \mathbb{N}(T_n) = \mathbb{R}(D_nA) = \mathbb{R}(D_n) = \mathbb{R}(H_n).$$

Beweis: Zunächst ist

$$(D_nA)^2 = H_n(C_n^*C_n)^{-1}C_n^*AH_n(C_n^*C_n)^{-1}C_n^*A = D_nA.$$

Wegen $T_n = I - D_nA$ folgt $T_n^2 = T_n$. Die Gleichung $D_nAx = 0$ liefert

$$AD_nAx = C_n(C_n^*C_n)^{-1}C_n^*Ax = 0.$$

Unter Beachtung von $\mathbb{N}(C_n) = \mathbb{N}(C_n(C_n^*C_n)^{-1}) = \{0\}$ bekommt man daraus $C_n^*Ax = 0$. Umgekehrt ist jede Lösung von $C_n^*Ax = 0$ auch Lösung von $D_nAx = 0$. Also gilt $\mathbb{R}(T_n) = \mathbb{N}(D_nA) = \mathbb{N}(C_n^*A)$. Wegen $\mathbb{R}(C_n^*) = Z_n$ und $(C_n^*C_n)^{-1} \in L(Z_n)$ ergibt sich $\mathbb{R}(C_n^*) = \mathbb{R}((C_n^*C_n)^{-1}C_n^*) = Z_n$ und $\mathbb{R}(D_n) = H_nZ_n = \mathbb{R}(H_n)$. Unter Ausnutzung der Beziehung $Y = \mathbb{N}(A^*) \oplus \mathbb{R}(A)$ gelangt man schließlich zu

$$\mathbb{R}(D_nA) = \mathbb{R}(H_n(C_n^*C_n)^{-1}H_n^*A^*A) = \mathbb{R}(D_n).$$

<u>Lemma 6.6:</u> Ist $Q \in L(Y)$ ein Orthoprojektor, dann ist jeder Operator $P \in L(X)$ mit $AP = QA$ ein A^*A-Orthoprojektor in $\mathbb{R}(A^*) = \overline{\mathbb{R}(A^*)}$.

<u>Beweis:</u> Aus $AP = QA$ entsteht $A^*Q^* = P^*A^*$. Analog wie in Lemma 5.6 kann man zeigen, daß P^* ein $(A^*A)^{-1}$-Orthoprojektor in $\mathbb{R}(A^*)$ ist. Das bedeutet auf $\mathbb{R}(A^*)$

$$((A^*A)^{-1}P^*)^* = P(A^*A)^{-1} = (A^*A)^{-1}P^*.$$

Daraus gewinnt man auf $\mathbb{R}(A^*)$ die Beziehungen

$$P = (A^*A)^{-1}P^*(A^*A) = P^2, \; A^*AP = P^*A^*A = (A^*AP)^*.$$

<u>Folgerung 6.7:</u> Die Operatoren T_n und $D_nA = I - T_n$ sind A^*A-Orthoprojektoren in $\mathbb{R}(A^*)$.

<u>Lemma 6.8:</u> Die Folgen (D_n) und $(C_n^*) = (H_n^*A^*)$ erzeugen äquivalente Ersatzsysteme (siehe Abschnitt 2.1).

<u>Beweis:</u> Nach Lemma 6.3 ist $\mathbb{N}(D_n) = \mathbb{N}(C_n^*)$ $(\forall n)$. Damit folgt die Behauptung aus Lemma 2.2.

<u>Lemma 6.9:</u> Es gilt mit

$$\mathbb{R} = \bigcap_n \mathbb{R}(S_n) = \bigcap_n \mathbb{N}(C_n^*), \; \mathbb{N} = \overline{\operatorname{lin} \bigcup_n \mathbb{N}(S_n)} = \overline{\operatorname{lin} \bigcup_n \mathbb{R}(C_n)}$$

die Beziehung $Y = \mathbb{R} \oplus \mathbb{N}$.

<u>Beweis:</u> Die Behauptungen ergeben sich aus Lemma 6.3. Insbesondere sind die S_n Orthoprojektoren, so daß die Gleichung $Y = \mathbb{R} \oplus \mathbb{N}$ besteht (siehe Abschnitt 3).

Zunächst kann Satz 4.4 für das Verfahren (6.3) konkretisiert werden.

<u>Satz 6.10:</u> Es seien folgende Voraussetzungen erfüllt:

<u>a</u>) Die Folge (H_n) ist endlich erzeugt.

<u>b</u>) Mindestens einer der Operatoren S_n ist kompakt auf $\mathbb{N} = \overline{\operatorname{lin} \bigcup_n \mathbb{R}(C_n)}$.

Dann ist (x_n) G-konvergent. Weiterhin stellt der Grenzwert x_∞ stets eine bezüglich $(H_n^*A^*)$ verallgemeinerte Lösung von (6.1) dar.

Beweis: Mit (H_n) ist auch (S_n) endlich erzeugt. Außerdem erweist sich (S_n) in Verbindung mit Lemma 6.3, Lemma 6.9 und Voraussetzung b) als kompakte Projektionsfolge. Aufgrund von Lemma 6.3 gilt außerdem $\mathbb{N}(D_n) = \mathbb{N}(AD_n)$ $(\forall n)$ und damit $\bigcap_n \mathbb{N}(AD_n) = \bigcap_n \mathbb{N}(D_n)$. Schließlich liefert Satz 4.4 die Behauptung, wenn man Lemma 6.8 berücksichtigt.

Versucht man Folgerung 6.7 zu nutzen, um die Sätze 4.2 und 4.3 für das Iterationsverfahren (6.3) zu spezialisieren, muß man für das gleiche Ergebnis stärkere Voraussetzungen als in Satz 6.10 fordern. Deshalb verzichten wir hier darauf.

Ist $X = R^N$, $Y = R^M$ und (H_n) eine zyklische Folge bestimmter Spaltenauswahlmatrizen, so entsteht das in /5/ untersuchte SPA-Verfahren (Spaltenapproximation). Dabei sind die Voraussetzungen des Satzes 6.10 von vornherein erfüllt. Dieser Satz verallgemeinert das Hauptergebnis der Arbeit /5/.

Literatur

/1/ Васильченко, Г.П., и Светлаков, А.А.: Проекционный алгоритм решения систем линейных алгебраических уравнений большой размерности. Ж. вычисл. мат. и мат. физ. 20, 3 - 10 (1980)

/2/ Berg, L.: General iteration methods for the evaluation of linear equations. Numer. Functional Anal. Optim. 1, 4, 365 - 381 (1979)

/3/ Kantorowitsch, L. W., und Akilow, G. P.: Funktionalanalysis in normierten Räumen. Berlin 1964

/4/ Maeß, G.: Iterative Lösung linearer Gleichungssysteme. Halle 1979

/5/ Maeß, G., und Peters, W.: Lösung inkonsistenter linearer Gleichungssysteme und Bestimmung einer Pseudoinversen für rechteckige Matrizen durch Spaltenapproximation. Z. Angew. Math. Mech. 58, 233 - 237 (1978)

/6/ Peters, W.: Lösung linearer Gleichungssysteme durch Projektion auf Schnitträume von Hyperebenen und Berechnung einer verallgemeinerten Inversen. Beiträge Numer. Math. 5, 129 - 146 (1976)

/7/ Schott, D.: Projektionskerne einer Operatorfolge. Beiträge Anal. (zum Druck eingereicht)

/8/ Schott, D.: Zur Konvergenz von Operatorfolgen im Zusammenhang mit Iterationsverfahren. Math. Nachr. (zum Druck eingereicht)

/9/ Tanabe, K.: Projection method for solving a singular system of linear equations and its applications. Numer. Math. 17, 203 - 214 (1971)

/10/ Taylor, A. E.: Introduction to Functional Analysis. New York 1958

eingegangen: 04. 07. 1980

Anschrift des Verfassers:

Dr. rer. nat. Dieter Schott
Wilhelm-Pieck-Universität Rostock
Sektion Mathematik
Universitätsplatz 1
DDR-2500 Rostock

Hinweise für Autoren

Manuskripte (in deutscher, ggf. auch in russischer oder englischer Sprache) bitten wir an die Schriftleitung zu schicken. Die gesamte Arbeit ist linksbündig zu schreiben. Eine Ausnahme hiervon bilden hervorzuhebende Formeln und das Literaturverzeichnis. Der Kopf der Arbeit soll folgende Form haben: Rostock. Math. Kolloq./ Leerzeile/ Vorname Name/ Leerzeile/ Titel der Arbeit/ 1 Zeilenumschaltung/ Unterstreichung/ Leerzeile. Der Text der Arbeit ist eineinhalbzeilig (= 3 Zeilenumschaltungen) zu schreiben mit maximal 63 Anschlägen je Zeile und maximal 37 Zeilen je Seite. Zwischenüberschriften sind wie folgt einzuordnen: 6 Zeilenumschaltungen/ Zwischenüberschrift/ Unterstreichung (ohne Zeilenumschaltung)/ 5 Zeilenumschaltungen. Hervorhebungen sind durch Unterstreichen und Sperren möglich. Ankündigungen wie Satz, Definition, Bemerkung, Beweis u. ä. sind zu unterstreichen und mit einem Doppelpunkt abzuschließen. Vor und nach Sätzen, Definitionen u. ä. ist ein Zeilenabstand von 5 Umschaltungen zu lassen. Fußnoten sind möglichst zu vermeiden. Sollte doch davon Gebrauch gemacht werden, so sind sie durch eine hochgestellte Ziffer im Text zu kennzeichnen und innerhalb des oben angegebenen Satzspiegels unten auf der gleichen Seite anzugeben. Formeln und Bezeichnungen sollen möglichst mit der Schreibmaschine zu schreiben sein. Hervorzuhebende Formeln sind drei Leerzeichen einzurücken und mit 6 Umschaltungen zum übrigen Text zu schreiben. Formelzähler sollen am rechten Rand stehen. Der Platz für Abbildungen ist beim Schreiben auszusparen; die Abbildungen selbst sind in der dem ausgesparten Platz entsprechenden Größe gesondert nach TGL-Vorschrift auf Transparentpapier beizufügen. Der zugehörige Begleittext ist im Manuskript mitzuschreiben. Sein Abstand nach unten beträgt 5 Umschaltungen. Literaturzitate im Text sind durch laufende Nummern in Schrägstrichen (vgl. /8/, /9/ und /10/) zu kennzeichnen und am Schluß der Arbeit unter der Zwischenüberschrift Literatur zusammenzustellen.

Beispiele:

/8/ Zariski, O., and Samuel, P.: Commutative Algebra. Princeton 1958

/9/ Steinitz, E.: Algebraische Theorie der Körper. J. Reine Angew. Math. 137, 167 - 309 (1920)

/10/ Gnedenko, B. W.: Über die Arbeiten von C. F. Gauß zur Wahrscheinlichkeitsrechnung. In: Reichard, H. (Ed.): C. F. Gauß, Gedenkband anläßlich des 100. Todestages. S. 193 - 204, Leipzig 1967

Die Angaben sollen in Originalsprache erfolgen; bei kyrillischen Buchstaben soll die bibliothekarische Transkription (Duden) verwendet werden.
Am Ende der Arbeit stehen folgende Angaben zum Autor und zur Arbeit: eingegangen: Datum/ Leerzeile/ Anschrift des Verfassers:/ Titel Initialen der Vornamen Name/ Institution/ Struktureinheit/ Straße Hausnummer/ Land Postleitzahl Ort.
Der Autor wird gebeten, eine Korrektur des Durchschlags vom Offsetmanuskript zu lesen und dabei die mathematischen Symbole einzutragen.

Ferner sollte er 1-2 Klassifizierungsnummern (entsprechend der "1980 Mathematics Subject Classification") zur inhaltlichen Einordnung seiner Arbeit angeben.