ISSN 0138-3248

ROSTOCKER

MATHEMATISCHES KOLLOQUIUM

Heft 45

UNIVERSITÄT ROSTOCK

In der Reihe **Rostocker Informatik-Berichte** sind bisher erschienen:

Heft 1	(1985):	20 Jahre Rechenzentrum/Sektion Informationsverarbeitung
Heft 2	(1985):	DIGRA '84 (Internationale wissenschaftliche Tagung 12.-16. November 1984, Ahrenshoop)
Heft 3	(1986):	Beiträge zur Digitalgraphik und ihre Anwendungen aus Institutionen und Kombinaten der DDR
Heft 4	(1986):	Arbeiten aus der Sektion Informatik
Heft 5	(1987):	Beiträge des Problemseminars „Graphisch-interaktive Systeme“
Heft 6	(1988):	„Graphische Standardisierung“ und „Simulation“
Heft 7	(1989):	DIGRA '88 (Internationale wissenschaftliche Tagung 7.-11. November 1988, Kühlungsborn)
Heft 8	(1989):	DIGRA '88
Heft 9	(1989):	DIGRA '88
Heft 10	(1990):	25 Jahre Rechenzentrum/Sektion Informatik
Heft 11	(1990):	Arbeiten des wissenschaftlichen Nachwuchses an der Sektion Informatik
Heft 12	(1991):	Arbeiten aus dem Fachbereich Informatik

Bezugsmöglichkeiten

Bestellungen an die Universität Rostock, Abt. Wissenschaftspublizistik, Vogelsang 13/14, O-2500 Rostock.

Ferner sind die Hefte im Rahmen des Schriftentausches über die Universität Rostock, Universitätsbibliothek, Tauschstelle, Universitätsplatz 5, Rostock, O-2500, zu beziehen.

Rostocker Mathematisches Kolloquium
Heft 45

Universität Rostock
Fachbereich Mathematik
1991

Wissenschaftliche Leitung: Prof. Dr. Hans-Wolfgang Stolle
(Sprecher des Fachbereichs Mathematik)
Dr. Werner Plischke

Redaktionelle Bearbeitung: Dr. Klaus-Dieter Drews
Herstellung der Druckvorlage: Dr. Andreas Straßburg

Universität Rostock
Fachbereich Mathematik
Universitätsplatz 1
O–2500 Rostock
Bundesrepublik Deutschland

Redaktionsschluß: 30. November 1991

Das **Rostocker Mathematische Kolloquium** erscheint dreimal im Jahr und ist im Rahmen des Schriftentausches über die Universität Rostock, Universitätsbibliothek, Tauschstelle, Universitätsplatz 5, O–2500 Rostock, Bundesrepublik Deutschland, zu beziehen.

Außerdem bestehen Bezugsmöglichkeiten für Bestellungen aus Deutschland und dem Ausland über die Universität Rostock, Abteilung Wissenschaftspublizistik, Vogelsang 13/14, O–2500 Rostock, Bundesrepublik Deutschland.

Zitat–Kurztitel: Rostock. Math. Kolloq. (1991) **45**

Universität Rostock
Abteilung Wissenschaftspublizistik
Vogelsang 13/14, O–2500 Rostock, Bundesrepublik Deutschland
Telefon 36 95 77
Druck: Ostsee-Zeitung, Verlag und Druck GmbH Rostock, BT Ribnitz
01000

https://doi.org/10.18453/rosdok_id00003875

Rostock. Math. Kolloq. 45, 3–8 (1991)

Subject Classification (AMS) 03G20, 08A40

DIETLINDE LAU

Ein Vollständigkeitskriterium für durch h–adisch elementare Relationen beschriebene maximale Klassen von P_k

Sei $E_k = \{0, 1, \ldots, k-1\}$, $k \geq 2$, P_k^n die Menge aller n–stelligen Funktionen über E_k, $P_k := \bigcup_{n\geq 1} P_k^n$ und $P_{k,l} := \bigcup_{n\geq 1}\{f^n \in P_k \mid f : E_k^n \longrightarrow E_l\}$ für $2 \leq l < k$. Mit $W(f)$ bezeichnen wir den Wertebereich von $f \in P_k$.

Als Operationen über P_k seien die üblichen Superpositionsoperationen $\zeta, \tau, \Delta, \nabla$ und $*$ zugelassen. Die Menge aller Funktionen, die aus Funktionen einer Menge $A \subseteq P_k$ mit Hilfe der Superpositionsoperationen in endlich vielen Schritten erzeugt werden können, heißt Abschluß von A und wird hier mit $[A]$ bezeichnet. Ist $A = [A]$, so nennen wir A abgeschlossene (Teil–) Menge bzw. Teilklasse von P_k. Eine echte Teilklasse A' von A nennt man maximale Klasse von A bzw. A–maximal, wenn für alle $f \in A \backslash A'$ stets $[A' \cup \{f\}] = A$ gilt. Eine Menge B ($\subseteq A$) heißt vollständig in A, wenn $[B] = A$ ist. Falls A eine gewisse endliche vollständige Teilmenge besitzt, so gibt es bekanntlich nur endlich viele A–maximale Klassen $A_1, A_2, \ldots, A_t$, und es gilt für beliebige $B \subseteq A$

$$[B] = A \Longleftrightarrow \forall i \in \{1, 2, \ldots, t\} : B \not\subseteq A_i. \tag{1}$$

Neben solchen Vollständigkeitskriterien kann man auch spezialisierte Kriterien (sogenannte verallgemeinerte Słupecki- Kriterien, siehe [7]) der Art

$$[B] = A \Longleftrightarrow A^1 \subseteq [B] \;\wedge\; (\forall i \in \{1, \ldots, r\} : B \not\subseteq A_i), \tag{2}$$

betrachten, wobei $A_1, \ldots, A_r$ diejenigen maximalen Klassen von A bezeichnen, die A^1 enthalten.

Bekanntlich liegt seit den Arbeiten von I.G. Rosenberg (siehe [8]) eine vollständige Beschreibung der maximalen Klassen von P_k und damit ein Kriterium der Form (1) für $A = P_k$ vor. Daran anschließend wurde von einer Reihe von Autoren versucht, für diejenigen maximalen Klassen, die endlich erzeugt sind, die in ihnen maximalen – die sogenannten submaximalen Klassen von P_k – zu ermitteln. Gelungen ist dies bisher vollständig nur für $k = 3$ sowie für einen der 6 möglichen Typen maximaler Klassen von P_k ([5], [10]). Daneben gibt es Ergebnisse für gewisse k (z.B. [1]) bzw. für spezielle Klassen ([6]). Für gewisse andere maximale Klassen (siehe [11]) kennt man bisher nur Kriterien der Art (2). Ein Beitrag zur letztgenannten Richtung soll hier für durch h–adisch elementare Relationen charakterisierbare maximale Klassen von P_k vorgestellt werden, womit ein Ergebnis aus [2] verallgemeinert wird. Wesentliches Hilfsmittel sind dabei Ergebnisse aus [4]. Wir übernehmen deshalb die dort angegebenen Begriffe und Bezeichnungen, die man im wesentlichen auch in [7] nachlesen kann. Insbesondere kennzeichnen wir nachfolgend mit $Pol_k\varrho$ (kurz $Pol\ \varrho$) die Menge aller Funktionen aus P_k, die die Relation ϱ bewahren, und bezeichnen die Relation

$$\left\{(a_0, a_1, \ldots, a_{h-1}) \in E_k^h \mid |\{a_0, \ldots, a_{h-1}\}| < h\right\}$$

mit ι_k^h.

Nachfolgend sei stets $k = h^m$, $h \geq 3$ und $m \geq 1$. Bezeichnet man mit $a^{(i)}$ die Ziffern der Zahl $a \in E_k$ im Zahlensystem mit der Grundzahl h, d.h.

$$a = a^{(m-1)} \cdot\ h^{m-1} + a^{(m-2)} \cdot\ h^{m-2} + \ldots + a^{(1)} \cdot\ h + a^{(0)} \quad \text{mit} \quad a^{(m-1)}, \ldots, a^{(0)} \in E_h,$$

so läßt sich eine h–adisch elementare Relation ξ_m ($\subseteq E_k^h$) wie folgt definieren:

$$(a_0, \ldots, a_{h-1}) \in \xi_m :\Longleftrightarrow \forall i \in E_m :\ (a_0^{(i)}, a_1^{(i)}, \ldots, a_{h-1}^{(i)}) \in \iota_h^h.$$

Außerdem können wir jede Funktion f^n aus P_k in der Form

$$f(\tilde{x}) = \sum_{i=0}^{m-1} f^{(i)}(\tilde{x}) \cdot h^i \tag{3}$$

$(f^{(0)}, \ldots, f^{(m-1)} \in P_{k,h})$ darstellen mit $f^{(i)}(\tilde{x}) := (f(\tilde{x}))^{(i)}$. In [4] wurde bewiesen

Lemma 1 *Eine Funktion $f^n \in P_k$ gehört genau dann zu $Pol\ \xi_m$, wenn für jedes $i \in E_m$ entweder $f^{(i)}$ höchstens $h-1$ verschiedene Werte annimmt oder sich eine Permutation s auf E_h, ein $j \in \{1,2,\ldots,n\}$ und ein $t \in E_m$ finden lassen, so daß*

$$f^{(i)}(x_1,\ldots,x_n) = s(x_j^{(t)}) \tag{4}$$

gilt.

Mit Hilfe der eben beschriebenen Darstellung der Funktionen von $Pol\ \xi_m$ lassen sich nun drei Teilmengen

$$A_1 := \bigcup_{n\geq 1} A_1^n, \quad A_2 := \bigcup_{n\geq 1} A_2^n \quad \text{und} \quad B := \bigcup_{n\geq 1} B^n$$

von $Pol\ \xi_m$ wie folgt definieren:

A_1^n sei die Menge aller Funktionen f der Form (3) aus $Pol\ \xi_m$ mit der Eigenschaft, daß für jedes $i \in E_m$ die Funktion $f^{(i)}$ entweder wesentlich von nur einer Variablen abhängt oder höchstens $h-2$ verschiedene Werte annimmt.

Mit A_2^n sei die Menge aller Funktionen der Form (3) aus $Pol\ \xi_m$ bezeichnet, deren Komponenten $f^{(i)}$ für beliebiges $i \in E_m$ entweder wesentlich von nur einer Variablen abhängen oder die quasilinear (d.h. von der Form

$$f^{(i)}(\tilde{x}) = g_0(g_1(x_1) + \ldots + g_n(x_n)\ mod\ 2)$$

für gewisse $g_0,\ldots,g_n \in P_{k,2}^1$) sind.

B^n sei die Menge aller Funktionen f der Form (3) aus $Pol\ \xi_m$, so daß sich unter den Funktionen $f^{(0)},\ldots,f^{(m-1)}$ keine zwei der Form (4) befinden, die von zwei verschiedenen Variablen wesentlich abhängen.

Mit Hilfe von Lemma 1 prüft man leicht nach, daß die Mengen A_1, A_2 und B abgeschlossen sind und $(Pol\ \xi_m)^1$ enthalten.

Lemma 2 *Sei $f^n \in Pol\ \xi_m \backslash B$ sowie $g^q \in Pol\ \xi_m \backslash A$ mit $A = A_1$ für $h \geq 4$ und $A = A_2$ für $h = 3$. Dann gilt*

$$[\{f,g\} \cup (Pol\ \xi_m)^1] = Pol\ \xi_m\ .$$

Beweis: Wenn $f^n \in Pol\ \xi_m$ nicht zu B gehört, können wir o.B.d.A. $f^{(i)}(\tilde{x}) = s_1(x_1^{(t_1)})$ und $f^{(j)}(\tilde{x}) = s_2(x_2^{(t_2)})$ $(i,j \in \{1,2,\ldots,n\},\ i \neq j,\ s_1,\ s_2$ *Permutationen auf* $E_h,\ t_1,\ t_2 \in E_m)$ annehmen. Zu $(Pol\ \xi_m)^1$ gehören für beliebige $u_i,\ r_1,\ r_2$ aus E_m die Funktionen h und q_i $(i \in \{1,2\})$ mit

$$h(x) := x^{(i)}h^{r_1} + x^{(j)}h^{r_2} \quad und \quad q_i(x) := (s_i^{-1}(x^{(u_i)}))h^{t_i}.$$

Folglich gilt

$$\begin{aligned} h_{u_1r_1;u_2r_2}(x,y) &:= h(f(q_1(x), q_2(y), y, \ldots, y)) \\ &= x^{(u_1)}h^{r_1} + y^{(u_2)}h^{r_2} \in [\{f,g\} \cup (Pol\ \xi_m)^1] \end{aligned}$$

für beliebige $u_1,\ u_2,\ r_1,\ r_2$ aus E_m. Wenn die Funktion g aus $Pol\ \xi_m$ nicht in A liegt, gibt es ein $i \in E_m$ mit (o.B.d.A.)

$$W(g^{(i)}) = E_{h-1} \quad und \quad g^{(i)} \notin [P_k^1] \quad für \quad h \geq 4$$

bzw. $g^{(i)}$ nicht quasilinear und $g^{(i)} \notin [P_k^1]$ für $h = 3$. Mit Hilfe der Vollständigkeitskriterien für P_k und $P_{k,l}$ ([8, 3]) überlegt man sich dann, daß als Superpositionen über

$$\left\{g^{(i)}\right\} \cup \left\{t \in P_k^1 \mid |W(t)| \leq h-1\right\} \tag{5}$$

sämtliche Funktionen aus $P_{k,h-1}$ zu erhalten sind. Die Funktionenmenge (5) gehört wegen $q(x) := x^{(i)} \in (Pol\ \xi_m)^1$, $q * g = g^{(i)}$ und $\iota_k^h \subseteq \xi_m$ zu $[\{f,g\} \cup (Pol\ \xi_m)^1]$.
Zusammenfassend gilt also

$$(Pol\ \xi_m)^1 \cup P_{k,h-1} \cup \{h_{u_1r_1;u_2r_2} \mid u_1,\ u_2,\ r_1,\ r_2 \in E_m\} \subseteq [(Pol\ \xi_m)^1 \cup \{f,g\}],$$

womit das dem Beweis von Satz 4.2 aus [4] zu entnehmende Erzeugendensystem für $Pol\ \xi_m$ in $[(Pol\ \xi_m)^1 \cup \{f,g\}]$ nachgewiesen ist. ∎

Da offenbar $A \nsubseteq B$ und $B \nsubseteq A$ gilt ($A = A_1$ für $h \geq 4$, $A = A_2$ für $h = 3$), folgt aus dem eben bewiesenen Lemma 2 der

Satz *Die Mengen*

$$A = \begin{cases} A_1 & f\ddot{u}r \;\; h \geq 4, \\ A_2 & f\ddot{u}r \;\; h = 3 \end{cases}$$

und B sind die einzigen maximalen Klassen von Pol ξ_m, die (Pol ξ_m)^1 enthalten.

Eine beliebige Teilmenge M von Pol ξ_m ist genau dann in Pol ξ_m vollständig, wenn M keine Teilmenge von A und B ist sowie im Abschluß (Pol ξ_m)^1 enthält.

Literatur

[1] **Bagyinszki, J.**, and **Demetrovics, J.**: *The lattice of linear classes in prime-valued logics.* Banach Center Publ. **7**, 105–123 (1982)

[2] **Burle, G.A.**: *Die Klassen der k-wertigen Logik, die alle Funktionen einer Veränderlichen enthalten.* (Russisch) Diskretnyj Analiz **10**, 3–7 (1967)

[3] **Lau, D.**: *Prävollständige Klassen von $P_{(k,l)}$.* Elektron. Informationsverarb. Kybernet. **11**, 624–626 (1975)

[4] **Lau, D.**: *Bestimmung der Ordnung maximaler Klassen von Funktionen der k-wertigen Logik.* Z. Math. Logik Grundlag. Math. **24**, 79–96 (1978)

[5] **Lau, D.**: *Submaximale Klassen von P_3.* Elektron. Informationsverarb. Kybernet. **18**, 227–243 (1982)

[6] **Lau, D.**: *Die maximalen Klassen von Pol_k(0).* Rostock. Math. Kolloq. **19**, 29–47 (1982)

[7] **Pöschel, R.**, und **Kalužnin, L.A.**: *Funktionen- und Relationenalgebren.* Berlin 1979

[8] **Rosenberg, I.G.**: *Über die funktionale Vollständigkeit in den mehrwertigen Logiken.* Rozpravy Čechoslovenské Akad. Věd Řada Mat. Přirod. Véd **80**, 3–93 (1970)

[9] **Rosenberg, I.G.**: *Completeness, closed classes and relations in multiple-valued logics.* In: Proc. Internat. Sympos. on multiple-valued Logics. Morgantown 1974

[10] **Rosenberg, I.G.**, and **Szendrei, Á.**: *Submaximal clones with a prime order automorphism.* Acta Sci. Math. (Szeged) **49**, 29–48 (1985)

[11] **Szabo, L.**, and **Szendrei, Á.**: *Slupecki-type criteria for quasilinear functions over a finite dimensional vector space.* Elektron. Informationsverarb. Kybernet. **17**, 601–611 (1961)

eingegangen: 12. September 1989

Verfasser:
Dr. D. Lau
Universität Rostock
Fachbereich Mathematik
Universitätsplatz 1
O–2500 Rostock
Germany

Rostock. Math. Kolloq. 45, 9–20 (1991)

Subject Classification (AMS)
05A20, 52A40

AHMED MOHAMMED; GUNTER TEUMER

On k-equivolume sets in $\mathbb{R}^d$

1 Introduction

A subset M of the Euclidean vector space $\mathbb{R}^d$ is called a k-equivolume set, if the volume of the parallelepiped spanned by any k vectors of M does not depend on the choice of these vectors. The question is to find an upper bound for the cardinality of M. The problem is a generalization of the equidistance sets considered in [1] and [2]. In [4] the case $k = 2$ is dealt with. All bounds obtained in the present paper are met at least for $k = 2$. The corresponding examples can be found in [4]. More examples for equidistance sets, i.e. where all vectors in M are of equal length and k is two, are listet up in [3]. Besides giving a general upper bound, the goal of the present paper is to consider closely the case, if amongst the parallelepipeds associated with M there is a rectangular one. It turns out that in this case the upper bound can be sharpened for certain values of the volume S in question and of the dimension d. Particularly, if the set is large there are only a finite number of values met by S.

2 The general bound

Definition 1 *A finite set M of vectors in $\mathbb{R}^d$ is called a* k-equivolume set *if the volume of each k-dimensional parallelepiped spanned by any k vectors of M is always the same.*

In the present paper k is supposed to be larger than one and the volume in question is denoted by S. Obviously, M is a k-equivolume set if and only if any principal minor of order k of the *Gram*-matrix of M is equal to S^2. Without loss of generality M can be considered to be embedded in the unit ball of $\mathbb{R}^d$ in such a way that at least one vector is a unit one. If $x_1, \dots, x_n$ are arbitrary vectors in $\mathbb{R}^d$, then $G(x_1, \dots, x_n)$ will denote the determinant of their *Gram*-matrix or the *Gram*-matrix itself. The sense in which it is used will be clear from the context.

Proposition 2 (Generalized *Hadamard's* inequality)

$$G(x_1,\dots,x_n) \le G(x_1,\dots,x_m)G(x_{m+1},\dots,x_n).$$

Equality holds iff either one of the determinants vanishes or the sets $\{x_1,\dots,x_m\}$ *and* $\{x_{m+1},\dots,x_n\}$ *are orthogonal to each other.* ■

Theorem 3 *The cardinality* v *of a* k*-equivolume set* M *does not exceed*

$$\binom{d(k-1)+2k-3}{2(k-1)}.$$

Proof. Define a system of homogenous polynomials of degree $2(k-1)$ in $d(k-1)$ variables by $F_{m_i}(x_1,\dots,x_{k-1}) := G(m_i,x_1,\dots,x_{k-1})$ for $i=1,\dots,v$.

The theorem is proved if this system is shown to be linearly independent in the vector space of the homogenous polynomials of degree $2(k-1)$ in $d(k-1)$ variables.

Therefore consider $\sum_{i=1}^{v}\alpha_i F_{m_i} \equiv 0$. Evaluating the right hand side of the equation at the point $P_1=(m_1,\dots,m_{k-1})$ yields

$$\sum_{i=k}^{v}\alpha_i = 0. \tag{1}$$

If $P_h=(m_1\dots,m_{h-1},m_{h+1},\dots,m_k)$ for each $h=1,\dots,k$, then $\sum_{i=1}^{v}\alpha_i F_{m_i}(P_h)=0$ implies

$$\alpha_h + \sum_{i=k+1}^{v}\alpha_i = 0. \tag{2}$$

Consequently, $\alpha_h = \alpha_k$.

Finally, if $P_q=(m_1,\dots,m_{k-1},m_q)$ for $q=k+1,\dots,v$, then again $\sum_{i=1}^{v}\alpha_i F_{m_i}(P_q)=0$ yields $\sum_{i=k-1}^{v}\alpha_i-\alpha_q=0$ and this together with (2) gives: $\alpha_{k-1}+\alpha_k-\alpha_h=\alpha_q$ from which we conclude $\alpha_q=\alpha_k$.

But then (1) leads to $a_k=0$.

Hence the set $\{F_{m_i}\}_{i=1}^{v}$ is linearly independent. On the other hand, it is well known that the dimension of the vector space of the homogenous polynomials of degree $2(k-1)$ in $d(k-1)$ variables is equal to $\binom{d(k-1)+2k-3}{2(k-1)}$. ■

3 Rectangular k-equivolume sets

Now let M be a k-eqivolume set that contains a rectangular parallelepiped with a side of unit length among the k-dimensional parallelepipeds spanned by the vectors of M. In the sequel, such k-equivolume sets will be reffered to as rectangular k-equivolume sets.

Lemma 4 *Any rectangular k-equivolume set M with cardinality v contains a subset with $v-k+1$ unit vectors.*

Proof. Assume that m_1 is a unit vector, i.e. $G(m_1) = 1$, in M which is a side of a rectangular parallelepiped. Consequently, there are k mutually orthogonal vectors $m_1, m_2, \ldots, m_k$ such that $G(m_1, \ldots, m_k) = S^2$. Consider the set $M' := M \backslash \{m_2, \ldots, m_k\}$ and take $m \in M'$. Then by definition

$$S^2 = G(m_2, \ldots, m_k, m) \leq G(m_2) \ldots G(m_k) G(m) = G(m_1) G(m_2) \ldots G(m_k) G(m) = S^2 G(m).$$

Hence $G(m) \geq 1$. Since all the vectors in M are of length at most one, it follows that $G(m) = 1$. Consequently all vectors in M' are unit vectors. Moreover M' is orthogonal to $\{m_2, \ldots, m_k\}$. Indeed, from the inequalities above it follows that $G(m_2, \ldots, m_k, m) = G(m_1) G(m_2) \ldots G(m_k) G(m)$ and thus as a consequence of Proposition 2, we obtain the desired result. ■

We remark that Lemma 4, together with the observation noted in the last part of its proof allows to regard M' as being embedded in the $(d-k)$ dimensional sphere of $\mathbb{R}^{d-k+1}$.

Lemma 5 *M', as defined in the proof of the lemma above, is an equidistance set and each vector in $M \backslash M'$ is of length $S^{1/k-1}$.*

Proof. If M' contains exactly one element, then there is nothing to prove and so assume that M' contains at least two elements, say m_i and m_j. Then it holds that

$$\begin{aligned} S^2 = G(m_2, \ldots, m_{k-1}, m_i, m_j) &= G(m_2, \ldots, m_{k-1}) G(m_i, m_j) \\ &= G(m_1) G(m_2, \ldots, m_{k-1}) G(m_i, m_j) \\ &= G(m_1, \ldots, m_{k-1}) G(m_i, m_j). \end{aligned}$$

Hence,

$$G(m_i, m_j) = \frac{S^2}{G(m_1, \ldots, m_{k-1})} = \|m_k\|^2.$$

Similarly, $G(m_i, m_j) = \|m_h\|^2$ for $h = 2, \ldots, k$.

Consequently, $S^2 = G(m_1, \ldots, m_k) = \|m_k\|^{2(k-1)}$.

Since m_i and m_j are unit vectors, $G(m_i, m_j) = 1 - < m_i, m_j >^2 = S^{2/(k-1)}$.

Hence $| < m_i, m_j > | = (1 - S^{2/(k-1)})^{1/2}$. Thus M' is an equiangular set. ■

Theorem 6 *The cardinality of any rectangular k-equivolume set does not exceed $\binom{d-k+1}{2}+k-1$. This bound is met for at least $k=2$.*

Proof. Consider the system of homogenous polynomials of degree 2 in $d-k+1$ variables defined by:

$$F_{m_i}(x) = < m_i, x >^2 + (1 - S^{2/(k-1)}) < x, x > \qquad \text{where } m_i \in M'.$$

Then

$$F_{m_i}(m_j) = \begin{cases} 0 & \text{for } i \neq j, \\ S^{2/(k-1)} & \text{for } i = j. \end{cases}$$

Consequently, the system $\{F_{m_i}\}_{i=1}^{v-k+1}$, where v is the cardinality of M, is a linearly independent set. Hence we have

$$v \leq \binom{d-k+1}{2} + k - 1.$$

■

For large rectangular k-equivolume sets, the value of the volume can not be chosen arbitrarily.

Theorem 7 *If* $\quad S \geq \left(1 - \frac{1}{d-k+3}\right)^{(k-1)/2}$, *then*

$$\text{Card}(M) \leq k + \frac{d-k}{(d-k+1)\left[S^{2/(k-1)} - \frac{d-k}{d-k+1}\right]}.$$

Proof. Let M be a rectangular k-equivolume set. Denote $(1-S^{2/(k-1)})^{1/2}$ by A. The entries in the diagonal of the *Gram*-matrix $G(M')$ are all ones and the off-diagonal entries are $\pm A$. Consider the matrix $B = \frac{1}{A}(G(M') - E)$, where E is the unit matrix. The diagonal entries of B are all zeros and the off-diagonal entries are all ± 1. Let the spectrum of $G(M')$ consist of λ_i, $i = 1, \ldots, v-k+1$, where

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > \lambda_{r+1} = \ldots = \lambda_{v-k+1} = 0$$

with $r = \text{rank}(G(M')) \leq d-k+1$. Consequently, the spectrum of B consists of $s_i = (\lambda_i - 1)/A$.

Now

$$0 = \mathrm{Tr}(B) = \sum_{i=1}^{d-k+1} s_i + \sum_{i=d-k+1}^{v-k+1} s_i = \sum_{i=1}^{d-k+1} s_i - (v-d)/A.$$

Therefore,

$$\sum_{i=1}^{d-k+1} s_i = (v-d)/A.$$

On the other hand:

$$(v-k+1)^2 - (v-k+1) = \mathrm{Tr}(B^2) = \sum_{i=1}^{d-k+1} s_i^2 + \sum_{i=d-k+2}^{v-k+1} s_i^2.$$

Consequently,

$$(v-k+1)(v-k) - (v-d)/A^2 = \sum_{i=1}^{d-k+1} s_i^2 \geq \frac{1}{d-k+1}\left(\sum_{i=1}^{d-k+1} |s_i|\right)^2 \geq \frac{1}{d-k+1}\left(\frac{v-d}{A}\right)^2.$$

Thus,

$$(v-k+1)(v-k)A^2 \geq \frac{(v-d)}{d-k+1}(d-k+1+v-d) = \frac{(v-d)(v-k+1)}{d-k+1}.$$

From this follows:

$$\text{If } S \geq \left(1 - \frac{1}{d-k+1}\right)^{(k-1)/2}, \text{ then } v \leq \frac{(1-S^{2/(k-1)})k(d-k+1)-d}{(d-k)(1-S^{2/(k-1)}) - S^{1/(k-1)}}$$
$$= k + \frac{d-k}{(d-k+1)\left[S^{2/(k-1)} - \dfrac{d-k}{d-k+1}\right]}.$$

This bound is not larger than the general bound obtained in Theorem 6, if

$$S \geq \left(1 - \frac{1}{d-k+3}\right)^{(k-1)/2}.$$

■

Now suppose that $k \leq d-3$.

Theorem 8 *If the cardinality of a rectangular k-equivolume set is larger than $2d-k+1$, then $(1-S^{2/(k-1)})^{-1/2}$ is an odd integer.*

Proof. Consider again the matrix $B = \frac{1}{A}(G(M') - E)$. Then $-1/A$ is an eigenvalue of B. Furthermore the characteristic polynomial of B is a polynomial of degree $v-k+1$, and its coefficients are rational numbers (in fact, they are integers). Hence $-1/A$ is an algebraic number. Assume that b is an algebraic number conjugate to $-1/A$. The multiplicity of the eigenvalue $-1/A$ is equal to $v-k+1 = \text{rank } G(M')$ which is not smaller than $v-k+1-(d-k+1) = v-d$. Hence, denoting the multiplicity of $-1/A$ by m, it holds $m \geq v-d$. Clearly the multiplicity of b is also m. If b and $-1/A$ were distinct then the degree of the characteristic polynomial of B will not be smaller than $2m$.
That is $v-k+1 \geq 2m \geq 2(v-d) \geq v-k+2$ and this is an obvious contradiction thereby forcing b and $-1/A$ to be the same. Therefore $-1/A$ is a rational number. Owing to the fact that the leading coefficient of the characteristic polynomial is 1 and all its coefficients are integers, any rational root is an integer. Consequently $1/A = (1-S^{2/(k-1)})^{-1/2}$ is an integer. Because in general $v \leq \binom{d-k+2}{2} + k - 1$, the condition $v \geq 2d-k+2$ can hold only if $k \leq d-3$.
In order to prove that $1/A$ is an odd integer, we consider the matrix $Q = \frac{1}{2}(B+E-J)$ where J is the matrix with all its entries equal to 1. Let us denote $1/A$ by P. The dimension of the kernel of J is equal to the order of Q minus 1, i.e. $v-k$. The multiplicity of P with respect to matrix B is not smaller than $v-d$. Thus, corresponding to the eigenvalue P there is an eigenvector x of B which belongs to the kernel of J. Consequently, $Qx = \frac{P+1}{2}x$ and thus, $\frac{P+1}{2}$ is an eigenvalue of Q. Since all the entries of Q are integers, any rational eigenvalue of Q is an integer. Therefore $P = 1/A$ is an odd integer as claimed. ∎

If $(1-S^{2/(k-1)})^{-1/2}$ is an integer, say m, then $S^2 = (1-1/m^2)^{k-1}$ with $m \geq 2$. The range of the parameter m is not arbitrary.

Theorem 9 *If the cardinality of a rectangular k-equivolume set exceeds $2d-k+1$, then $S^2 = (1-1/m^2)^{k-1}$, where $9 \leq m^2 \leq \dfrac{2(d-k+1)^2}{d-k+2}$ and m is odd.*

Proof. Let $u = v-(k-1) > 2(d-k+1)$. This u is the cardinality of the underlying equidistance set M'. $G(M')$ is of order u and its rank does not exceed $d-k+1$.
Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{d-k+1} \geq \lambda_{d-k+2} = \ldots = \lambda_u = 0$ be the spectrum of $G(M')$. It follows that

$$\sum_{i=1}^{d-k+1} \lambda_i = u.$$

Consequently,

$$u^2 = \left(\sum_{i=1}^{d-k+1} \lambda_i\right)^2 = \sum_{i=1}^{d-k+1} \lambda_i^2 + 2\sum_{i>j} \lambda_i\lambda_j.$$

Because M' is an equidistance set it holds that

$$2\sum_{i>j}\lambda_i\lambda_j = \binom{u}{2} S^{2/(k-1)}.$$

Hence,

$$\begin{aligned} u \geq u(u-1)S^{2/(k-1)} + \sum_{i=1}^{d-k+1}\lambda_i^2 &\geq u(u-1)S^{2/(k-1)} + \frac{1}{d-k+1}\left(\sum_{i=1}^{d-k+1}\lambda_i\right)^2 \\ &= u(u-1)S^{2/(k-1)} + \frac{u^2}{d-k+1}. \end{aligned}$$

Therefore, $u\left(1 - S^{2/(k-1)} - \frac{1}{d-k+1}\right) \geq -S^{2/(k-1)}$.

Since the hypotheses of Theorem 8 are fulfilled, $S^{2/(k-1)} = 1 - 1/m^2$ for some odd integer $m \geq 3$.

Thus, $u\left(1/m^2 - \frac{1}{d-k+1}\right) \geq 1/m^2 - 1$ or $(u-1)/m^2 \geq \frac{u}{d-k+1} - 1 > 1$.

Finally

$$m^2 \leq \frac{(u-1)(d-k+1)}{u-d+k-1} \leq \frac{2(d-k+1)^2}{d-k+2}$$

as desired. ∎

The rectangular k-equivolume sets of maximal cardinality are unique sets.

Theorem 10 *Let M be a rectangular k-equivolume set, and v be the cardinality of M. If $v = \binom{d-k+1}{2} + k - 1$, then $d-k+3$ is an odd square integer and it holds*

$$S^2 = \left(1 - \frac{1}{d-k+3}\right)^{k-1}.$$

Proof. Consider the system of polynomials defined by:

$$F_{m_i}(x) = < m_i, x >^2 \quad \text{for } i = 1, \ldots, v-k+1 \quad \text{and} \quad F_0(x) = (1 - S^{2/(k-1)}) < x, x >$$

where $m_i \in M'$ and $x \in \mathbb{R}^{d-k+1}$.

These are polynomials of degree two in $d-k+1$ variables. If the system is linearly independent it will hold

$$v - k + 1 < \binom{d-k+1}{2}$$

contradicting the maximality of v stated in the hypothesis of the theorem. Consequently, the system must be dependent. Consider

$$\sum_{i=1}^{v-k+1} \alpha_i F_{m_i}(x) + \beta F_0(x) \equiv 0.$$

Putting in m_j yields

$$(1 - S^{2/(k-1)}) \sum_{i=1}^{v-k+1} \alpha_i + \beta(1 - S^{2/(k-1)}) + \alpha_j S^{2/(k-1)} = 0.$$

Hence, the α_j's are equal to the same constant, say C. Consequently,

$$\beta = -\frac{C[(v-k+1)(1 - S^{2/(k-1)}) + S^{2/(k-1)}]}{1 - S^{2/(k-1)}}.$$

Using these values in the above linear combination we obtain

$$C\left(\sum_{i=1}^{v-k+1} < m_i, x >^2 - [(v-k+1)(1 - S^{2/(k-1)}) + S^{2/(k-1)}] < x, x >\right) \equiv 0.$$

Taking for x the vectors of an orthonormal basis, and denoting by $m_{i,k}$ the k-th coordinate of m_i with respect to this orthonormal basis in the above equation, we get

$$C\left(\sum_{i=1}^{v-k+1} m_{i,k}^2 - [(v-k+1)(1 - S^{2/(k-1)}) + S^{2/(k-1)}]\right) = 0$$

and summing this over k gives

$$C\left((v-k+1) - (d-k+1)[(v-k+1)(1 - S^{2/(k-1)}) + S^{2/(k-1)}]\right) = 0.$$

C must be different from zero since the set $\{F_{m_i}, F_0\}$ is linearly dependent.

Hence

$$(v-k+1) - (d-k+1)[(v-k+1)(1 - S^{2/(k-1)}) + S^{2/(k-1)}] = 0$$

or

$$(v-k+1)[(k-d) + (d-k+1)S^{2/(k-1)}] = (d-k+1)S^{2/(k-1)}.$$

Consequently, the last formula together with the fact that

$$v = \binom{d-k+2}{2} + (k-1)$$

leads to

$$S^{2/(k-1)} = \frac{d-k+2}{d-k+3}.$$

Therefore, we have $S^2 = \left(1 - \frac{1}{d-k+3}\right)^{k-1}$. Furthermore, in accordance with Theorem 8, $d-k+3 = m^2$ for some odd integer m. ■

Finally, we consider k-equivolume sets whose vectors are spread over the $(d-1)$-dimensional unit sphere, i.e. whose vectors are unit vectors.

Theorem 11 *If all vectors of a k-equivolume set M are of equal length, then its cardinality does not exceed*

$$\binom{d(k-1)+2k-3}{2(k-1)} - (k-1) \quad \text{for} \quad k \geq 3.$$

Proof. Consider the system of homogenous polynomials of degree $2(k-1)$ in $d(k-1)$ variables, defined by

$$F_{m_i}(x_1,\ldots,x_{k-1}) = G(m_i,x_1,\ldots,x_{k-1}) \quad \text{and} \quad Q_j(x_1,\ldots,x_{k-1}) = < x_j, x_j >^{k-1}$$

for $i = 1,\ldots,v$ and $j = 1,\ldots,k-1$, where $v = \mathrm{Card}(M)$, and $k \geq 3$.

We maintain that the above system is linearly independent. To this end, consider the linear combination

$$\sum_{i=1}^{v} \alpha_i F_{m_i}(x_1,\ldots,x_{k-1}) + \sum_{i=1}^{k-1} \beta_i Q_i(x_1,\ldots,x_{k-1}) \equiv 0.$$

Evaluating this at $(m_1,m_2,\ldots,m_{k-2},m_1)$ gives $\sum_{i=1}^{k-1} \beta_i = 0$.

At $(m_1,m_2,\ldots,m_{k-1})$, we get

$$S^2 \sum_{i=k}^{v} \alpha_i + \sum_{i=1}^{k-1} \beta_i = 0$$

and hence

$$\sum_{i=k}^{v} \alpha_i = 0. \tag{3}$$

At $(m_1,m_2,\ldots,m_{k-2},m_h)$, where $h \geq k$, we obtain

$$\sum_{i=k}^{v} \alpha_i - \alpha_h + \alpha_{k-1} = 0.$$

Consequently, $\alpha_h = \alpha_{k-1}$ for all $h \geq k$.

Finally, evaluation at the point $(m_1,\ldots,m_{h-1},m_{h+1},\ldots,m_k)$ for $h \leq k-1$ yields

$$\sum_{i=k+1}^{v} \alpha_i + \alpha_h = 0.$$

But from (3) follows $\sum_{i=k+1}^{v} \alpha_i = -\alpha_k = -\alpha_{k-1}$.

Therefore, $\alpha_h = \alpha_{k-1}$ for $h \leq k-1$.

Now we are left with the sum

$$\sum_{i=1}^{k-1} \beta_i Q_i(x_1, \ldots, x_{k-1}) \equiv 0.$$

But this implies $\beta_i = 0$ for all $i = 1, \ldots, k-1$ (take $x_i = 0$ for $i \neq j$ and $\|x_j\| = 1$). Thus

$$v + k - 1 \leq \binom{d(k-1) + 2k - 3}{2(k-1)}.$$

■

The case $k = 2$ is considered in [4]. In this case, M is an equidistance set and the results are obtained similarly in the case of rectangular k-equivolume sets.

Proposition 12 *If M is a 2-equidistance set such that all its vectors are of the same length, then $Card(M) = v \leq \binom{d+1}{2}$. Moreover:*

(i) *If $v = \binom{d+1}{2}$, then $S^2 = (d+1)(d+2)^{-1}$.*

(ii) *If $v = \binom{d+1}{2}$ and $d \geq 5$, then $d+2$ is a square integer.*

(iii) *If $v = 2d+1$ and $d \geq 5$, then $S^2 = 1 - m^{-2}$, where m is an integer with $4 \leq m^2 \leq 2d-1$.* (c.f. [4]).

Acknowledgement. The camera-ready document was produced on the French Desktop Publishing Station at the Computer center of the Department of Mathematics at Addis Abeba University. The authors would like to express their gratitude to the Computer center for getting this privilege.

References

[1] **Blokhuis, A.**, and **Seidel, J.J.**: *Few-distance sets in $\mathbb{R}^{p,q}$.* In: Combinatorica. Convegno Roma, 23–26 maggio 1983. Sympos. Math. **28**, 145–158, New York 1986

[2] **Blokhuis, A.**: *Few-distance sets.* Thesis, CWI Tracts 7 Amsterdam, Mathematisch Centrum 1984

[3] **Seidel, J.J.**: *The pentagon.* In: Baayen, P.C. (Ed.): Proceedings bicentennial congress wiskundig Genootschap, part 1. Mathematical Centre Tracts **100**, 81–96, Amsterdam, Mathematisch Centrum 1979

[4] **Teumer, G.**: *Equiareal sets in $\mathbb{R}^d$.* (to appear in Discrete Comput. Geom.)

eingegangen: 2. Januar 1990

Authors:

Dr. A. Mohammed
Addis-Abeba-University
P.O. Box 1176
Dpt. Mathematics
Addis Abeba
Ethiopia

Dr. G. Teumer
Ernst-Moritz-Arndt-Universität Greifswald
Fachbereich Mathematik
F.-L.-Jahn-Straße 15a
O–2200 Greifswald
Germany

Rostock. Math. Kolloq. 45, 21–24 (1991)

Subject Classification (AMS)
16A32, 16A66, 47C05

JÜRGEN SYNNATZSCHKE

Zur Erzeugung des Ideals der endlichdimensionalen Elemente einer Algebra durch zwei Unteralgebren eindimensionaler Elemente mit trivialem Quadrat

Ist X ein Banach-Raum mit $\dim X > 1$, so wird, wie W. Żelazko [3] kürzlich zeigte, die Algebra $\mathfrak{B}(X)$ all seiner stetigen linearen Operatoren in der starken Operatortopologie von zwei Unteralgebren mit jeweils trivialem Quadrat erzeugt. Bekanntlich enthält $\mathfrak{B}(X)$ aber als Teilmenge, die dicht in der starken Operatortopologie ist, das Ideal aller endlichdimensionalen Operatoren. Der Beweis des oben zitierten Faktes reduziert sich damit zum Beweis dessen, daß dieses Ideal von zwei Unteralgebren mit jeweils trivialem Quadrat erzeugt wird. Wir werden zeigen, daß dieser Fakt für eine recht allgemeine Klasse beliebiger Algebren $\mathfrak{A}$ gültig bleibt: In einer nichtkommutativen Algebra $\mathfrak{A}$ mit hinreichend vielen eindimensionalen Elementen wird das Ideal $\mathfrak{F}_e(\mathfrak{A})$ all ihrer endlichdimensionalen Elemente von zwei Unteralgebren eindimensionaler Elemente mit trivialem Quadrat erzeugt.

Wir erinnern daran, daß ein Element $w \neq 0$ einer Algebra $\mathfrak{A}$ *eindimensional* genannt wird, falls mit einem Funktional $g_w \neq 0$ die Beziehung $wzw = g_w(z)w$ für alle $z \in \mathfrak{A}$ gilt. Die Gesamtheit aller eindimensionalen Elemente von $\mathfrak{A}$ wird mit $\mathfrak{F}_1(\mathfrak{A})$ bezeichnet. Sie heißt *o-total*, falls für beliebige Elemente $y, z \in \mathfrak{A}$ mit $y, z \neq 0$ ein Element $w \in \mathfrak{F}_1(\mathfrak{A})$ mit $ywz \neq 0$ existiert. Im Falle $y = z$ ist das äquivalent dazu, daß $g_w(z) \neq 0$ für ein Element $w \in \mathfrak{F}_1(\mathfrak{A})$ ist (vgl. Definition 2.1 und Theorem 2.1 aus [1]). Mit $\mathfrak{F}_e(\mathfrak{A})$ wird die in $\mathfrak{A}$ von $\mathfrak{F}_1(\mathfrak{A})$ erzeugte Unteralgebra bezeichnet. Deren Elemente werden wir sinnvollerweise *endlichdimensional* nennen. Diese Unteralgebra ist, falls $\mathfrak{F}_1(\mathfrak{A})$ o-total ist, gleichzeitig das kleinste Ideal in $\mathfrak{A}$ (s. Theorem 1.4 und die sich anschließende Bemerkung aus [1]). Ist X ein Vektorraum mit totalem Dualraum und $\mathfrak{L}(X)$ die Algebra all seiner linearen Operatoren, so fällt in dieser der Begriff des eindimensionalen Elementes mit dem des eindimensionalen Operators zusammen, und $\mathfrak{F}_e(\mathfrak{L}(X))$ ist das gewöhnliche Ideal aller endlichdimensionalen Operatoren. Eindimensionale Elemente in beliebigen Algebren waren in [1] und [2] untersucht worden, und wir verweisen auf diese Arbeiten betreffs gewisser Fakten, die im folgenden noch benötigt werden sollten.

Lemma *Es sei $\mathfrak{A}$ eine Algebra mit o-totaler Menge eindimensionaler Elemente $\mathfrak{F}_1(\mathfrak{A})$. Es möge ein idempotentes Element $w \in \mathfrak{F}_1(\mathfrak{A})$ mit $g_w^{-1}(0)g_w^{-1}(0) \subset g_w^{-1}(0)$ existieren. Dann ist das Funktional g_w multiplikativ und die Algebra $\mathfrak{A}$ kommutativ.*

Beweis. Wegen der geforderten Idempotentheit des Elementes $w \in \mathfrak{F}_1(\mathfrak{A})$ ist bekanntlich $g_w(wz) = g_w(z) = g_w(zw)$ für alle $z \in \mathfrak{A}$ (s. Theorem 1.1/(iv) aus [1]). Aus der Zerlegung $\mathfrak{A} = \mathbb{R}w \oplus g_w^{-1}(0)$ und der geforderten Inklusion $g_w^{-1}(0)g_w^{-1}(0) \subset g_w^{-1}(0)$ ergibt sich nun leicht, daß $g_w(xy) = g_w(yx)$ für alle $x, y \in \mathfrak{A}$ ist. Wendet man dies auf beliebige Elemente der Form $x = wb$ und $y = cw$ mit $b, c \in \mathfrak{A}$ an, so folgt, daß $g_w(bc) = g_w(wbcw) = g_w(cwb) = g_w(b)g_w(c)$ (s. Theorem 1.1/(v) aus [1]), d.h. daß g_w multiplikativ ist. Sei nun $u \in \mathfrak{F}_1(\mathfrak{A})$ beliebig. Jedes fixierte eindimensionale Element und insbesondere w kann alle übrigen eindimensionalen Elemente erzeugen, genauer ist $\mathfrak{F}_1(\mathfrak{A}) = \mathfrak{A}w\mathfrak{A}\backslash\{0\}$ (s. Lemma 3.1 aus [2]). Also gibt es Elemente $b, c \in \mathfrak{A}$ mit $u = bwc$. Aus der bereits gezeigten Multiplikativität des Funktionals g_w ergibt sich dann die Beziehung $g_u(xy) = g_w(cxyb) = g_w(cyxb) = g_u(yx)$ (s. Theorem 1.4/(iii) aus [1]). Wegen der o-Totalität von $\mathfrak{F}_1(\mathfrak{A})$ ist das nur bei $xy = yx$ möglich. ∎

Theorem *Es sei $\mathfrak{A}$ eine nichtkommutative Algebra mit o-totaler Menge $\mathfrak{F}_1(\mathfrak{A})$ und $w \in \mathfrak{F}_1(\mathfrak{A})$ ein idempotentes Element. Dann sind $wg_w^{-1}(0)$ und $g_w^{-1}(0)w$ Unteralgebren eindimensionaler Elemente mit trivialem Quadrat, die zusammen das Ideal $\mathfrak{F}_e(\mathfrak{A})$ aller eindimensionalen Elemente von $\mathfrak{A}$ erzeugen.*

Beweis. Offensichtlich sind $wg_w^{-1}(0)$ und $g_w^{-1}(0)w$ Unteralgebren mit trivialem Quadrat. Daß jedes ihrer nichttrivialen Elemente eindimensional ist, ergibt sich leicht aus der o-Totalität von $\mathfrak{F}_1(\mathfrak{A})$ (s.a. Theorem 1.4/(ii) aus [1]). Es genügt nun der Nachweis dessen, daß die von $wg_w^{-1}(0)$ und $g_w^{-1}(0)w$ in $\mathfrak{A}$ erzeugte Unteralgebra jedes Element $w' \in \mathfrak{F}_1(\mathfrak{A})$ enthält. Dazu werden wir getrennt die Fälle $g_w(w') = 0$ und $g_w(w') \neq 0$ betrachten. Der erste von ihnen läßt sich weiter in die drei Fälle $ww' = 0 = w'w$, $ww' = 0$ und $0 \neq w'w$, sowie $ww' \neq 0$ und $0 = w'w$ aufspalten.

1. *Der Fall $ww' = 0 = w'w$:* Es ist dann offensichtlich $g_w(w'z) = 0 = g_w(zw')$ für alle $z \in \mathfrak{A}$. Da $\mathfrak{F}_1(\mathfrak{A})$ o-total und gleich $\mathfrak{A}w\mathfrak{A}\backslash\{0\}$ ist, lassen sich Elemente $b, c \in \mathfrak{A}$ mit $g_{w'}(bwc) = 1$ finden. Wir setzen nun $u = wcw'$ und $v = w'bw$. Offensichtlich sind dann $u \in wg_w^{-1}(0)$ und $v \in g_w^{-1}(0)w$. Bleibt zu vermerken, daß die Beziehung $w' = g_{w'}(bwc)w' = (w'bw)(wcw') = vu$ gilt.

2. *Der Fall $ww' = 0$, $0 \neq w'w$:* Es ist dann offensichtlich $g_w(w'z) = 0$ für alle $z \in \mathfrak{A}$ und $w'w \in \mathfrak{F}_1(\mathfrak{A})$. Wir betrachten nun nichttriviale Elemente $c \in g_{w'w}^{-1}(0)$. Wenigstens für eines von ihnen muß ein Element $b \in \mathfrak{A}$ mit $g_{w'}(bwc) \neq 0$ existieren, denn wegen der Zerlegung $\mathfrak{A} = \mathbb{R}w'w \oplus g_{w'w}^{-1}(0)$ und der Gleichheit $ww' = 0$ wäre sonst $g_{w'}(bwc) = 0$ für alle $b, c \in \mathfrak{A}$, wegen $\mathfrak{A}w\mathfrak{A} = \mathfrak{F}_1(\mathfrak{A}) \cup \{0\}$ also $g_u(w') = g_{w'}(u) = 0$ für alle $u \in \mathfrak{F}_1(\mathfrak{A})$ im Widerspruch zur

o-Totalität von $\mathfrak{F}_1(\mathfrak{A})$. Es kann dann $g_{w'}(bwc) = 1$ angenommen werden, und wie in Fall 1 erhält man eine Darstellung $w' = vu$ mit $u \in wg_w^{-1}(0)$ und $v \in g_w^{-1}(0)w$.

3. *Der Fall* $ww' \neq 0,\ 0 = w'w$: Dieser Fall kann analog zum vorigen behandelt werden, weshalb wir irgendwelche Ausführungen zu ihm unterlassen.

4. *Der Fall* $g_w(w') \neq 0$: Wir fixieren ein beliebiges Element $z \in \mathfrak{A}$ mit $g_{w'}(z) = 1$ und können dann die Identität

$$w' = ww'w + ww'z(w' - w'w) + (w' - ww')zw'w + (w' - ww')zw'z(w' - w'w)$$

notieren. Wir setzen als erstes $u = ww'z(w'-w'w)$. Da w idempotent, also $g_w(w) = 1$ ist, gilt $g_w(ww') = g_w(w')$ und folglich $g_w(u) = g_w(ww' - ww'w) = 0$. Also haben wir $u \in wg_w^{-1}(0)$. Analog kommen wir zur Inklusion $v \in g_w^{-1}(0)w$ für das Element $v := (w' - ww')zw'w$. Die obige Zerlegung des Elementes w' kann nun in der Form $w' = ww'w + u + v + \beta vu$ mit $\beta = 1/g_w(w')$ geschrieben werden. Es verbleibt, sich davon zu überzeugen, daß auch das Element $ww'w$ hierin in der von $wg_w^{-1}(0)$ und $g_w^{-1}(0)w$ erzeugten Unteralgebra enthalten ist. Dazu betrachten wir beliebige Elemente $b, c \in g_w^{-1}(0)$. Gemäß obigem Lemma können diese so gewählt werden, daß $g_w(cb) \neq 0$ ist. Dann folgt die gewünschte Inklusion aus der Darstellung $ww'w = \gamma(wc)(bw)$ mit $\gamma = g_w(w')/g_w(cb)$. ∎

Beispiel. Wir betrachten den n-dimensionalen Raum $I\!R^n$, dessen Elemente $x = (x_1, \ldots, x_n)^T$ wir als Spaltenvektoren schreiben werden. Es sei $\{e^{(i)}\}_{i=1,\ldots,n}$ seine natürliche Basis, d.h., die i-te Koordinate des Elementes $e^{(i)}$ sei gleich 1, und seine übrigen Koordinaten seien gleich 0. Mit $x' = (x'_1, \ldots, x'_n)$ wird das von einem Element $x \in \mathfrak{A}$ erzeugte Funktional und mit A der von einer Matrix $(a_{ij})_{i,j=1,\ldots,n}$ erzeugte Operator bezeichnet. Wir definieren nun den eindimensionalen Operator $w = e^{(i)'} \otimes e^{(i)}$, in dessen Bezeichnung wir der Einfachheit halber den Index i unterdrücken. Offensichtlich ist $w \in \mathfrak{F}_1(\mathfrak{L}(I\!R^n))$ und $w^2 = w$. Für $A \in \mathfrak{L}(I\!R^n)$ ist $g_w^{-1}(A) = 0$ äquivalent zu $a_{ij} = 0$. Folglich bestehen die Unteralgebren $wg_w^{-1}(0)$ und $g_w^{-1}(0)w$ aus allen eindimensionalen Operatoren der Form $x' \otimes e^{(i)}$ ($x \in I\!R^n$) mit $x'_i = 0$ bzw. $e^{(i)'} \otimes y$ ($y \in I\!R^n$) mit $y_i = 0$.

Ist s der natürliche Vektorraum aller abzählbar unendlichen Folgen, so bleibt das Beispiel mit gewissen Modifikationen für Algebren $\mathfrak{A} \subset \mathfrak{L}(s)$ gültig. Wichtiger sind, analog zur eingangs geschilderten Situation $\mathfrak{A} = \mathfrak{B}(X)$ für einen Banach-Raum X, topologische Algebren $\mathfrak{A}$ mit „hinreichend" umfangreichem schwachem Abschluß $\overline{\mathfrak{F}_1(\mathfrak{A})}$. Obwohl offensichtlich, sei schließlich erwähnt, daß unser Resultat auch unter dem Blickpunkt der Frage von Interesse ist, wann eine nichtkommutative Algebra durch kommutative Unteralgebren erzeugt werden kann.

Literatur

[1] **Synnatzschke, J.** : *Über eindimensionale Elemente linearer Algebren und einige algebraische Eigenschaften eines elementaren Operators.* Math. Nachr. **123**, 121–130 (1985)

[2] **Synnatzschke, J.** : *Stetige Elemente bedingt vollständiger* 1-*Algebren.* Math. Nachr. **135**, 223–236 (1988)

[3] **Żelazko, W.** : $B(X)$ *is generated in strong operator topology by two subalgebras with square zero.* Proc. Royal Irish Acad. **88.A**, 19–21 (1988)

eingegangen: 26. März 1990 **revidierte Fassung:** 11. Juli 1990

Verfasser:

Dr. J. Synnatzschke
Universität Leipzig
Sektion Mathematik
Augustusplatz 10
O–7010 Leipzig
Germany

Rostock. Math. Kolloq. 45, 25–34 (1991)

Subject Classification (AMS)
16A32, 46F10

ANDREAS GUNDLACH

Klassifikation von Distributionenalgebren mit Hilfe der Peirceschen Zerlegung[1]

1 Einleitung

Der Gegenstand dieser Arbeit hat seinen Ursprung im Problem der Einführung geeigneter Produktdefinitionen für Distributionen. Ausgangspunkt der Betrachtungen sind dabei die *Heaviside*sche Sprungfunktion $h(t)$ und die *Dirac*sche Delta-Distribution $\delta(t)$ mit $\delta = h'$.

Bereits 1954 zeigte L. Schwarz in [19] die Unlösbarkeit dieses Problems innerhalb der Distributionen. In der analytischen Theorie der Distributionen (wie zum Beispiel bei P. Antosik, J. Mikusiński und R. Siborski in [1]) sind solche Produkte nur in einigen speziellen Fällen mit Hilfe von Delta-Folgen definiert, wie z.B. $h\delta = \delta h = \frac{1}{2}\delta$. Das Produkt $\delta\delta$ existiert dort nicht.

1982 und 1983 stellte J. F. Colombeau in [14] - [16] geeignete kommutative Produkte vor, die deshalb von besonderer Bedeutung sind, weil seine Theorie der nichtlinearen verallgemeinerten Funktionen (vgl. [12, 13]) eine Verallgemeinerung der klassischen C^∞-Funktionen darstellt und eine Anwendung in der Physik möglich ist.

Jedoch bereits 1975 gelang L. Berg in [3] eine formale Produktdefinition, die auch den Fall $\delta\delta$ erfaßt. In einer Reihe von Publikationen (s. [2] - [5]) entwickelt er die Theorie und untersucht nichtkommutative Differentiationsalgebren, die durch die *Heaviside*sche Sprungfunktion $h(t)$ erzeugt werden und damit auch die *Dirac*sche Delta-Distribution $\delta(t)$ enthalten. Auch wenn bisher keine Anwendungen der Theorie von L. Berg über die nichtkommutativen Differentiationsalgebren bekannt sind, so gelingt doch innerhalb dieser Strukturen die Einführung von Integralen, Pseudoinversen und Wurzeln (vgl. L. Berg [5] - [9]).

Besonders die Deutung von Distributionen als Operatoren, wie zum Beispiel in [5] oder in [11], läßt eine Anwendung dieser Theorie in der Operatorenrechnung erwarten.

Deshalb untersucht auch die vorliegende Arbeit die Struktur von Distributionenalgebren, die in [3] von L. Berg wie folgt definiert wurden:

[1] Vortrag, gehalten auf dem Bremen-Rostocker-Kolloquium, Juni 1990

Eine Distributionenalgebra D ist eine assoziative Differentiationsalgebra über einem kommutativen Körper P mit mindestens einem Element h, das den Bedingungen

$$h^2 = h \tag{1}$$

sowie

$$h' \neq 0 \qquad \text{(und damit auch } h \neq 0) \tag{2}$$

genügt.

Die Differentiation hat dabei die Eigenschaften

$$\begin{aligned} &(a) \quad (x+y)' = x' + y', \\ &(b) \quad (xy)' = x'y + xy' \qquad \textit{sowie} \\ &(c) \quad (ax)' = ax' \end{aligned} \tag{3}$$

für alle Elemente x, y von D und $a \in P$.

Mit der Bezeichnung $\delta = h'$, für die (2) dann

$$\delta \neq 0 \tag{2'}$$

lautet, und den üblichen Bezeichnungen für höhere Ableitungen wurde in [3] gezeigt, daß jede Distributionenalgebra nichtkommutativ ist und in ihr für beliebige ganze Zahlen $n \geq 0$ die Gleichungen

$$h\delta^{(n)} = \delta^{(n)} - \sum_{i=1}^{n} \binom{n+1}{i} \delta^{(n-i)}\delta^{(i-1)} - \delta^{(n)}h, \tag{4}$$

$$h\delta^{2n+1} = \delta^{2n+1} - \delta^{2n+1}h \tag{5}$$

sowie

$$h\delta^{2n} = \delta^{2n}h \tag{6}$$

erfüllt sind.

Die zugehörige freie Distributionenalgebra D_0 aus [3] wurde in [10] konstruiert.

In der vorliegenden Arbeit wird an einem konkreten Beispiel gezeigt, daß außer (4), (5) und (6) noch weitere lineare Identitäten zwischen den Elementen gültig sind. Im Abschnitt 2 wird gezeigt, wie wir mit Hilfe der *Peirce*schen Zerlegung weitere mögliche lineare Identitäten analysieren können.

Eine allgemeinere Beschreibung dieser Methode sowie weitere Beispiele findet man in [18]. Eine Klassifikation erfolgt hier nur in dem Sinne, daß wir verschiedene Distributionenalgebren dadurch unterscheiden, welche linearen Identitäten in einer Algebra gültig sind und welche nicht.

2 Analyse linearer Identitäten mit Hilfe der Peirceschen Zerlegung

Wir betrachten nun Distributionenalgebren D, die von h und den Ableitungen von h erzeugt werden.

Wegen (4), (5) und (6) sind alle Elemente, die h als linken Faktor enthalten, Linearkombinationen von Elementen, die h nur als rechten Faktor oder gar nicht enthalten. Wenn wir nach möglichen linearen Identitäten zwischen Elementen aus D fragen, können wir uns also auf Elemente dieser Art beschränken.

Als Beispiel wählen wir hier δ', $\delta' h$, δ^2 und $\delta^2 h$ und suchen alle möglichen linearen Identitäten der Art

$$c_1\delta' + c_2\delta' h + c_3\delta^2 + c_4\delta^2 h = 0 \tag{7}$$

mit c_1, c_2, c_3, $c_4 \in P$.

Dabei benutzen wir, daß D eine Algebra mit einem Idempotent h ist. Danach erhalten wir die zweiseitige *Peirce*sche Zerlegung (vgl. M. Deuring in [17]) für jedes Element x von D in der Form

$$x = x_{00} + x_{01} + x_{10} + x_{11} \tag{8}$$

mit

$$x_{00} := x - xh - hx + hxh, \tag{9a}$$

$$x_{01} := xh - hxh, \tag{9b}$$

$$x_{10} := hx - hxh, \tag{9c}$$

$$x_{11} := hxh. \tag{9d}$$

Wir setzen nun nacheinander (7), d.h. $x = c_1\delta' + c_2\delta' h + c_3\delta^2 + c_4\delta^2 h = 0$, in (9a), (9b), (9c) und (9d) ein. Indem wir die Elemente, die h als linken Faktor enthalten, mit Hilfe von (4), (5) und (6) durch Linearkombinationen von Elementen ersetzen, die h nicht als linken Faktor enthalten, ergeben sich

$$\delta^2 - \delta^2 h = 0, \tag{10a}$$

$$\delta' h + 2\delta^2 h = 0, \tag{10b}$$

$$\delta' - \delta' h - 2\delta^2 + 2\delta^2 h = 0, \tag{10c}$$

$$\delta^2 h = 0, \tag{10d}$$

d.h., daß alle linearen Identitäten der Art (7), die in einer Distributionenalgebra D erfüllt sind, Linearkombinationen von (10a), (10b), (10c) und (10d) sind.

Damit können wir 16 Distributionenalgebren unterscheiden, je nachdem, ob (10a), (10b), (10c) und (10d) in einer Algebra erfüllt sind oder nicht.

Für den Fall, daß (10a) in einer Algebra nicht erfüllt ist, schreiben wir, es gilt $(\overline{10a})$. Bei (10b), (10c) und (10d) verfahren wir analog.

3 Die Existenz der Distributionenalgebren

In diesem Abschnitt zeigen wir mit Hilfe geeigneter Matrixdarstellungen, daß die 16 Distributionenalgebren aus Abschnitt 2 tatsächlich existieren. Als Idempotent h wählen wir

$$h = \text{diag}\,(h_1, h_2, \ldots, h_{n-1}, h_n) \quad \text{mit } h_i = 0 \text{ oder } 1 \text{ für alle } i = 1, 2, \ldots, n. \tag{11}$$

Wegen (2), (2') und (4) mit $n = 0$ setzen wir voraus, daß h nicht die Nullmatrix oder die Einheitsmatrix ist. Wie in [7] und [8] benutzen wir die Differentiation

$$x' = x\sigma - \sigma x, \tag{12}$$

wobei σ eine Matrix vom gleichen Format $n \times n$ wie h ist. Für unsere Beispiele wählen wir σ wie in [7] und [8] mit

$$\sigma = (\sigma_{ik}), \quad \text{wobei } \sigma_{i,i+1} = 1 \text{ und } \sigma_{ik} = 0 \text{ für alle } k \neq i + 1. \tag{13}$$

Unsere Distributionenalgebra D ist eine Matrixalgebra über $I\!R$, die aus einem Element h der Form (11) und der Differentiation (12) mit σ aus (13) erzeugt wird. Solche Matrixalgebren finden wir auch in [2] und [7].

Mit $\delta' = ((\delta')_{ik})$, $\delta' h = ((\delta' h)_{ik})$, $\delta^2 = ((\delta^2)_{ik})$ und $\delta^2 h = ((\delta^2 h)_{ik})$ sowie mit (11), (12) und (13) erhalten wir die Gleichungen

$$
\begin{array}{rcllll}
(\delta')_{i,i+2} & = & h_i - 2h_{i+1} + h_{i+2} & \text{und} \quad (\delta')_{ik} = 0 & \text{für alle} \quad k \neq i+2, & (14)\\
(\delta' h)_{i,i+2} & = & (\delta')_{i,i+2} \cdot h_{i+2} & \text{und} \quad (\delta' h)_{ik} = 0 & \text{für alle} \quad k \neq i+2, & (15)\\
(\delta^2)_{i,i+2} & = & (h_i - h_{i+1})(h_{i+1} - h_{i+2}) & \text{und} \quad (\delta^2)_{ik} = 0 & \text{für alle} \quad k \neq i+2, & (16)\\
(\delta^2 h)_{i,i+2} & = & (\delta^2)_{i,i+2} \cdot h_{i+2} & \text{und} \quad (\delta^2 h)_{ik} = 0 & \text{für alle} \quad k \neq i+2. & (17)
\end{array}
$$

Definition 1 *Drei aufeinanderfolgende Komponenten aus h der Art h_i, h_{i+1}, h_{i+2} bezeichnen wir als Segment von h der Länge 3.*

Wegen (11) gibt es genau 8 verschiedene Segmente der Länge 3. Für $j \in \{1, 2, \ldots, 8\}$ bezeichnen wir mit S_j das Segment der Länge 3, in dem die hintereinander geschriebenen Komponenten $h_i h_{i+1} h_{i+2}$ die dreistellige Dualzahl $\varrho(j-1)$ bilden, vgl. die spätere Tabelle 1. Im folgenden bedeutet $\overline{S_j}$, daß h das Segment S_j nicht enthält.

Satz 1 *Die Matrix h, wie unter* (11), *mit der Differentiation* (12) *und σ, wie unter* (13), *erzeugt genau dann eine Distributionenalgebra, in der*

($\overline{10\text{a}}$), ($\overline{10\text{b}}$), ($\overline{10\text{c}}$), ($\overline{10\text{d}}$) *gilt, wenn h die Segmente* S_3, $S_2 \vee S_4$, $S_5 \vee S_7$, S_6 *enthält,*

($\overline{10\text{a}}$), ($\overline{10\text{b}}$), ($\overline{10\text{c}}$), (10d) *gilt, wenn h die Segmente* S_3, $S_2 \vee S_4$, $S_5 \vee S_7$, $\overline{S_6}$ *enthält,*

($\overline{10\text{a}}$), ($\overline{10\text{b}}$), (10c), ($\overline{10\text{d}}$) *gilt, wenn h die Segmente* S_3, $S_2 \vee S_4$, $\overline{S_5 \vee S_7}$, S_6 *enthält,*

($\overline{10\text{a}}$), ($\overline{10\text{b}}$), (10c), (10d) *gilt, wenn h die Segmente* S_3, $S_2 \vee S_4$, $\overline{S_5 \vee S_7}$, $\overline{S_6}$ *enthält,*

($\overline{10\text{a}}$), (10b), ($\overline{10\text{c}}$), ($\overline{10\text{d}}$) *gilt, wenn h die Segmente* S_3, $\overline{S_2 \vee S_4}$, $S_5 \vee S_7$, S_6 *enthält,*

($\overline{10\text{a}}$), (10b), ($\overline{10\text{c}}$), (10d) *gilt, wenn h die Segmente* S_3, $\overline{S_2 \vee S_4}$, $S_5 \vee S_7$, $\overline{S_6}$ *enthält,*

($\overline{10\text{a}}$), (10b), (10c), ($\overline{10\text{d}}$) *gilt, wenn h die Segmente* S_3, $\overline{S_2 \vee S_4}$, $\overline{S_5 \vee S_7}$, S_6 *enthält,*

($\overline{10\text{a}}$), (10b), (10c), (10d) *gilt, wenn h die Segmente* S_3, $\overline{S_2 \vee S_4}$, $\overline{S_5 \vee S_7}$, $\overline{S_6}$ *enthält,*

(10a), ($\overline{10\text{b}}$), ($\overline{10\text{c}}$), ($\overline{10\text{d}}$) *gilt, wenn h die Segmente* $\overline{S_3}$, $S_2 \vee S_4$, $S_5 \vee S_7$, S_6 *enthält,*

(10a), ($\overline{10\text{b}}$), ($\overline{10\text{c}}$), (10d) *gilt, wenn h die Segmente* $\overline{S_3}$, $S_2 \vee S_4$, $S_5 \vee S_7$, $\overline{S_6}$ *enthält,*

(10a), ($\overline{10\text{b}}$), (10c), ($\overline{10\text{d}}$) *gilt, wenn h die Segmente* $\overline{S_3}$, $S_2 \vee S_4$, $\overline{S_5 \vee S_7}$, S_6 *enthält,* (10a), $(\overline{10b})$, (10c), (10d) *gilt, wenn* h *die Segmente* $\overline{S_3}$, $S_2 \vee S_4$, $\overline{S_5 \vee S_7}$, $\overline{S_6}$ *enthält,*

(10a), (10b), $(\overline{10c})$, $(\overline{10d})$ *gilt, wenn* h *die Segmente* $\overline{S_3}$, $\overline{S_2 \vee S_4}$, $S_5 \vee S_7$, S_6 *enthält,*

(10a), (10b), $(\overline{10c})$, (10d) *gilt, wenn* h *die Segmente* $\overline{S_3}$, $\overline{S_2 \vee S_4}$, $S_5 \vee S_7$, $\overline{S_6}$ *enthält,*

(10a), (10b), (10c), $(\overline{10d})$ *gilt, wenn* h *die Segmente* $\overline{S_3}$, $\overline{S_2 \vee S_4}$, $\overline{S_5 \vee S_7}$, S_6 *enthält,*

(10a), (10b), (10c), (10d) *gilt, wenn* h *die Segmente* $\overline{S_3}$, $\overline{S_2 \vee S_4}$, $\overline{S_5 \vee S_7}$, $\overline{S_6}$ *enthält.*

Beweis. Wegen (11), (12) und (13) sind (1) und (3) und somit auch (4), (5) und (6) erfüllt. Aus (14), (15), (16) und (17) erkennt man, daß für die Elemente δ', $\delta' h$, δ^2 und $\delta^2 h$ nur Segmente der Länge 3 (vgl. Definition 1) von Bedeutung sind. Mit Hilfe von (14), (15), (16) und (17) können wir die Tabelle 1 aufstellen, in der alle Segmente der Länge 3 erfaßt sind. In der letzten Spalte ist vermerkt, welche der Identitäten (10a), (10b), (10c) und (10d) nicht erfüllt sind. Im Fall $\{(\overline{10a}), (\overline{10b}), (\overline{10c}), (\overline{10d})\}$ erkennt man aus Tabelle 1 leicht, daß h genau die Segmente S_3, $S_2 \vee S_4$, $S_5 \vee S_7$ und S_6 enthalten muß. Die anderen 15 Fälle im Satz 1 lassen sich analog verifizieren.

$S_j = (h_i, h_{i+1}, h_{i+2})$	$(\delta')_{i,i+2}$	$(\delta' h)_{i,i+2}$	$(\delta^2)_{i,i+2}$	$(\delta^2 h)_{i,i+2}$	
$S_1 = (0,\ 0,\ 0)$	0	0	0	0	-
$S_2 = (0,\ 0,\ 1)$	1	1	0	0	$(\overline{10b})$
$S_3 = (0,\ 1,\ 0)$	−2	0	−1	0	$(\overline{10a})$
$S_4 = (0,\ 1,\ 1)$	−1	−1	0	0	$(\overline{10b})$
$S_5 = (1,\ 0,\ 0)$	1	0	0	0	$(\overline{10c})$
$S_6 = (1,\ 0,\ 1)$	2	2	−1	−1	$(\overline{10d})$
$S_7 = (1,\ 1,\ 0)$	−1	0	0	0	$(\overline{10c})$
$S_8 = (1,\ 1,\ 1)$	0	0	0	0	–

Tabelle 1

Der folgende Graph in Abbildung 1 macht deutlich, welche Reihenfolge die Segmente S_j der Länge 3 in h einnehmen können.

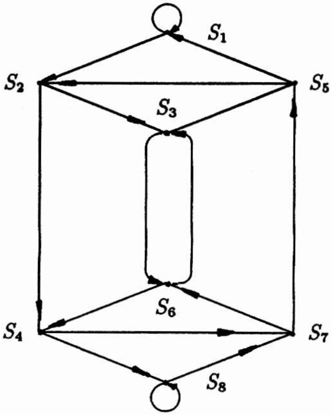

Abbildung 1

Mit Hilfe von Abbildung 1 und Satz 1 erhalten wir nun sehr einfach Beispiele für die einzelnen Distributionenalgebren. So zum Beispiel für $\{(10a), \overline{(10b)}, \overline{(10c)}, (10d)\}$, indem wir aus Satz 1 entnehmen, daß in diesem Fall $\overline{S_3}$, $S_2 \vee S_4$, $S_5 \vee S_7$ und $\overline{S_6}$ in h enthalten sind. Wir dürfen in Abbildung 1 also nur Segmente durchlaufen, die h für den jeweiligen Fall enthält.

Eine Möglichkeit wäre zum Beispiel

$$S_2 \longrightarrow S_4 \longrightarrow S_7.$$

Damit erhalten wir (vgl. auch Tabelle 1)

$$h = \operatorname{diag}\ (\underbrace{0,\ \overbrace{0,\ 1,}^{S_4}}_{S_2}\ \underbrace{1,\ 0}_{S_7}).$$

Eine weitere Möglichkeit wäre

$$S_2 \longrightarrow S_4 \longrightarrow S_7 \longrightarrow S_5 \longrightarrow S_2 \longrightarrow S_4 \longrightarrow S_8$$

und damit

$$h = \text{diag}\,(\,0,\;\;0,\;\;1,\;\;1,\;\;0,\;\;0,\;\;1,\;\;1,\;\;1\,).$$

In den anderen Fällen können wir analoge Beispiele konstruieren.

Der Fall {(10a), (10b), (10c), (10d)} bildet jedoch eine Ausnahme. In diesem Fall dürfen keine Segmente außer S_1 oder S_8 in h enthalten sein. Dies bedeutet aber, daß h entweder die Null- oder die Einheitsmatrix ist, was aber immer im Widerspruch zu (2') steht.
Indem wir aber einfach h = diag (0, 1) oder h = diag (1, 0) wählen, erhalten wir zwei Beispiele für den Fall {(10a), (10b), (10c), (10d)}, in denen auch (2') gilt.

Literatur

[1] **Antosik, P., Mikusiński, J.**, and **Sikorski, R.**: *Theory of Distributions, the Sequential Approach.* Amsterdam, Warzawa 1973

[2] **Berg, L.** : *Noncommutative operations with distributions.* In: I. Dimovski (Ed.): *Generalized Functions and Operational Calculus.* Proceedings of the conference, Varna, Sept. 29 - Okt. 6, 1975. pp. 27–34, Sofia 1979

[3] **Berg, L.** : *Multiplication of distributions.* Math. Nachr. **76**, 195–202 (1977)

[4] **Berg, L.** : *Produkte der Diracschen Delta-Distribution und ihre Ableitungen.* Nova Acta Leopoldina (N.F.) **44**, Supplementum Nr. 8, 69–77 (1976)

[5] **Berg, L.** : *Representations for distributions algebras.* Z. Angew. Math. Mech. **56**, 177–181 (1976)

[6] **Berg, L.** : *Differentiationsalgebren mit einem kubischen Element.* Z. Anal. Anwendungen **1**, 4, 25–30 (1982)

[7] **Berg, L.** : *Erweiterung von Distributionenalgebren.* Rostock. Math. Kolloq. **1**, 5–13 (1976)

[8] **Berg, L.** : *Distributionenalgebren mit Wurzeln.* Rostock. Math. Kolloq. **13**, 73–80 (1979)

[9] **Berg, L.**: *Distribution algebras with reflexive inverses.* Resultate Math. **3**, 7–16 (1980)

[10] **Berg, L.**: *Construction of distribution algebras.* Math. Nachr. **82**, 255–262 (1978)

[11] **Berg, L.**: *The division of distributions by the independent variable.* Math. Nachr. **78**, 327–338 (1977)

[12] **Biagioni, H.A.**, and **Colombeau, J.F.**: *New generalized functions and C^∞-functions with values in generalized complex numbers.* J. London Math. Soc. (2) **33**, 169–179 (1986)

[13] **Colombeau, J.F.**: *New generalized functions and multiplication of distributions.* North-Holland Math. Studies **84**, Amsterdam 1984

[14] **Colombeau, J.F.**: *A multiplication of distributions.* J. Math. Anal. Appl. **94**, 1, 96–115 (1983)

[15] **Colombeau, J.F.**: *New generalized functions. Multiplication of distributions. Physical applications.* Portugal. Math. **41**, 1–4, 57–69 (1982)

[16] **Colombeau, J.F.**: *Une multiplication generale dès distributions.* C.R. Acad. Sci. Paris Sér. I Math. **296**, 357–360 (1983)

[17] **Deuring, M.**: *Algebren.* Berlin 1935

[18] **Gundlach, A.**: *Klassifikation von Distributionenalgebren durch lineare Identitäten.* Dissertation, Universität Rostock 1990

[19] **Schwartz, L.**: *Séminaire Schwartz, 2e année 1954–55. Équations aux dériviées partielles.* Exposé N^o **7**, p. 7–05. Faculté des Siences de Paris, Paris 1955

eingegangen: 20. Dezember 1990

Verfasser:

Dr. A. Gundlach

Abbenser Str. 22

W–3155 Edemissen

Germany

Rostock. Math. Kolloq. 45, 35–42 (1991)

Subject Classification (AMS)
35G10, 35C10

JÜRGEN PÜNGEL

Verallgemeinerte Separierbarkeit formaler Anfangswertprobleme

1 Einleitung

Für zwei Typen von Kernfunktionen, die bei Integraldarstellungen für die Lösungen einer Klasse linearer selbstadjungierter Differentialgleichungen zweiter Ordnung von Bedeutung sind, wurden in [3] sogenannte Additionsformeln hergeleitet. Eine davon (für Riemann-Funktionen) wurde bereits 1952 in [1] angegeben; vgl. auch [2]. Das in [3] vorgeschlagene gemeinsame Prinzip zur Konstruktion solcher Additionsformeln besteht darin, nach Wahl geeigneter Variablen das jeweils zu lösende Anfangswertproblem in ein nach diesen Variablen separierbares Anfangswertproblem zu transformieren, d.h. Produktlösungen zu finden, deren zwei Faktoren jeweils nur von einem Teil der Variablen abhängen, und somit das Anfangswertproblem auf zwei einfachere Anfangswertprobleme zurückzuführen. Auf diese Weise gelingt es, Riemann-Funktionen und Bergman-Erzeugende für Differentialgleichungen der Form $u_{z\zeta} = [f(z+\zeta)+g(z-\zeta)]u$ durch jene für $u_{z\zeta} = f(z+\zeta)u$ und $u_{z\zeta} = g(z-\zeta)u$ auszudrücken. Die eine Transformation (für Riemann-Funktionen) ist

$$\sum_{\mu=0}^{\infty} p_\mu(\alpha,\beta)\omega^\mu/\mu! \longrightarrow \sum_{\mu=0}^{\infty} p_\mu(\alpha,\beta)\omega^\mu,$$

die andere (für Bergman-Erzeugende) vom Typ

$$\sum_{\mu=0}^{\infty} p_\mu(\alpha,\beta)\omega^\mu/\mu! \longrightarrow \sum_{\mu=0}^{\infty} p_\mu(\alpha,\beta)(2\mu)!(\omega/4)^\mu/(\mu!)^2$$

mit jeweils geeigneten komplexen Variablen α, β, ω. Beide Transformationen sind aus der Klasse

$$\sum_{\mu=0}^{\infty} p_\mu(\alpha,\beta)\omega^\mu/\mu! \longrightarrow \sum_{\mu=0}^{\infty} (\lambda)_\mu p_\mu(\alpha,\beta)\omega^\mu/\mu!$$

mit $(\lambda)_\mu := \lambda(\lambda+1)\dots(\lambda+\mu-1)$, $(\lambda)_0 := 1$, $\lambda \in \mathbb{C}$, $-\lambda \notin I\!N$ (für $\lambda = 1$ bzw. $\lambda = 1/2$ ergeben sich die beiden erstgenannten Transformationen).

Da bei diesen Transformationen aber ersichtlich i.allg. die Konvergenz der Bildreihe nicht gesichert ist, „formal" damit aber doch die erwähnten Additionsformeln hergeleitet werden können, ist es zweckmäßig, diese Transformationen auf geeigneten Mengen formaler Potenzreihen über einem Ring holomorpher Funktionen zu definieren. Hier werden noch allgemeinere Transformationen der Art

$$\sum_{\mu=0}^{\infty} p_\mu(\alpha,\beta)\omega^\mu/\mu! \longrightarrow \sum_{\mu=0}^{\infty}(\lambda_1)_\mu(\lambda_2)_\mu\dots(\lambda_n)_\mu p_\mu(\alpha,\beta)\omega^\mu/\mu!$$

betrachtet. Sie sind durch n weitgehend beliebige Parameter $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$ bestimmt (nur die Einschränkung $-\lambda_j \notin \mathbb{N}$ für $j \in \{1,\dots,n\}$ möge die Invertierbarkeit sichern). Dies eröffnet, wie im folgenden gezeigt wird, die Möglichkeit, eine sehr allgemeine Klasse von Anfangswertproblemen beliebig hoher Ordnung in ein separierbares Anfangswertproblem zweiter Ordnung zu transformieren und mittels dieser „verallgemeinerten Separierbarkeit" weitere Additionsformeln zu gewinnen.

2 Voraussetzungen

Für zwei nichtleere Gebiete $G_1, G_2 \subset \mathbb{C}$ bezeichne $\mathbf{R}$ den Ring aller auf $G_1 \times G_2 \subset \mathbb{C}^2$ definierten und dort holomorphen Funktionen von zwei (unabhängigen) komplexen Variablen $\alpha \in G_1, \beta \in G_2$. Mit $\partial_1, \partial_2 : \mathbf{R} \to \mathbf{R}$ seien die Operatoren der komplexen partiellen Ableitungen nach der ersten bzw. zweiten Komponente in $G_1 \times G_2$ bezeichnet.

Sei $\mathbf{P} := \mathbf{R}^{\mathbb{N}}$ die Menge aller Folgen $p : \mathbb{N} \to \mathbf{R}$. Wie üblich werden mit $p_\mu \in \mathbf{R}$ die Bilder der $\mu \in \mathbb{N}$ unter $p \in \mathbf{P}$ notiert. Mit den durch

$$(p+q)_\mu := p_\mu + q_\mu, \quad (pq)_\mu := \sum_{\nu=0}^{\mu}\binom{\mu}{\nu}p_\nu q_{\mu-\nu} \tag{1}$$

(für $p, q \in \mathbf{P}$, $\mu \in \mathbb{N}$) gegebenen inneren Kompositionen ist auf $\mathbf{P}$ eine Ringstruktur definiert. Mit der vermöge $(ap)_\mu := ap_\mu$ (für alle $a \in \mathbf{R}$, $p \in \mathbf{P}$, $\mu \in \mathbb{N}$) festgelegten äußeren Komposition ist $\mathbf{P}$ zudem ein $\mathbf{R}$-Modul.

Für $X, Y : \mathbf{P} \to \mathbf{P}$ und $q \in \mathbf{P}$ seien die Operatoren $X+Y$, XY, $qX : \mathbf{P}\to\mathbf{P}$ in kanonischer Weise durch

$$(X+Y)p := Xp + Yp, \quad (XY)p := X(Yp), \quad (qX)p := q(Xp) \tag{2}$$

(für alle $p \in \mathbf{P}$) erklärt.

Es bezeichne $\mathbf{H} \subset \mathbf{P}$ jene Teilmenge von Folgen $p \in \mathbf{P}$, für die ein $\varepsilon > 0$ existiert, so daß durch

$$\hat{p}(\alpha,\beta;\omega) := \sum_{\mu=0}^{\infty} p_\mu(\alpha,\beta)\omega^\mu/\mu! \tag{3}$$

eine in $G_1 \times G_2 \times U_\varepsilon$ ($U_\varepsilon := \{\omega \in \boldsymbol{C} | \, |\omega| < \varepsilon\}$) holomorphe Funktion $\hat{p}$ bestimmt ist. Sei $\hat{\mathbf{H}}$ die Menge dieser Funktionen. Die Restriktionen der durch (1) definierten Verknüpfungen auf $\mathbf{H}$ stellen dann ersichtlich Addition und Multiplikation der gemäß (3) zugeordneten Funktionen in $\hat{\mathbf{H}}$ dar. Jeder Abbildung (jedem Operator) $X : \mathbf{P} \to \mathbf{P}$ mit $X[\mathbf{H}] \subset \mathbf{H}$ kann mittels

$$(\hat{X}\hat{p})(\alpha, \beta; \omega) := \sum_{\mu=0}^{\infty} (Xp)_\mu(\alpha, \beta)\omega^\mu / \mu! \tag{4}$$

(für $\hat{p} \in \hat{\mathbf{H}}$ nach (3)) eine Abbildung $\hat{X} : \hat{\mathbf{H}} \to \hat{\mathbf{H}}$ zugeordnet werden.

Die Ableitungen ∂_1, ∂_2 induzieren gemäß

$$(D_i p)_\mu := \partial_i p_\mu \tag{5}$$

(für $p \in \mathbf{P}$, $\mu \in I\!N$, $i \in \{1, 2\}$) kommutierende Derivationen $D_1, D_2 : \mathbf{P} \to \mathbf{P}$. Durch

$$(Dp)_\mu := p_{\mu+1} \tag{6}$$

(für $p \in \mathbf{P}$, $\mu \in I\!N$) ist eine weitere, mit D_1, D_2 vertauschbare Derivation $D : \mathbf{P} \to \mathbf{P}$ gegeben. Für $p \in \mathbf{H}$ entsprechen $D_1 p, D_2 p$ bzw. Dp den partiellen Ableitungen $\partial\hat{p}/\partial\alpha$, $\partial\hat{p}/\partial\beta$ bzw. $\partial\hat{p}/\partial\omega$ der p gemäß (3) zugeordneten Funktion $\hat{p} \in \mathbf{H}$; d.h., es gilt

$$\begin{aligned}
(\hat{D}_1\hat{p})(\alpha, \beta; \omega) &= \frac{\partial}{\partial\alpha}\hat{p}(\alpha, \beta; \omega), \\
(\hat{D}_2\hat{p})(\alpha, \beta; \omega) &= \frac{\partial}{\partial\beta}\hat{p}(\alpha, \beta; \omega), \\
(\hat{D}\hat{p})(\alpha, \beta; \omega) &= \frac{\partial}{\partial\omega}\hat{p}(\alpha, \beta; \omega)
\end{aligned}$$

mit $\hat{D}_1$, $\hat{D}_2$ bzw. $\hat{D}$ nach (4). Die Mengen

$$\mathbf{P}^* := \{p \in \mathbf{P} | \, Dp = 0\}, \quad \mathbf{P}_i := \{p \in \mathbf{P} | \, D_{3-i}p = 0\}, \quad \mathbf{P}_i^* := \mathbf{P}_i \cap \mathbf{P}^*$$

für $i \in \{1, 2\}$, sowie $\mathbf{P}_1 \cap \mathbf{P}_2$, $\mathbf{P}_1^* \cap \mathbf{P}_2^*$ sind Teilringe von $\mathbf{P}$ (Konstantenringe).

Jeder Konstanten $\kappa \in \boldsymbol{C}$ sei der mittels

$$(S_\kappa p)_\mu := (\mu + \kappa)p_\mu \tag{7}$$

definierte Operator $S_\kappa : \mathbf{P} \to \mathbf{P}$ zugeordnet. S_κ entspricht

$$(\hat{S}_\kappa\hat{p})(\alpha, \beta; \omega) = (\omega\frac{\partial}{\partial\omega} + \kappa)\hat{p}(\alpha, \beta; \omega)$$

in $\hat{\mathbf{H}}$. Dem durch

$$(Kp)_\mu := p_{\mu+1}/(\mu+1) \tag{8}$$

gegebenen Operator $K : \mathbf{P} \to \mathbf{P}$ ist nach (4)

$$(\hat{K}\hat{p})(\alpha,\beta;\omega) = [\hat{p}(\alpha,\beta;\omega) - \hat{p}(\alpha,\beta;0)]/\omega$$

zuzuordnen. D_1, D_2 (nicht aber D) sind mit K vertauschbar. Für alle $p, q \in \mathbf{P}$ gilt

$$\begin{aligned} S_\kappa(pq) &= (S_\kappa p)q + p(S_0 q), \\ K(pq) &= (Kp)q + p_0(Kq) = p(Kq) + (Kp)q_0. \end{aligned}$$

S_0 ist somit Derivation von $\mathbf{P}$.

Für jedes n-Tupel $\lambda = \{\lambda_1, \ldots, \lambda_n\} \in \boldsymbol{C}^n$ mit $-\lambda_j \notin I\!N$ für alle $j \in \{1, \ldots, n\}$ ist der vermöge

$$(T_\lambda p)_\mu := p_\mu/(\lambda)_\mu \quad \text{mit} \quad (\lambda)_\mu := \prod_{j=1}^{n} (\lambda_j)_\mu \tag{9}$$

erklärte Operator $T_\lambda : \mathbf{P} \to \mathbf{P}$ invertierbar. Mit $\Pi_\lambda : \mathbf{P} \to \mathbf{P}$, definiert durch

$$\Pi_\lambda := \prod_{j=1}^{n} S_{\lambda_j} \tag{10}$$

$(\lambda \in \boldsymbol{C}^n)$, gilt

$$(\Pi_\lambda p)_\mu = (\lambda)_{\mu+1} p_\mu/(\lambda)_\mu. \tag{11}$$

In $\hat{\mathbf{H}}$ ist

$$(\hat{\Pi}_\lambda \hat{p})(\alpha,\beta;\omega) = (\omega\frac{\partial}{\partial\omega} + \lambda_n)\ldots(\omega\frac{\partial}{\partial\omega} + \lambda_1)\hat{p}(\alpha,\beta;\omega).$$

T_λ ist mit D_1, D_2 und S_0 vertauschbar:

$$D_i T_\lambda = T_\lambda D_i \tag{12}$$

für $i \in \{1, 2\}$ und

$$S_0 T_\lambda = T_\lambda S_0. \tag{13}$$

Dagegen gilt

$$\Pi_\lambda K T_\lambda = T_\lambda K, \tag{14}$$

$$\Pi_\lambda D T_\lambda = T_\lambda D. \tag{15}$$

Jedem $\lambda = (\lambda_1, \dots, \lambda_n) \in \boldsymbol{C}^n$ mit $-\lambda_j \notin \mathbb{N}$ für alle $j \in \{1, \dots, n\}$ sei durch

$$(p \overset{*}{\lambda} q)_\mu := \sum_{\nu=0}^{\mu} \binom{\mu}{\nu} (\lambda)_\nu (\lambda)_{\mu-\nu} p_\nu q_{\mu-\nu} / (\lambda)_\mu \tag{16}$$

ein Produkt $\overset{*}{\lambda} : \mathbf{P} \times \mathbf{P} \to \mathbf{P}$ zugeordnet. Mit (9) gilt

$$p \overset{*}{\lambda} q = T_\lambda((T_\lambda^{-1} p)(T_\lambda^{-1} q)) \tag{17}$$

für alle $p, q \in \mathbf{P}$.

3 Verallgemeinerte Separierbarkeit

Im weiteren wird ein („formales") Anfangswertproblem der Form

$$M_\lambda p = (f^1 + f^2)p \quad \text{mit} \quad p_0 = 1 \tag{18}$$

für $p \in \mathbf{P}$ betrachtet (hochgestellte $i \in \{1, 2\}$ mögen hier und im folgenden stets Indizes bezeichnen). Dabei seien die f^i aus $\mathbf{P}_i^*$, und der Operator $M_\lambda : \mathbf{P} \to \mathbf{P}$ sei durch

$$M_\lambda := (a^1 D_1 + b^1) D_1 + (a^2 D_2 + b^2) D_2 + l S_0 + \Pi_\lambda (c^1 K D_1 + c^2 K D_2 + m D) \tag{19}$$

mit $a^i, b^i, c^i \in \mathbf{P}_i^*$, $m, l \in \mathbf{P}_1^* \cap \mathbf{P}_2^*$ sowie $\lambda \in \boldsymbol{C}^n$ mit $-\lambda_j \notin \mathbb{N}$, gegeben. In $\hat{\mathbf{H}}$ entspricht (18) einem Anfangswertproblem der Ordnung $n + 1$

$$\begin{aligned}
& \hat{a}^1(\alpha)\hat{p}_{\alpha\alpha}(\alpha,\beta;\omega) + \hat{a}^2(\beta)\hat{p}_{\beta\beta}(\alpha,\beta;\omega) \\
+ \quad & \hat{b}^1(\alpha)\hat{p}_\alpha(\alpha,\beta;\omega) + \hat{b}^2(\beta)\hat{p}_\beta(\alpha,\beta;\omega) + \hat{l}\omega\hat{p}_\omega(\alpha,\beta;\omega) \\
+ \quad & (\omega\frac{\partial}{\partial\omega} + \lambda_n)\dots(\omega\frac{\partial}{\partial\omega} + \lambda_1)\left(\frac{\hat{c}^1(\alpha)}{\omega}\hat{p}_\alpha(\alpha,\beta;\omega) + \frac{\hat{c}^2(\beta)}{\omega}\hat{p}_\beta(\alpha,\beta;\omega) + \hat{m}\hat{p}_\omega(\alpha,\beta;\omega)\right) \\
= \quad & (\hat{f}^1(\alpha) + \hat{f}^2(\beta))\hat{p}(\alpha,\beta;\omega) \\
& \text{mit} \quad \hat{p}(\alpha,\beta;0) = 1 \quad \text{für} \quad \alpha \in G_1, \; \beta \in G_2
\end{aligned}$$

für in $G_1 \times G_2 \times U_\varepsilon$ holomorphe $\hat{p}$ mit in G_i holomorphen Koeffizienten $\hat{a}^i$, $\hat{b}^i$, $\hat{c}^i$, $\hat{f}^i$ ($i \in \{1, 2\}$), $\hat{m}, \hat{l} \in \boldsymbol{C}$ und $-\lambda \in \boldsymbol{C} \setminus \mathbb{N}$ $(j \in \{1, \dots, n\})$.

Ist speziell $p^i \in \mathbf{P}_i$, so gilt insbesondere

$$M_\lambda p^i = M_\lambda^i p^i$$

mit den Operatoren $M_\lambda^1, M_\lambda^2 : \mathbf{P} \to \mathbf{P}$ der Form

$$M_\lambda^i := (a^i D_i + b^i) D_i + l S_0 + \Pi_\lambda (c^i K D_i + m D). \tag{20}$$

Das Anfangswertproblem (18) ist in folgendem Sinn nicht separierbar: Für beliebige Koeffizienten in (19) und beliebige $f^i \in \mathbf{P}_i^*$ hat (18) keine nichttrivialen Lösungen $p \in \mathbf{P}$ der Form $p = p^1 p^2$ mit $p^i \in \mathbf{P}_i$. Dagegen gilt der

Satz *Sind $p^i \in \boldsymbol{P}_i$, $i \in \{1,2\}$, Lösungen von*

$$M_\lambda^i p^i = (f^i + (3-2i)\kappa) p^i \quad \text{mit} \quad p_0^i = 1 \tag{21}$$

für $f^i \in \boldsymbol{P}_i^*$, $\kappa \in \boldsymbol{P}_1 \cap \boldsymbol{P}_2$ und M_λ^i nach (20), dann ist

$$p = p^1 \underset{\lambda}{*} p^2 \tag{22}$$

Lösung von (18) mit M_λ nach (19).

Beweis. Infolge (12), (13), (14), (15) ist

$$M_\lambda T_\lambda = T_\lambda L, \; M_\lambda^i T_\lambda = T_\lambda L^i \quad (i \in \{1,2\}) \tag{23}$$

mit $L, L^1, L^2 : \mathbf{P} \to \mathbf{P}$ von der Gestalt

$$\begin{aligned} L &= (a^1 D_1 + b^1) D_1 + (a^2 D_2 + b^2) D_2 + l S_0 + c^1 K D_1 + c^2 K D_2 + m D, \\ L^i &= (a^i D_i + b^i) D_i + l S_0 + c^i K D_i + m D. \end{aligned} \tag{24}$$

Wegen $D_{3-i} p^i = 0$ gilt auch $D_{3-i} q^i = 0$ für $q^i := T_\lambda^{-1} p^i$, was

$$L(q^1 q^2) = (L^1 q^1) q^2 + q^1 (L^2 q^2) \tag{25}$$

zur Folge hat. Die Voraussetzung (21) liefert mit (23) und $p^i = T_\lambda q^i$ die Gleichungen

$$L^i q^i = T_\lambda^{-1} M_\lambda^i T_\lambda q^i = T_\lambda^{-1} M_\lambda^i p^i = T_\lambda^{-1} (f^i \pm \kappa) p^i = (f^i \pm \kappa) q^i$$

($+\kappa$ für $i = 1$, $-\kappa$ für $i = 2$) und mit (25)

$$L(q^1q^2) = (f^1+\kappa)q^1q^2 + (f^2-\kappa)q^1q^2 = (f^1+f^2)q^1q^2;$$

(22) mit (17), d.h. $p = T_\lambda(q^1q^2)$, und (23) ergibt

$$M_\lambda p = T_\lambda L(q^1q^2) = T_\lambda((f^1+f^2)q^1q^2) = (f^1+f^2)p.$$

Schließlich führt $p_0^i = 1$ wegen $q_0^i = 1$ zu $p_0 = 1$. ■

Bezüglich der verallgemeinerten Multiplikation $\overset{*}{\lambda}$ nach (16) bzw. (17) ist demnach das formale Anfangswertproblem (18) separierbar, d.h. $p \in \mathbf{P}$, gegeben durch die „Additionsformel“

$$p_\mu = \sum_{\nu=0}^{\mu} \binom{\mu}{\nu} (\lambda)_\nu (\lambda)_{\mu-\nu} p_\nu^1 p_{\mu-\nu}^2 / (\lambda)_\mu,$$

ist Lösung von (18), wenn $p^1 \in \mathbf{P}_1$ und $p^2 \in \mathbf{P}_2$ Lösungen von $M_\lambda^1 p^1 = f^1p^1$ und $M_\lambda^2 p^2 = f^2p^2$ mit $p_0^1 = p_0^2 = 1$ sind.

4 Verallgemeinerungen

Aus zunächst nur formalen Gründen bietet sich an, für die Folgen in $\mathbf{P} = \mathbf{R}^N$ (Koeffizienten der formalen Potenzreihen) anstelle einer Menge $\mathbf{R}$ holomorpher Funktionen einen beliebigen Differentialring $\mathbf{R}$ mit zwei (oder mehreren, u.U. nicht vertauschbaren) Derivationen zu wählen.

Sollte eine Anwendung auf Differentialgleichungssysteme ins Auge gefaßt werden, so scheint es zweckmäßig, auch nichtkommutative Ringe $\mathbf{R}$ zuzulassen.

Bezogen auf die jeweils gegebene Transformation T könnte sich folgende Definition einer „verallgemeinerten Separierbarkeit” als geeignet herausstellen.

Ein (additiver) Operator $\Omega : \mathbf{P} \to \mathbf{P}$ heiße „T-separierbar”, wenn $\Omega T = T\Omega_0$ mit einem separierbaren Operator $\Omega_0 : \mathbf{P} \to \mathbf{P}$ ist. Dabei heiße Ω_0 „separierbar” (nach D_1, D_2), wenn

$$\Omega_0(p^1p^2) = (\Omega_0^1 p^1)p^2 + p^1(\Omega_0^2 p^2)$$

und

$$D_1\Omega_0^2 p^2 = D_2\Omega_0^1 p^1 = 0$$

gilt für alle p^i mit $D_2p^1 = D_1p^2 = 0$ und geeigneten, Ω_0 zugeordneten Operatoren Ω_0^1, Ω_0^2 (vgl. (25)).

Die durch (9) beschriebene enge Klasse von Transformationen sollte dabei ebenfalls eine Erweiterung erfahren: etwa auf Transformationen T der Art

$$(Tp)_\mu := \sum_{\nu=0}^{\mu} k_{\mu\nu} p_\nu$$

oder allgemeiner

$$(Tp)_\mu := \sum_{\nu=0}^{\infty} k_{\mu\nu} p_\nu$$

mit $k : \mathbb{N} \times \mathbb{N} \to \mathbb{C}$ bzw. $k : \mathbb{N} \times \mathbb{N} \to \mathbf{R}_0$, $\mathbf{R}_0 \subset \mathbf{R}$.

Literatur

[1] **Olevskij, M.N.:** *On the Riemann function for the differential equations* $\partial^2 u/\partial x^2 - \partial^2 u/\partial t^2 + [\rho_1(x) + \rho_2(t)]u = 0$. Dokl. Akad. Nauk SSSR **87**, 337–340 (1952)

[2] **Papadakis, J.S.**, and **Wood, D.H.:** *An addition formula for Riemann functions.* J. Differential Equations **24**, 397–411 (1977)

[3] **Püngel, J.:** *Additionsformeln für Riemannfunktionen und Bergmanerzeugende.* In: Withalm, C. (Ed.): *Complex Methods on Partial Differential Equations.* Math. Res. **53**, 126–136, Berlin 1989

eingegangen: 13. November 1989 **revidierte Fassung:** 23. April 1990

Verfasser:

Doz. Dr. J. Püngel
Institut für Mathematik D
Technische Universität Graz
Steyrergasse 17/5
A-8010 Graz
Austria

Rostock. Math. Kolloq. **45**, 43–56 (1991)

Subject Classification (AMS)
46F10

EMIN ÖZÇAĞ; BRIAN FISHER

On partial derivatives of the Beta function

The Beta function $B(\lambda,\mu)$ is usually defined by

$$B(\lambda,\mu)=\int_0^1 t^{\lambda-1}(1-t)^{\mu-1}dt$$

for $\lambda,\mu>0$. It then follows that

$$B(\lambda,\mu)=\frac{\Gamma(\lambda)\Gamma(\mu)}{\Gamma(\lambda+\mu)},$$

where Γ denotes the Gamma function, and this expression is then used to define the Beta function for $\lambda,\mu<0$ and $\lambda,\mu\neq -1,-2,\dots$.

It can then be shown, see Gel'fand and Shilov [3], that

$$\begin{aligned} B(\lambda,\mu) &= \int_0^{1/2} t^{\lambda-1}\left[(1-t)^{\mu-1}-\sum_{i=0}^{r}\frac{(-1)^i(\mu-1)_i}{i!}t^i\right]dt \\ &\quad+\sum_{i=0}^{r}\frac{(-1)^i(\mu-1)_i}{2^{\lambda+i}i!(\lambda+i)} \\ &\quad+\int_{1/2}^{1}(1-t)^{\mu-1}\left[t^{\lambda-1}-\sum_{i=0}^{s}\frac{(-1)^i(\lambda-1)_i}{i!}(1-t)^i\right]dt \\ &\quad+\sum_{i=0}^{s}\frac{(-1)^i(\lambda-1)_i}{2^{\mu+i}i!(\mu+i)}, \end{aligned}$$

for $\lambda>-r,\ \mu>-s,\ \lambda\neq 0,-1,\dots,-r+1$ and $\mu\neq 0,-1,\dots,-s+1$, where

$$(\lambda)_i=\begin{cases} 1, & i=0,\\ \prod_{j=0}^{i-1}(\lambda-j), & i\geq 1. \end{cases}$$

In [2] it was shown that

$$\begin{aligned}\int_\epsilon^{1-\epsilon} t^{\lambda-1}(1-t)^{\mu-1}dt &= \int_\epsilon^{1/2} t^{\lambda-1}\left[(1-t)^{\mu-1} - \sum_{i=0}^{r}\frac{(-1)^i(\mu-1)_i}{i!}t^i\right]dt \\ &+\sum_{i=0}^{r}\frac{(-1)^i(\mu-1)_i}{i!(\lambda+i)}(2^{-\lambda-i}-\epsilon^{\lambda+i}) \\ &+\int_{1/2}^{1-\epsilon}(1-t)^{\mu-1}\left[t^{\lambda-1}-\sum_{i=0}^{s}\frac{(-1)^i(\lambda-1)_i}{i!}(1-t)^i\right]dt \\ &+\sum_{i=0}^{s}\frac{(-1)^i(\lambda-1)_i}{i!(\mu+i)}(2^{-\mu-i}-\epsilon^{\mu+i})\end{aligned}$$

for $\lambda > -r,\ \mu > -s,\ \lambda \neq 0,-1,\dots,-r+1$ and $\mu \neq 0,-1,\dots,-s+1$, so that

$$B(\lambda,\mu) = \operatorname*{N-lim}_{\epsilon\to 0}\int_\epsilon^{1-\epsilon} t^{\lambda-1}(1-t)^{\mu-1}\,dt,$$

where N is the neutrix, see van der Corput [1], having domain $N' = \{\epsilon:\ 0<\epsilon<\frac{1}{2}\}$ with negligible functions finite linear sums of the functions

$$\epsilon^\lambda \ln^{r-1}\epsilon,\quad \ln^r\epsilon\quad (\lambda<0,\ r=1,2,\dots)$$

and all functions of ϵ which converge to zero in the usual sense as ϵ tends to zero.

This suggests the following definition, given in [2], for

$$B_{p,q}(\lambda,\mu) = \frac{\partial^{p+q}}{\partial^p\lambda\partial^q\mu}B(\lambda,\mu)$$

for all values of λ,μ and $p,q=0,1,2,\dots$.

Definition *The function $B_{p,q}(\lambda,\mu)$ is defined by*

$$B_{p,q}(\lambda,\mu) = \operatorname*{N-lim}_{\epsilon\to 0}\int_\epsilon^{1-\epsilon} t^{\lambda-1}\ln^p t(1-t)^{\mu-1}\ln^q(1-t)dt \tag{1}$$

for $p,q=0,1,2,\dots$ and all λ,μ.

It is not immediately obvious that the neutrix limit in equation (1) exists and it was proved in [2] that this neutrix limit existed only for the case $p=q=0$. In the following, we prove that this neutrix limit exists for $p,q=0,1,2,\dots$ and all λ,μ so that $B_{p,q}(\lambda,\mu)$ is well defined.

We first of all need the following

Lemma *The neutrix limits as ϵ tends to zero of the functions*

$$\int_\epsilon^{1/2} t^\lambda \ln^p t \ln^q(1-t)\,dt, \quad \int_{1/2}^{1-\epsilon} (1-t)^\lambda \ln^p t \ln^q(1-t)dt$$

exist for $p, q = 0, 1, 2, \ldots$ and all λ.

Proof. Suppose first of all that $p = q = 0$. Then

$$\int_\epsilon^{1/2} t^\lambda dt = \begin{cases} \dfrac{2^{-\lambda-1} - \epsilon^{\lambda+1}}{\lambda+1}, & \lambda \neq -1, \\ -\ln 2 - \ln\epsilon, & \lambda = -1 \end{cases}$$

and so

$$\operatorname*{N-lim}_{\epsilon \to 0} \int_\epsilon^{1/2} t^\lambda dt$$

exists for all λ.

Now suppose that $q = 0$ and that

$$\operatorname*{N-lim}_{\epsilon \to 0} \int_\epsilon^{1/2} t^\lambda \ln^p t dt$$

exists for some nonnegative integer p and all λ. Then

$$\int_\epsilon^{1/2} t^\lambda \ln^{p+1} t dt = \begin{cases} \dfrac{-2^{-\lambda-1}\ln 2 - \epsilon^{\lambda+1}\ln\epsilon}{\lambda+1} - \dfrac{p+1}{\lambda+1}\displaystyle\int_\epsilon^{1/2} t^\lambda \ln^p t\,dt, & \lambda \neq -1, \\ \dfrac{(-1)^p \ln^{p+2} 2 - \ln^{p+2}\epsilon}{p+2}, & \lambda = -1 \end{cases}$$

and it follows by induction that

$$\operatorname*{N-lim}_{\epsilon \to 0} \int_\epsilon^{1/2} t^\lambda \ln^p t dt$$

exists for $p = 0, 1, 2, \ldots$ and all λ. Finally we note that we can write

$$\ln^q(1-t) = \sum_{i=1}^{\infty} \alpha_{i_q} t^i$$

for $q = 1, 2, \ldots$, the expansion being valid for $|t| < 1$. Choosing a positive integer k such that $\lambda + k > -1$, we have

$$\int_{\epsilon}^{1/2} t^{\lambda} \ln^p t \ln^q(1-t) dt = \sum_{i=1}^{k-1} \alpha_{iq} \int_{\epsilon}^{1/2} t^{\lambda+i} \ln^p t dt + \sum_{i=k}^{\infty} \alpha_{iq} \int_{\epsilon}^{1/2} t^{\lambda+i} \ln^p t\, dt.$$

It follows from what we have just proved that

$$\operatorname*{N-lim}_{\epsilon \to 0} \sum_{i=1}^{k-1} \alpha_{iq} \int_{\epsilon}^{1/2} t^{\lambda+i} \ln^p t dt$$

exists and further

$$\begin{aligned} \operatorname*{N-lim}_{\epsilon \to 0} \sum_{i=k}^{\infty} \alpha_{iq} \int_{\epsilon}^{1/2} t^{\lambda+i} \ln^p t\, dt &= \lim_{\epsilon \to 0} \sum_{i=k}^{\infty} \alpha_{iq} \int_{\epsilon}^{1/2} t^{\lambda+i} \ln^p t\, dt \\ &= \sum_{i=k}^{\infty} \alpha_{iq} \int_{0}^{1/2} t^{\lambda+i} \ln^p t\, dt, \end{aligned}$$

proving that

$$\operatorname*{N-lim}_{\epsilon \to 0} \int_{\epsilon}^{1/2} t^{\lambda} \ln^p t \ln^q(1-t) dt$$

exists for $p, q = 0, 1, 2, \ldots$ and all λ.

Making the substitution $1 - t = u$ in

$$\int_{1/2}^{1-\epsilon} (1-t)^{\lambda} \ln^p t \ln^q(1-t)\, dt,$$

it follows that

$$\operatorname*{N-lim}_{\epsilon \to 0} \int_{1/2}^{1-\epsilon} (1-t)^{\lambda} \ln^p t \ln^q(1-t) dt$$

also exists for $p, q = 0, 1, 2, \ldots$ and all λ.

Theorem 1 *The function $B_{p,q}(\lambda,\mu)$ exists for $p,q = 0,1,2,\ldots$ and all λ,μ.*

Proof. Choose positive integers r, s such that $\lambda > -r$ and $\mu > -s$. Then we can write

$$\begin{aligned}
&\int_{\epsilon}^{1-\epsilon} t^{\lambda-1}\ln^p t(1-t)^{\mu-1}\ln^q(1-t)dt \\
&= \int_{\epsilon}^{1/2} t^{\lambda-1}\ln^p t\ln^q(1-t)\left[(1-t)^{\mu-1} - \sum_{i=0}^{r}\frac{(-1)^i(\mu-1)_i}{i!}t^i\right]dt \\
&+ \sum_{i=0}^{r}\frac{(-1)^i(\mu-1)_i}{i!}\int_{\epsilon}^{1/2} t^{\lambda+i-1}\ln^p t\ln^q(1-t)dt \\
&+ \int_{1/2}^{1-\epsilon} \ln^p t(1-t)^{\mu-1}\ln^q(1-t)\left[t^{\lambda-1} - \sum_{i=0}^{s}\frac{(-1)^i(\lambda-1)_i}{i!}(1-t)^i\right]dt \\
&+ \sum_{i=0}^{s}\frac{(-1)^i(\lambda-1)_i}{i!}\int_{1/2}^{1-\epsilon}(1-t)^{\mu+i-1}\ln^p t\ln^q(1-t)\,dt.
\end{aligned}$$

We have

$$\lim_{\epsilon\to 0}\int_{\epsilon}^{1/2} t^{\lambda-1}\ln^p t\ln^q(1-t)\left[(1-t)^{\mu-1} - \sum_{i=0}^{r}\frac{(-1)^i(\mu-1)_i}{i!}t^i\right]dt$$

$$= \int_{0}^{1/2} t^{\lambda-1}\ln^p t\ln^q(1-t)\left[(1-t)^{\mu-1} - \sum_{i=0}^{r}\frac{(-1)^i(\mu-1)_i}{i!}t^i\right]dt$$

and

$$\begin{aligned}
&\lim_{\epsilon\to 0}\int_{1/2}^{1-\epsilon} \ln^p t(1-t)^{\mu-1}\ln^q(1-t)\left[t^{\lambda-1} - \sum_{i=0}^{s}\frac{(-1)^i(\lambda-1)_i}{i!}(1-t)^i\right]dt \\
&= \int_{1/2}^{1} \ln^p t(1-t)^{\mu-1}\ln^q(1-t)\left[t^{\lambda-1} - \sum_{i=0}^{s}\frac{(-1)^i(\lambda-1)_i}{i!}(1-t)^i\right]dt,
\end{aligned}$$

the integrals being convergent. Further, from the lemma we see that the neutrix limit of the function

$$\sum_{i=0}^{r} \frac{(-1)^i(\mu-1)_i}{i!} \int_{\epsilon}^{1/2} t^{\lambda+i-1} \ln^p t \ln^q(1-t)dt$$
$$+\sum_{i=0}^{s} \frac{(-1)^i(\lambda-1)_i}{i!} \int_{1/2}^{1-\epsilon} (1-t)^{\mu+i-1} \ln^p t \ln^q(1-t)dt$$

exists, implying that

$$\operatorname*{N-lim}_{\epsilon \to 0} \int_{\epsilon}^{1-\epsilon} t^{\lambda-1} \ln^p t(1-t)^{\mu-1} \ln^q(1-t)dt$$

exists. This proves the existence of the function $B_{p,q}(\lambda, \mu)$ for $p, q = 0, 1, 2, \ldots$ and all λ, μ.

Theorem 2

$$B_{p,q}(\lambda, \mu) = B_{q,p}(\mu, \lambda)$$

for $p, q = 0, 1, 2, \ldots$ *and all* λ, μ.

The proof of this theorem is trivial.

In the following, we now evaluate some particular values of $B_{p,q}(\lambda, \mu)$. In order to simplify the proofs, we note that

$$B_{p,q}(\lambda, \mu) = \operatorname*{N-lim}_{\epsilon \to 0} \int_{\epsilon}^{1} t^{\lambda-1} \ln^p t(1-t)^{\mu-1} \ln^q(1-t)dt$$

if $\mu > 0$, since the integral is then convergent in the neighbourhood of the point $t = 1$.

Theorem 3

$$B_{p,0}(0, 1) = 0$$

for $p = 1, 2, \ldots$.

Proof. We have

$$\int_{\epsilon}^{1} t^{-1} \ln^p t\, dt = -\frac{\ln^{p+1} \epsilon}{p+1}$$

and so

$$B_{p,0}(0, 1) = \operatorname*{N-lim}_{\epsilon \to 0} \int_{\epsilon}^{1} t^{-1} \ln^p t\, dt = 0$$

for $p = 1, 2, \ldots$.

Theorem 4

$$B_{p,0}(0,r+1) = \sum_{i=1}^{r} \binom{r}{i} \frac{(-1)^{p+i} p!}{i^{p+1}}$$

for $p, r = 1, 2, \dots$ *, where*

$$\binom{r}{i} = \frac{r!}{i!(r-i)!}.$$

Proof. We have

$$\int_{\epsilon}^{1} t^{-1} \ln^p t (1-t)^r dt = \int_{\epsilon}^{1} t^{-1} \ln^p t \, dt + \sum_{i=1}^{r} (-1)^i \binom{r}{i} \int_{\epsilon}^{1} t^{i-1} \ln^p t \, dt$$

and so

$$\begin{aligned} B_{p,0}(0,r+1) &= \operatorname*{N-lim}_{\epsilon \to 0} \int_{\epsilon}^{1} t^{-1} \ln^p t (1-t)^r dt \\ &= B_{p,0}(0,1) + \sum_{i=1}^{r} (-1)^i \binom{r}{i} \int_{0}^{1} t^{i-1} \ln^p t \, dt \\ &= \sum_{i=1}^{r} \binom{r}{i} \frac{(-1)^{p+i} p!}{i^{p+1}}, \end{aligned}$$

for $p, r = 1, 2, \dots$, since it is easily proved that

$$\int_{0}^{1} t^i \ln^p t \, dt = \frac{(-1)^p p!}{(i+1)^{p+1}}, \tag{2}$$

for $i = 0, 1, 2, \dots$.

Theorem 5

$$B_{p,0}(-n,1) = -\frac{p!}{n^{p+1}} \tag{3}$$

for $p, n = 0, 1, 2, \dots$.

Proof. Integrating by parts we have

$$\int_{\epsilon}^{1} t^{-n-1} \ln t dt = n^{-1} \epsilon^{-n} \ln \epsilon + n^{-1} \int_{\epsilon}^{1} t^{-n-1} dt$$

and so

$$B_{1,0}(-n,1) = \underset{\epsilon\to 0}{\text{N-lim}} \int_\epsilon^1 t^{-n-1}\ln t\,dt = -n^{-2}$$

proving equation (3) for $p=1$ and $n=1,2,\ldots$.

More generally we have

Theorem 6

$$B_{p,0}(-n,r+1) = \begin{cases} \sum_{i=0}^{r}\binom{r}{i}\frac{(-1)^{i+1}p!}{(n-i)^{p+1}}, & r<n, \\ \sum_{i=0}^{n-1}\binom{r}{i}\frac{(-1)^{i+1}p!}{(n-i)^{p+1}}, & r=n \end{cases} \tag{4}$$

for $p,n=1,2,\ldots$ and $r=0,1,\ldots,n$ and

$$B_{p,0}(-n,r+1) = \sum_{\substack{i=0\\ i\neq n}}^{r}\binom{r}{i}\frac{(-1)^{i+1}p!}{(n-i)^{p+1}} \tag{5}$$

for $p,n=1,2,\ldots$ and $r=n+1,n+2,\ldots$.

Proof. We have

$$\int_\epsilon^1 t^{-n-1}\ln^p t(1-t)^r dt = \sum_{i=0}^{r}(-1)^i\binom{r}{i}\int_\epsilon^1 t^{i-n-1}\ln^p t\,dt \tag{6}$$

and so

$$\begin{aligned} B_{p,0}(-n,r+1) &= \sum_{i=0}^{r}(-1)^i\binom{r}{i}\underset{\epsilon\to 0}{\text{N-lim}}\int_\epsilon^1 t^{i-n-1}\ln^p t\,dt \\ &= \sum_{i=0}^{r}(-1)^i\binom{r}{i}B_{p,0}(i-n,1) \end{aligned}$$

for $r=0,1,\ldots,n$. Equation (4) follows on using Theorems 3 and 5.

When $r\geq n+1$, equation (6) again holds, but this time we have

$$B_{p,0}(-n,r) = \sum_{i=0}^{n}(-1)^i\binom{r}{i}B_{p,0}(i-n,1) + \sum_{i=n+1}^{r}(-1)^i\binom{r}{i}\int_0^1 t^{i-n-1}\ln^p t\,dt$$

and equation (5) follows on using Theorems 3 and 5 and equation (2).

Theorem 7

$$B_{p,0}(0,0) = B_{p,0}(1,0) = (-1)^p p! \zeta(p+1), \tag{7}$$

$$B_{p,0}(-1,0) = -p! + (-1)^p p! \zeta(p+1), \tag{8}$$

for $p = 1, 2, \ldots$, *where*

$$\zeta(p) = \sum_{i=1}^{\infty} i^{-p}, \qquad (p > 1)$$

denotes the zeta function.

Proof. We have

$$\int_{\epsilon}^{1-\epsilon} t^{-1} \ln^p t (1-t)^{-1} dt = \int_{\epsilon}^{1-\epsilon} [t^{-1} + (1-t)^{-1}] \ln^p t \, dt,$$

and so

$$B_{p,0}(0,0) = B_{p,0}(0,1) + B_{p,0}(1,0) = B_{p,0}(1,0)$$

on using Theorem 3.

Further

$$\begin{aligned} B_{p,0}(1,0) &= \operatorname*{N-lim}_{\epsilon \to 0} \int_0^{1-\epsilon} (1-t)^{-1} \ln^p t \, dt \\ &= \int_0^1 (1-t)^{-1} \ln^p t \, dt \\ &= \sum_{i=0}^{\infty} \int_0^1 t^i \ln^p t \, dt \\ &= \sum_{i=0}^{\infty} \frac{(-1)^p p!}{(i+1)^{p+1}} \\ &= (-1)^p p! \zeta(p+1), \end{aligned}$$

on using equation (2), proving equations (7).

To prove equation (8), we note that

$$B_{p,0}(-1,0) = \operatorname*{N-lim}_{\epsilon \to 0} \int_{\epsilon}^{1-\epsilon} t^{-2} \ln^p t (1-t)^{-1} dt$$

$$
\begin{aligned}
&= \operatorname*{N-lim}_{\epsilon \to 0} \int_{\epsilon}^{1-\epsilon} [t^{-2} + t^{-1} + (1-t)^{-1}] \ln^p t \, dt \\
&= B_{p,0}(-1,1) - B_{p,0}(0,1) + B_{p,0}(1,0) \\
&= -p! + (-1)^p p! \zeta(p+1)
\end{aligned}
$$

on using Theorems 3 and 5 and equations (7).

Theorem 8

$$
B_{0,1}(-n, r+1) = \begin{cases} \displaystyle\sum_{i=0}^{r} \frac{(-1)^i}{n-i} \binom{r}{i} \left[\phi(n-i) - \frac{2}{n-i}\right], & r < n, \\ \displaystyle\sum_{i=0}^{n-1} \frac{(-1)^i}{n-i} \binom{r}{i} \left[\phi(n-i) - \frac{2}{n-i}\right] - (-1)^n \zeta(2), & r = n, \end{cases} \tag{9}
$$

for $n = 1, 2, \ldots,$ *and* $r = 0, 1, \ldots, n$ *and*

$$
\begin{aligned}
B_{0,1}(-n, r+1) &= \sum_{i=0}^{n-1} \frac{(-1)^i}{n-i} \binom{r}{i} \left[\phi(n-i) - \frac{2}{n-i}\right] \\
&\quad -(-1)^n \binom{r}{n} \zeta(2) - \sum_{i=n+1}^{r} \frac{(-1)^i}{i-n} \binom{r}{i} \phi(i-n),
\end{aligned} \tag{10}
$$

for $n = 1, 2, \ldots$ *and* $r = n+1, n+2, \ldots,$ *where*

$$
\phi(n) = \begin{cases} 0, & n = 0 \\ \displaystyle\sum_{i=1}^{n} \frac{1}{i}, & n \geq 1. \end{cases}
$$

Proof. We have

$$
\int_{\epsilon}^{1} t^{-n-1} \ln(1-t) dt = n^{-1} \epsilon^{-n} \ln(1-\epsilon) - n^{-1} \int_{\epsilon}^{1} t^{-n} (1-t)^{-1} dt
$$

and it follows that

$$
\begin{aligned}
B_{0,1}(-n, 1) &= -n^{-2} - n^{-1} B(-n+1, 0) \\
&= -n^{-2} + n^{-1} \phi(n-1) \\
&= n^{-1} [\phi(n) - 2n^{-1}],
\end{aligned} \tag{11}
$$

since it was proved in [2] that

$$B(-n,-r) = -\frac{(n+r)!}{n!r!}[\phi(n)+\phi(r)-2\phi(n+r)], \tag{12}$$

for $n, r = 0, 1, 2, \ldots$. Equation (9) is therefore proved for the case $r = 0$ and $n = 1, 2, \ldots$. More generally we have

$$\int_\epsilon^1 t^{-n-1}\ln(1-t)(1-t)^r dt = \sum_{i=0}^r (-1)^i \binom{r}{i} \int_\epsilon^1 t^{i-n-1}\ln(1-t)dt$$

and it follows that

$$B_{0,1}(-n, r+1) = \sum_{i=0}^r (-1)^i \binom{r}{i} B_{0,1}(-n+i, 1). \tag{13}$$

Equations (9) now follow on using equations (7) and (11).

To prove equation (10) we note that

$$\begin{aligned} B_{0,1}(s,1) &= \frac{\partial}{\partial\mu}\frac{\Gamma(\lambda)\Gamma(\mu)}{\Gamma(\lambda+\mu)}\Bigg]_{\substack{\lambda=s\\ \mu=1}} \\ &= \frac{\Gamma'(1)}{s} - \frac{\Gamma'(s+1)}{ss!} \\ &= -s^{-1}\phi(s) \end{aligned}$$

and so

$$B_{0,1}(-n+i,1) = -(i-n)^{-1}\phi(i-n),$$

for $i = n+1, n+2, \ldots$. Equation (10) now follows from equation (13).

Theorem 9

$$B_{1,0}(-n,0) = -\sum_{i=1}^n i^{-2} - \zeta(2), \tag{14}$$

$$B_{1,0}(-n+1,-1) = -n\sum_{i=1}^n i^{-2} - n\zeta(2) - 1 + \phi(n), \tag{15}$$

for $n = 1, 2, \ldots$.

Proof. We have

$$\int_{\epsilon}^{1-\epsilon} t^{-2}\ln t(1-t)^{-1}dt = \int_{\epsilon}^{1-\epsilon} t^{-1}\ln t[t^{-1}+(1-t)^{-1}]dt$$

and so

$$\begin{aligned} B_{1,0}(-1,0) &= B_{1,0}(-1,1)+B_{1,0}(0,0) \\ &= -1-\zeta(2), \end{aligned}$$

on using equations (3) and (7). Equation (14) is therefore proved for the case $n=1$. Now assume that equation (14) holds for some positive n. Then

$$\int_{\epsilon}^{1-\epsilon} t^{-n-2}\ln t(1-t)^{-1}dt = \int_{\epsilon}^{1-\epsilon} t^{-n-1}\ln t[t^{-1}+(1-t)^{-1}]dt$$

and so

$$\begin{aligned} B_{1,0}(-n-1,0) &= B_{1,0}(-n-1,1)+B_{1,0}(-n,0) \\ &= -\sum_{i=1}^{n+1} i^{-2}-\zeta(2), \end{aligned}$$

on using equation (4) and our assumption. Equation (14) now follows by induction for $n=1,2,\dots$.

To prove equation (15) we note that

$$\begin{aligned} \int_{\epsilon}^{1-\epsilon} t^{-n-1}\ln t(1-t)^{-1}dt &= -n^{-1}\int_{\epsilon}^{1-\epsilon} \ln t(1-t)^{-1}dt^{-n} \\ &= -n^{-1}\left[t^{-n}\ln t(1-t)^{-1}\right]_{\epsilon}^{1-\epsilon} \\ &\quad +n^{-1}\int_{\epsilon}^{1-\epsilon}[t^{-1}(1-t)^{-1}+\ln t(1-t)^{-2}]t^{-n}dt \end{aligned}$$

and so

$$\begin{aligned} B_{1,0}(-n,0) &= n^{-1}[1+B(-n,0)+B_{1,0}(-n+1,-1)] \\ &= n^{-1}[1-\phi(n)+B_{1,0}(-n+1,-1)] \end{aligned}$$

on using equation (12). Equation (15) follows on using equation (14).

Theorem 10

$$B_{1,0}(-n,-r) = -\frac{1}{n!r!}\sum_{j=0}^{r-1}\sum_{i=0}^{j}\frac{(n+r+i-j-1)!j!}{(j-i+1)i!}$$

$$-\frac{(n+r)!}{n!r!}\left\{\sum_{i=n+1}^{n+r}\frac{\phi(i)}{i}+\sum_{i=0}^{r-1}\frac{\phi(i)}{n+r-i}-2\phi(n+r)[\phi(n+r)-\phi(n)]+\sum_{i=1}^{n+r}i^{-2}+\zeta(2)\right\} \tag{16}$$

for $n, r = 1, 2, \ldots$.

Proof. We note first of all that equation (16) holds for $r = 1$ and $n = 1, 2, \ldots$ by equation (15). We therefore assume that equation (16) holds for some $r - 1$ and $n = 1, 2, \ldots$.
We have

$$\begin{aligned}\int_{\epsilon}^{1-\epsilon} t^{-n-1}\ln t(1-t)^{-r-1}dt &= r^{-1}\int_{\epsilon}^{1-\epsilon} t^{-n-1}\ln t\, d(1-t)^{-r} \\ &= r^{-1}\left[t^{-n-1}\ln t(1-t)^{-r}\right]_{\epsilon}^{1-\epsilon} \\ &\quad -r^{-1}\int_{\epsilon}^{1-\epsilon}\left[t^{-n-2}-(n+1)t^{-n-2}\ln t\right](1-t)^{-r}\,dt,\end{aligned}$$

where

$$\operatorname*{N-lim}_{\epsilon\to 0}\,(1-\epsilon)^{-n-1}\ln(1-\epsilon)\epsilon^{-r} = -\sum_{i=0}^{r-1}\frac{(n+i)!}{(r-i)i!n!},$$

$$\operatorname*{N-lim}_{\epsilon\to 0}\,\epsilon^{-n-1}\ln\epsilon(1-\epsilon)^{-r} = 0$$

and so

$$\begin{aligned}B_{1,0}(-n,-r) &= -\sum_{i=0}^{r-1}\frac{(n+i)!}{r(r-i)i!n!} - r^{-1}B(-n-1,-r+1) \\ &\quad +\frac{n+1}{r}B_{1,0}(-n-1,-r+1)\end{aligned}$$

$$= -\sum_{i=0}^{r-1}\frac{(n+i)!}{r(r-i)i!n!} - \frac{(n+r)!}{r!(n+1)!}[\phi(n+1)+\phi(r-1)-2\phi(n+r)]$$

$$-\frac{1}{n!r!}\sum_{j=0}^{r-2}\sum_{i=0}^{j}\frac{(n+r+i-j+1)!j!}{(j-i+1)i!}$$

$$-\frac{(n+r)!}{n!r!}\left\{\sum_{i=n+2}^{n+r}\frac{\phi(i)}{i} + \sum_{i=0}^{r-2}\frac{\phi(i)}{n+r-i} - 2\phi(n+r)[\phi(n+r)-\phi(n+1)] + \sum_{i=1}^{n+r} i^{-2} + \zeta(2)\right\}$$

on using our assumption. This equation can be rearranged to give equation (16) which now follows by induction.

References

[1] **van der Corput, J.G.**: *Introduction to the neutrix calculus.* J. Analyse Math. **7**, 291–398 (1959-60)

[2] **Fisher, B.**, and **Kuribayashi, Y.**: *Neutrices and Beta function.* Rostock. Math. Kolloq. **32**, 56-66 (1987)

[3] **Gel'fand, I.M.**, and **Shilov, G.E.**: *Generalized Functions.* Vol. I. New York 1964

received: December 21, 1990

Authors:

Dr. E. Özçağ
Dr. B. Fisher
Department of Mathematics
The University
Leicester, LE1 7RH
England

Rostock. Math. Kolloq. 45, 57–74 (1991)

Subject Classification (AMS)
62F05, 62G10

LUDWIG HOTHORN; FRIEDRICH LIESE

Adaptive Umbrellatests – Simulationsuntersuchungen

1 Tests im k-Stichprobenproblem und deren Effizienz

Allen Betrachtungen in dieser Arbeit liegt das folgende Modell zugrunde, das gegebenenfalls noch spezialisiert wird. Wir setzen voraus, daß die $\underline{x}_{i,j}$ unabhängige Zufallsvariablen sind, die bei festem $1 \leq i \leq k$ für $1 \leq j \leq n_j$ identisch verteilt sind und die Verteilungsfunktion F_i haben mögen. Dabei deuten wir $\underline{x}_{i,j}$ als j-te Messung in der i-ten Patientengruppe (bzw. i-ten Behandlungsmethode, i-ten Dosisgruppe). Die Nullhypothese lautet $H_0 \ : \ F_1 = F_2 = \ldots = F_k$, während die globale Alternative darin besteht, daß ein Paar (i, j) existiert mit $F_i \neq F_j$:

$$H_A^g \ : \ \exists (i,j) : F_i \neq F_j. \tag{1}$$

Diese Alternative ist sehr groß, und es existieren deshalb keine Tests mit befriedigender Güte in allen Teilgebieten der Alternative. In vielen Fällen liegen aber Zusatzinformationen über mögliche Alternativen vor. Für deren Formulierung verwenden wir die stochastische Ordnung zwischen zwei Verteilungsfunktionen: F heißt stochastisch kleiner oder gleich G, (in Zeichen $F \leq G$), falls $1 - F(x) \leq 1 - G(x)$ für alle reellen x gilt. Hat die Zufallsvariable $\underline{x}$ die Verteilungsfunktion F und bezeichnet G die Verteilungsfunktion von $\underline{x} + a$, wobei a eine reelle Zahl ist, so gilt

$$F \leq G \longleftrightarrow a \geq 0. \tag{2}$$

Eine oft benutzte Einschränkung der globalen Alternative ist die Trendalternative (ordered alternative)

$$H_A^t \ : \ F_1 \leq F_2 \leq \ldots \leq F_k \qquad \text{mit} \qquad F_1 \neq F_k. \tag{3}$$

Dagegen spricht man von einer Umbrella-Alternative, falls ein Umschlagpunkt l mit $1 \leq l \leq k$ existiert:

$$H_A^u \;:\; F_1 \leq \ldots \leq F_l \geq F_{l+1} \geq \ldots \geq F_k \qquad \text{mit} \qquad F_1 \neq F_l. \tag{4}$$

Gehen wir von einem linearen Modell

$$\underline{x}_{i,j} = \mu_i + \underline{\varepsilon}_{i,j}, \;\; 1 \leq j \leq n_i, \;\; 1 \leq i \leq k, \tag{5}$$

mit unabhängig und identisch verteileten $\underline{\varepsilon}_{i,j}$ aus, so lautet die globale Alternative

$$H_A^g \;:\; \exists(i,j) \;:\; \mu_i \neq \mu_j \tag{1'}$$

und wegen (2) die Trendalternative

$$H_A^t \;:\; \mu_1 \leq \ldots \leq \mu_k \qquad \text{mit} \qquad \mu_1 \neq \mu_k. \tag{3'}$$

bzw. die Umbrella-Alternative

$$H_A^u \;:\; \mu_1 \leq \ldots \leq \mu_l \geq \mu_{l+1} \geq \ldots \geq \mu_k \qquad \text{mit} \qquad \mu_1 \neq \mu_l. \tag{4'}$$

Für die Alternativen (1), (1'), (3), (3'), (4), (4') gibt es eine große Anzahl von speziellen Tests. Für Übersichten verweisen wir z.B. auf [2], [8] und [23].

Zur Konstruktion adaptiver Tests gehen wir zunächst wie in ROTHE [19] vor. Deshalb nehmen wir an, daß $\underline{\varepsilon}_{i,j}$ unabhängig und identisch normalverteilt sind mit Erwartungswert Null und Varianz $\sigma^2 > 0$. Weiter sei

$$N = \sum_{i=1}^{k} n_i, \qquad \tau = (\tau_1, \ldots \tau_k) \qquad \text{mit} \qquad \tau_i = n_i/N$$

und

$$\mu_{\tau,i} = \mu_i - \sum_{i=1}^{k} \tau_i \mu_i \qquad \text{sowie}$$

$$\underline{x}_{i.} = 1/n_i \sum_{j=1}^{n_i} \underline{x}_{i,j}, \qquad\qquad \underline{x}_{..} = 1/k \sum_{i=1}^{k} \underline{x}_{i.}.$$

Die globale parametrische Alternative (1') läßt sich durch den F-Test testen, der die Nullhypothese für große Werte von

$$\underline{T} = \frac{\frac{1}{k-1} \sum\limits_{i=1}^{k} n_i(\underline{x}_{i.} - \underline{x}_{..})^2}{\frac{1}{N-k} \sum\limits_{i=1}^{k} \sum\limits_{j=1}^{n_i} (\underline{x}_{i,j} - \underline{x}_{i.})^2}$$

ablehnt. $\underline{T}$ besitzt unter der Modellvoraussetzung (5) eine F-Verteilung mit (k-1, N-k) Freiheitsgraden und den Nichtzentralitätsparameter

$$\delta^2 = \frac{1}{\sigma^2}\sum_{i=1}^{k} n_i \mu_{\tau,i}^2.$$

Sei nun $b = (b_1, \ldots, b_k)$ ein beliebiger Vektor. Wir setzen $b_{\tau,i} = b_i - \sum_{l=1}^{k} \tau_l b_l$ und definieren die Teststatistik $\underline{T}^{(b)}$ als verallgemeinerte t-Statistik analog [19]:

$$\underline{T}^{(b)} = \frac{\sum\limits_{i=1}^{k} b_{\tau,i} n_i (\underline{x}_{i.} - \underline{x}_{..})}{\left(\frac{1}{N-k}\sum\limits_{i=1}^{k}\sum\limits_{j=1}^{n_i}(\underline{x}_{i,j} - \underline{x}_{i.})^2 \sum\limits_{i=1}^{k} n_i b_{\tau,i}^2\right)^{1/2}}.$$

Unter der Modellvoraussetzung (5) besitzt $\underline{T}^{(b)}$ eine t-Verteilung mit $N-k$ Freiheitsgraden und den Nichtzentralitätsparameter

$$\sigma_b = \frac{1}{\sigma}\frac{\sum\limits_{i=1}^{k} n_i b_{\tau,i}\mu_{\tau,i}}{(\sum\limits_{i=1}^{k} n_i b_{\tau,i}^2)^{1/2}}.$$

Weil die nichtzentrale t-Verteilung monoton von $\sum_{i=1}^{k} n_i b_{\tau,i}\mu_{\tau,i}$ abhängt, ergibt sich, daß auf der Alternative $\mu = (\mu_1, \ldots, \mu_k)$ die Teststatistik $\underline{T}^{(b^0)}$ mit $b_{\tau,i}^0 = c\mu_{\tau,i}$, wobei c ein beliebiger Faktor ist, in der Klasse der Teststatistiken $\underline{T}^{(b)}$ die größte Güte liefert. Wie in [19] ausgeführt wurde, liefert darüber hinaus ein Vergleich zwischen den auf T bzw. $T^{(b^0)}$ beruhenden Tests, indem man einschlägige Tafelwerke verwendet, daß der durch $\underline{T}^{(b^0)}$ definierte Test auf der Alternative $\mu = (\mu_1, \ldots, \mu_k)$ eine größere Güte als der F-Test besitzt. Einerseits ist eine formelmäßige Behandlung dieser Frage sehr aufwendig. Andererseits ist bekannt, daß beide Tests gute Robustheitseigenschaften haben gegenüber der Verletzung der Voraussetzung, daß die $\underline{\varepsilon}_{i,j}$ normalverteilt sind, was auf der Wirkung des zentralen Grenzwertsatzes beruht. Darüber hinaus ist ein Gütevergleich für nicht normalverteilte $\underline{\varepsilon}_{i,j}$ überhaupt nicht mehr möglich, weil die Verteilungen von $\underline{T}$ und $\underline{T}^{(b)}$ nicht mehr zugänglich sind. Die genannten Schwierigkeiten lassen sich aber durch asymptotische Effizienzbetrachtungen überwinden.
Es sei $G(\alpha, \mu, N)$ bzw. $G^b(\alpha, \mu, N)$ die Güte des Tests, der die Nullhypothese $H_0 : \mu_1^0 = \ldots = \mu_k^0$ für große Werte von $\underline{T}$ bzw. $\underline{T}^{(b)}$ ablehnt, wobei $\mu = (\mu_1, \ldots, \mu_k)$ aus der Alternative ist. Wie in [19] setzen wir für $\sigma^2 > 0$ und $0 < \alpha < \beta < 1$

$$\begin{aligned} N(\alpha, \beta, \mu, \delta^2) &= \mathrm{Min}(N : G(\alpha, \mu, N) \geq \beta), \\ N_b(\alpha, \beta, \mu, \delta^2) &= \mathrm{Min}(N : G^b(\alpha, \mu, N) \geq \beta). \end{aligned}$$

Die exakte relative Effizienz des durch $\underline{T}^{(b)}$ definierten Tests gegenüber dem F-Test ist definiert durch

$$e_b(\alpha,\beta,\mu,\delta^2) = \frac{N(\alpha,\beta,\mu,\delta^2)}{N_b(\alpha,\beta,\mu,\delta^2)}.$$

Wie bereits ausgeführt wurde, ist eine direkte Berechnung von $e_b(\alpha,\beta,\mu,\delta^2)$ sehr aufwendig und für nicht normale Grundgesamtheiten praktisch unmöglich. Entsprechend dem Effizienzkonzept nach PITMAN [17] untersucht man jetzt $\lim_{\mu\to\mu^0} e(\alpha,\beta,\mu,\delta^2)$, wobei μ^0 ein Punkt aus der Nullhypothese ist, μ in der Alternative liegt und sich der Nullhypothese im Punkt μ^0 nähert. Da sich bei festem N die Güte der Tests bei Annäherung an die Nullhypothese verschlechtert, ergibt sich aus $\mu \to \mu^0$, daß $N(\alpha,\beta,\mu,\delta^2)$ und $N_b(\alpha,\beta,\mu,\delta^2)$ gegen unendlich streben. Wir nehmen an, daß die Anzahl der Messungen $n_i(N)$ so gegen unendlich strebt, daß

$$\lim_{N\to\infty} \frac{n_i(N)}{N} = \tau_i > 0$$

gilt. Zur Formulierung des folgenden Satzes bezeichnen wir für $0 < \alpha < \beta < 1$ durch $d^2(\alpha,\beta,k)$ denjenigen eindeutig bestimmten Nichtzentralitätsparameter, so daß das α-Quantil einer zentralen χ^2-Verteilung mit k Freiheitsgraden und das β-Quantil einer χ^2-Verteilung mit k Freiheitsgraden und dem Nichtzentralitätsparameter $d^2(\alpha,\beta,k)$ übereinstimmen. Weiterhin sei für $\tau = (\tau_1,\ldots,\tau_k)$, $\mu = (\mu_1,\ldots,\mu_k)$ und $b = (b_1,\ldots,b_k)$

$$R(\tau,\mu,b) = \frac{\sum_{i=1}^{k} \tau_i \mu_{\tau,i} b_{\tau,i}}{(\sum_{i=1}^{k} \tau_i \mu_{\tau,i}^2)^{1/2} (\sum_{i=1}^{k} \tau_i b_{\tau,i}^2)^{1/2}}.$$

Der folgende Satz wurde in [19] bewiesen.

Satz 1 *Die Zufallsvariablen $\underline{x}_{i,j}$, $1 \le j \le n_i$, $1 \le i \le k$, seien durch das Modell* (5) *definiert, wobei die $\underline{\varepsilon}_{i,j}$ unabhängig und identisch verteilt sind und positive Varianz σ^2 besitzen. Dann gilt*

$$\lim_{\mu\to\mu^0} \frac{e_b(\alpha,\beta,\mu,\delta^2)}{R^2(\tau,\mu,b)} = \frac{d^2(\alpha,\beta,k-1)}{(\Phi^{-1}(\alpha) - \Phi^{-1}(\beta))^2}, \tag{6}$$

wobei μ^0 aus der Nullhypothese fest gewählt ist und μ aus der Alternative (1') ist.

Bemerkung 1 Weil die Effizienzbetrachtungen asymptotisch durchgeführt werden, wird die Normalverteiltheit der $\underline{\varepsilon}_{i,j}$ in Satz 1 nicht benötigt. Satz 1 wurde von ROTHE in [19] nur für μ aus der Alternative (3′) bewiesen. Eine Analyse des Beweises zeigt jedoch, daß die Behauptung auch für die Alternative (1′) gilt.

Bemerkung 2 Auf Grund der Schwarzschen Ungleichung gilt $R^2(\tau,\mu,\beta) \leq 1$, wobei das Gleichheitszeichen genau dann steht, wenn $b_{\tau,i} = c\mu_{\tau,i}$ mit einer Konstanten c für alle $1 \leq i \leq k$ ist. In diesem Falle hat der durch $\underline{T}^{(b)}$ definierte Test auf der Alternative μ die höchste asymptotische Effizienz. Bezeichnet $g(\alpha,\beta)$ die rechte Seite in (6), so gilt nach [19] die Beziehung $\lim_{\alpha\to 0} g(\alpha,\beta) = 1$. Sind jetzt die b_i beliebig gewählt, so gilt $R^2(\tau,\alpha,b) < 1$, und deshalb gibt es ein hinreichend kleines α, so daß der durch $\underline{T}^{(b)}$ definierte Test schlechter als der F-Test ist.

Zu Satz 1 gibt es ein nichtparametrisches Analogon. Es sei $\Psi : [0,1] \to R$ eine meßbare Funktion. Weiterhin sei für jedes $N = 1,2,\ldots$ auf $\{1,\ldots,N+1\}$ eine reelle Funktion $a^{(N)}$ definiert. Dabei möge gelten

$$\int_0^1 \Psi^2(t)\,dt < \infty, \tag{7}$$

$$\int_0^1 \Psi^2(t)\,dt - \left(\int_0^1 \Psi(t)\,dt\right)^2 > 0, \tag{8}$$

$$\lim_{N\to\infty} \int_0^1 (a^{(N)}(1+[tN]) - \Psi(t))^2\,dt = 0. \tag{9}$$

Es bezeichne $\underline{R}_{i,j}$ den Rang von $\underline{x}_{i,j}$ in der „gepoolten“ Stichprobe. Unter der Nullhypothese besitzt der N-dimensionale zufällige Vektor $(\underline{R}_{1,1},\ldots,\underline{R}_{k,n_k})$ eine Gleichverteilung auf der Menge der $N!$ Permutationen. Für festes $a : (1,\ldots,N) \to R$ setzen wir

$$\underline{T}^{(a)} = \frac{\sum\limits_{i=1}^{k} \frac{1}{n_i} \sum\limits_{j=1}^{n_i} (a(\underline{R}_{i,j}) - a_.)^2}{\sum\limits_{i=1}^{N} (a(i) - a_.)^2},$$

wobei $a_. = \frac{1}{N}\sum_{i=1}^{N} a(i)$ gesetzt wurde. Für $a(i) = \frac{i}{N+1}$ erhält man die Statistik nach KRUSKAL/WALLIS [13]. Deshalb werden auch für beliebiges a die durch $\underline{T}^{(a)}$ definierten Tests als Tests vom KRUSKAL/WALLIS-Typ bezeichnet.

Analog zu $\underline{T}^{(b)}$ definieren wir $\underline{T}^{(b,a)}$ durch

$$\underline{T}^{(b,a)} = \frac{(\sum\limits_{i=1}^{k} \frac{b_i}{n_i} \sum\limits_{j=1}^{n_i} (a(\underline{R}_{i,j}) - a_.))}{(\frac{1}{N-1}\sum\limits_{j=1}^{N} (a(j) - a_.)^2 \sum\limits_{i=1}^{k} n_i b_{\tau,i}^2)^{1/2}}.$$

Betrachtet man jetzt statt einer festen Funktion a eine Folge von Funktionen $a^{(N)}$, so läßt sich die Effizienz $e_b^{a^{(N)}}(\alpha,\beta,\mu)$ analog zu $e_b(\alpha,\beta,\mu,\delta^2)$ einführen. Setzt man voraus, daß die Verteilungsfunktion F der $\underline{\varepsilon}_{i,j}$ im Modell (5) eine differenzierbare Dichte f hat und gilt $\int_0^1 \Psi_f^2(x)\,dx < \infty$ mit $\Psi_f = -\frac{f' \circ F^{-1}}{f \circ F^{-1}}$ als *score*-Funktion, so gilt nach [19]

Satz 2 *Unter den oben angegebenen Bedingungen an die Verteilung der $\underline{\varepsilon}_{i,j}$ und an die $a^{(N)}$ gilt*

$$\lim_{\mu\to\mu^0} \frac{e_b^{a^{(N)}}(\alpha,\beta,\mu)}{R^2(\tau,\mu,\beta)} = \frac{d^2(\alpha,\beta,k-1)}{(\Phi^{-1}(\alpha)-\Phi^{-1}(\beta))^2},$$

wobei μ,μ_0 und $d^2(\alpha,\beta,k-1)$ die gleiche Bedeutung wie in Satz 1 haben.

Eine analoge Aussage wie in Bemerkung 2 zu Satz 1 gilt auch im nichtparametrischen Fall auf Grund von Satz 2. Dies bedeutet insbesondere, daß der Test nach JONCKHEERE [11], für den $a(\underline{R}_{i,j}) = \underline{R}_{i,j}$, $b_i = i$ gilt, schlechter als der KRUSKAL/WALLIS-Test sein kann. Dies gilt für hinreichend kleines α auf Grund der Struktur von $R(\tau,\mu,\beta)$ dann, wenn die Alternative $\mu = (\mu_1 \ldots, \mu_k)$ stark von einem linearen Verlauf abweicht. Dann sind die linear anwachsenden Gewichte $b_i = i$ nicht angemessen.

2 Adaptive Tests

Wie die im vorigen Abschnitt dargestellten Effizienzbetrachtungen nachwiesen, ist die Verwendung der Teststatistiken $\underline{T}^{(b)}$ bzw. $\underline{T}^{(b,a)}$ gegenüber $\underline{T}$ bzw. $\underline{T}^{(a)}$ dann vorteilhaft, wenn man die mögliche Alternative als Punkt im Alternativbereich kennt. Diese Situation ist aber für die meisten praktischen Anwendungen unrealistisch. Deshalb versucht man, mit geschätzten Koeffizienten zu arbeiten. Tests mit geschätzten Koeffizienten werden auch als *adaptive Tests* bezeichnet.

Wir nehmen an, daß das Modell (5) mit normalverteilten und unabhängigen $\underline{\varepsilon}_{i,j}$ vorliegt. Da die optimale Teststatistik $\underline{T}^{(b)}$ einen Koeffizientenvektor $b = (b_1,\ldots,b_k)$ mit $b_{\tau,i} = c\mu_{\tau,i}$ hat, mit einer Konstanten c, die die Struktur der Statistik $\underline{T}^{(b)}$ nicht beeinflußt, setzen wir $\hat{b}_i = \underline{x}_{i.}$ Dann gilt $\hat{b}_{\tau,i} = \underline{x}_{i.} - \underline{x}_{..}$. Wir setzen jetzt in $\underline{T}^{(b)}$ die Schätzungen $\hat{b}_i$ ein und gehen somit zur Teststatistik $\underline{T}^{(\hat{b})}$ über:

$$\underline{T}^{(\hat{b})} = \frac{\sum\limits_{i=1}^{k} n_i(\underline{x}_{i.}-\underline{x}_{..})(\underline{x}_{i.}-\underline{x}_{..})}{\left(\frac{1}{N-k}\sum\limits_{i=1}^{k}\sum\limits_{j=1}^{n_i}(\underline{x}_{i,j}-\underline{x}_{i.})^2 \sum\limits_{i=1}^{k} n_i(\underline{x}_{i.}-\underline{x}_{..})^2\right)^{1/2}}.$$

Dann gilt aber $\underline{T}^{(\hat{b})} = \underline{T}^{1/2}$, wobei $\underline{T}$ die Teststatistik des F-Tests ist. Weil der Übergang von $\underline{T}$ zu $\underline{T}^{1/2}$ mit einer monotonen Funktion erfolgt, definieren beide Teststatistiken die gleichen α-Tests, wobei die Ablehnung jeweils für große Werte von $\underline{T}$ bzw. $\underline{T}^{1/2}$ erfolgt. Damit ist also der F-Test interpretierbar als ein Test mit einer verallgemeinerten t-Statistik mit geschätzten Gewichten. Gegenüber dem globalen F-Test ergeben sich durch dieses adaptive Herangehen keine Vorteile . Dies war eigentlich auch zu erwarten, weil in der Schätzung der Gewichte, also die Schätzung der Alternative, keinerlei Strukturvoraussetzungen über den Alternativbereich eingingen. Jetzt schränken wir diesen Bereich ein und bestimmen die Gewichte in Abhängigkeit von der Stichprobe so, daß der geschätzte Vektor auch im Alternativbereich liegt. Damit haben der Koeffizientenvektor und der Vektor μ, der die Alternative darstellt, einen ähnlichen Verlauf, und es besteht so die Hoffnung, daß man zu einer Teststatistik gelangt, die eine höhere Güte liefert als der globale F-Test. Geht man etwa von der Umbrella-Alternative (4′) mit bekanntem Umbrellapunkt 1 aus, so hat die von MACK und WOLFE in [16] eingeführte Statistik

$$Z^{-1}\Big(\sum_{1\leq i<j\leq 1} n_i n_j(\underline{x}_{j.} - \underline{x}_{i.}) + \sum_{l+1\leq i<j\leq k} n_i n_j(\underline{x}_{i.} - \underline{x}_{j.})\Big),$$

mit $Z = (\mathrm{MQR} \sum_{i=1}^{k} c_i^2/n_i)^{1/2}$ und c_i als Koeffizient bei $\underline{x}_{i.}$ im Zähler der obigen Statistik, gerade die Struktur $\underline{T}^{(b)}$ mit einem geeigneten Koeffizientenvektor b. Simulationsuntersuchungen in [5] und [6] haben gezeigt, daß solche Umbrellatests eine hohe Güte bei der Alternative (4′) aufweisen und dort insbesonere dem globalen F-Test überlegen sind. Da genaue Kenntnisse über den Umbrellapunkt in der Praxis selten sind, wird er in den folgenden Tests zunächst geschätzt. Dann wird der Koeffizientenvektor b in geeigneter Weise aus der zu (4′) gehörigen Menge $M_l = (\mu \ : \ \mu = (\mu_1, \ldots, \mu_k),\ \mu_1 \leq \ldots \leq \mu_l \geq \mu_{l+1} \geq \ldots \geq \mu_k)$ gewählt. In diesem Falle ist gesichert, daß sich die Vektoren b und $\mu = (\mu_1, \ldots, \mu_k)$ ähnlich sind, falls tatsächlich eine Alternative mit Umbrellapunkt l vorliegt. Auf Grund von Bemerkung 2 zu Satz 1 besteht die Hoffnung, so zu Tests mit hoher Güte auf der Umbrella-Alternative zu gelangen.

Für die Schätzung des Umbrellapunktes werden die Gruppenmittelwerte im parametrischen Modell und die mittleren Ränge im nichtparametrischen Modell benutzt. Wir setzen

$$\underline{l}_p \ = \ \operatorname*{Max}_{1\leq i\leq k} \ (i \ : \ \underline{x}_{i.} \geq \underline{x}_{j.} \ \text{ für alle } 2 \leq j \leq k), \tag{10}$$

$$\underline{l}_{np} \ = \ \operatorname*{Max}_{1\leq i\leq k} \ \Big(i \ : \ \frac{1}{n_i}\sum_{l=1}^{n_i} \underline{R}_{i,l} \geq \frac{1}{n_j}\sum_{l=1}^{n_j} \underline{R}_{j,l} \ \text{ für alle } 2 \leq j \leq k\Big). \tag{11}$$

Bei festem Wert von $\underline{l}_p$ bzw. $\underline{l}_{np}$ sollen jetzt die Koeffizienten $b_1 \ldots, b_k$ gewählt werden.

I. Wir nehmen an, daß die Umbrella-Alternative stückweise linear ist, d.h. bis l linear ansteigt und dann linear fällt. Dann setzen wir

$$b_{i,p}^{(I)} = \begin{cases} k-1-(\underline{l}_p - i) & \text{für } i = 1, \ldots, \underline{l}_p, \\ k-1+(\underline{l}_p - i) & \text{für } i = \underline{l}_p + 1, \ldots, k \end{cases} \tag{12}$$

und definieren $b_{i,np}^{(I)}$ analog unter Verwendung von $\underline{l}_{np}$.

Ist $\underline{l}_p$ bzw. $\underline{l}_{np}$ gleich 1 oder gleich k, so sind die $b_{i,p}^{(I)}$ bzw. $b_{i,np}^{(I)}$ gerade die Gewichtskoeffizienten, die zu einem parametrischen bzw. nichtparametrischen Trendtest gehören.

II. Die Koeffizienten $b_1, \ldots, b_k$ sollen jetzt in der Menge M_l nach einem Max-Min-Prinzip gewählt werden. Der Koeffizientenvektor $\underline{b}_p^{(II)}$ sei so definiert, daß

$$\operatorname*{Max}_{b \in M_l} \operatorname*{Min}_{\mu \in M_l} R(\tau, \mu, b) = \operatorname*{Min}_{\mu \in M_l} R(\tau, \mu, \underline{b}^{(II)}) \tag{13}$$

gilt, falls $\underline{l}_p$ den Wert l annimmt. Analog wird $\underline{b}_{np}^{(II)}$ mit Hilfe von $\underline{l}_{np}$ gewählt. Da $R(\tau, \mu, b)$ ein Maß für die Güte des zu $\underline{T}^{(b)}$ gehörenden Tests ist, wird durch diese Wahl von $\underline{b}_p^{(II)}$ bzw. $\underline{b}_{np}^{(II)}$ erreicht, daß für ungünstigste Parameterkonfiguration die Güte maximiert wird. Damit wird auf ganz M_l eine bestimmte Mindestgüte garantiert.

Wir führen für zwei Vektoren μ und b das Skalarprodukt ein durch $(\mu, b) = \sum_{i=1}^{k} \tau_i \mu_i b_i$. Offensichtlich kann man sich im Max-Min-Problem (13) auf Einheitsvektoren μ und b beschränken. Somit lautet das Max-Min-Problem mit Hilfe des neu eingeführten Skalarproduktes

$$\operatorname*{Min}_{\mu \in M_l, |\mu|=1} (\mu, b^{(II)}) = \operatorname*{Max}_{b \in M_l, |b|=1} \operatorname*{Min}_{\mu \in M_l, |\mu|=1} (\mu, b). \tag{14}$$

Da sich die Statistik $\underline{T}^{(b)}$ nicht ändert, wenn zu jeder Komponente des Vektors b der gleiche Wert addiert wird, können wir zusätzlich in (14) fordern, daß b einen linearen Kontrast bildet, also $\sum_{l=1}^{k} b_l = 0$ gilt. Lineare Kontraste, die das Max-Min-Problem (14) lösen, wurden in ABELSON and TUKEY [1] charakterisiert. Dort wurde allerdings vorausgetzt, daß (μ, b) das Euklidische Skalarprodukt ist, also $\tau_1 = \tau_2 = \ldots = \tau_k = 1$ gilt. Eine Analyse der Beweise zeigt jedoch, daß die Aussagen auch für den allgemeinen Fall richtig bleiben. In der genannten Arbeit ist auch ein Algorithmus zu finden, wie man b berechnen kann. Im Fall $k = 5$ und gleicher Stichprobenumfänge $n_i = \text{const.}$ wurden die entsprechenden Rechnungen in [12] durchgeführt. Dabei ergeben sich die Koeffizienten der Tabelle 1.

	$l=1$	$l=2$	$l=3$	$l=4$	$l=5$
$b_1^{(II)}$	a	b	$-a$	$-a$	$-a$
$b_2^{(II)}$	b	a	$-b$	$-b$	$-b$
$b_3^{(II)}$	0	0	c	0	0
$b_4^{(II)}$	$-b$	$-b$	$-b$	a	b
$b_5^{(II)}$	$-a$	$-a$	$-a$	b	a

mit $a = 0.8944,\ b = 0.2010,\ c = 2.1908$

Tabelle 1

Für das parametrische Modell (5) ergeben sich also zwei adaptive Teststatistiken durch folgenden Algorithmus:

(a) Der Umbrellapunkt wird mit $\underline{l}_p$ aus (10) geschätzt.

(b) Bei festliegendem Wert von $\underline{l}_p$ wird der optimale Koeffizientenvektor $\underline{b}^{(I)}$ bzw. $\underline{b}^{(II)}$ gemäß (12) mit $k = 5$ bzw. Tabelle 1 ermittelt und in die Teststatistik $\underline{T}^{(b)}$ eingesetzt.

Die durch (a) und (b) definierten Teststatistiken bezeichnen wir mit $\underline{T}^{(\underline{b}(I))}$ bzw. $\underline{T}^{(\underline{b}(II))}$.

Analog kann man nichtparametrische adaptive Teststatistiken konstruieren, die durch Satz 2 motiviert sind und auf den Rängen der gepoolten Stichprobe beruhen. In [12] wurde folgende Teststatistik $\underline{T}_{np}^{(b)}$ betrachtet, die sich von der im Satz 2 benutzten nur durch eine andere Normierung unterscheidet:

$$\underline{T}_{np}^{(b)} = \frac{\sum\limits_{i=1}^{k} n_i b_i (\underline{R}_{i.} - \underline{R}_{..})}{\left(\frac{1}{N-k} \sum\limits_{i=1}^{k} \sum\limits_{j=1}^{n_i} (\underline{R}_{i,j} - \underline{R}_{i.}^2) \sum\limits_{i=1}^{k} n_i b_i^2 \right)^{1/2}}.$$

Zu den beiden nichtparametrischen Teststatistiken $\underline{T}_{np}^{(\underline{b}(I))}$ bzw. $\underline{T}_{np}^{(\underline{b}(II))}$ gelangt man so:

(a) Man schätze den Umbrellapunkt l mit $\underline{l}_{np}$ aus (11).

(b) Bei festem Wert von $\underline{l}_{np}$ wird der Koeffizientenvektor $\underline{b}^{(I)}$ bzw. $\underline{b}^{(II)}$ gemäß (12) oder (15) gewählt, was zur Teststatistik $\underline{T}_{np}^{(\underline{b}(I))}$ bzw. $\underline{T}_{np}^{(\underline{b}(II))}$ führt.

Die in den Schritten (a) und (b) konstruierten parametrischen bzw. nichtparametrischen adaptiven Teststatistiken sind ganz auf die speziellen Alternativen zugeschnitten und haben dort, wie später noch genauer ausgeführt wird, eine große Güte. Ihr Nachteil besteht darin, daß die zur Konstruktion der Tests notwendigen Quantile unter der Nullhypothese nicht zugänglich sind. Der ursprünglichen Vermutung, daß die konstruierten Teststatistiken unter der Nullhypothese näherungsweise t- bzw. normalverteilt sind, widersprechen Simulationsuntersuchungen. Bei Verwendung von Quantilen einer t-Verteilung mit $N-k$ Freiheitsgraden zur Konstruktion der für den Test notwendigen kritischen Werte, wurde bei den durchgeführten Simulationsuntersuchungen der vorgegebene Fehler erster Art wesentlich überschritten. Die Ursache ist darin zu sehen, daß durch die Auswahl eines Umbrellapunktes im Schritt (a) unter der Nullhypothese, wo es ja in Wirklichkeit keinen Umbrellapunkt gibt, in die Teststatistik noch eine zusätzliche Variabilität hineinkommt, die zu einer wesentlich breiteren Verteilung führt.

Für den Fall $k = 5$ und ausgewählte Stichprobenumfänge wurden die Quantile unter der Nullhypothese für die adaptiven Teststatistiken in [12] ermittelt, indem entsprechende Simulationen durchgeführt wurden, s. Tabelle 2.

$\underline{T}_p^{(\underline{b}(I))}$	$\underline{T}_{np}^{(\underline{b}(I))}$	$\underline{T}_p^{(\underline{b}(II))}$	$\underline{T}_{np}^{(\underline{b}(II))}$	t	
$(1-\alpha) = 0.99$ *Quantile*					
3.150	3.285	3.150	3.285	2.528	$n_i = 5$
2.834	2.834	2.975	2.834	2.403	$n_i = 11$
2.595	2.781	2.775	2.859	2.364	$N_i = 21$
$(1-\alpha) = 0.95$ *Quantile*					
2.374	2.350	2.460	2.450	1.725	$n_i = 5$
2.260	2.250	2.280	2.265	1.676	$n_i = 11$
2.225	2.209	2.313	2.272	1.660	$n_i = 21$
$(1-\alpha) = 0.90$ *Quantile*					
1.980	1.960	2.050	2.080	1.325	$n_i = 5$
1.940	1.921	2.002	2.010	1.299	$n_i = 11$
1.938	1.965	2.000	2.035	1.290	$n_i = 21$

Tabelle 2

Dieser Tabelle ist zu entnehmen, daß bei gleicher Art, die Gewichte b_i zu wählen, zwischen den Quantilen der parametrischen und nichtparametrischen Teststatistik nur ein geringer Unterschied besteht. Die Abweichung von den entsprechenden t-Quantilen ist für kleine

Stichprobenumfänge beträchtlich und verringert sich mit wachsenden Stichprobenumfängen. So ist beispielsweise das 0.99-Quantil einer t-Verteilung mit 20 Freiheitsgraden ($n_i = 5$) 2.528. Die durch Simulation erhaltenen Werte sind aber alle größer oder gleich 3.150.

3 Simulationsuntersuchungen

Simulationsuntersuchungen zu Umbrellatests wurden gelegentlich in der Literatur publiziert, siehe z.B. SCHUMACHER [20], ROTHE [18], HOTHORN [8]. Ziel der hier durchgeführten Simultionsuntersuchungen war die Charakterisierung des Verhaltens (Fehler 1. und 2. Art) oben abgeleiteter adaptiver Umbrellatests im Vergleich zu ausgewählten, alternativen Teststatistiken.

Diese Simulationsstudie ([12]) war Teil einer umfangreichen Studie (vgl. [5, 9]), wo die technischen Details ausführlicher dargestellt sind.

- Verhalten des Fehlers 1. Art α

In Tabelle 3 sind die Schätzwerte $\hat{\alpha}$ der adaptiven Umbrellatests für verschiedene Stichprobenumfänge, Verteilungen {N(0,1); Fleishmann (0,1,2,7)} sowie nominelle α-Niveaus dargestellt.

Geht man von einer $20\% - \alpha$-Robustheit nach HERRENDÖRFER et al. [7] aus, so sind die fett gekennzeichneten Situationen als nicht-robust zu bezeichnen (die Schätzwerte für $\alpha = 0.01$ sind wegen des geringen Simulationsumfangs von 500 Wiederholungen nur als orientierende Angaben zu verstehen).

- Güteverhalten

Für die Betrachtung eines einzelnen Punktes der Gütefunktion mit Güte π $(= 1 - \beta)$ sei hier die Abhängigkeit zu verschiedenen Erwartungswertprofilen für den Fall $n_i =$ const. $=$ 11, $N(0,1)$, $\alpha = 0.01$ in Tabelle 4 dargestellt.

Es ist deutlich, daß für verschiedene Typen von Erwartungswertprofilen jeweils andere Tests ein Güteoptimum (Fettdruck) aufweisen, z. T. mit erheblichem Abstand zum nächstbesten Test, d.h., es existiert bezogen auf Umbrellaprofile kein gleichmäßig bester Test.

Vergleicht man die adaptiven Umbrellatests mit den traditionellen Tests, so zeigt sich die erwartet höhere Güte nur dann, wenn der Umbrellapunkt l_j sich bei nicht zu kleinem $j \in (1, \ldots, k-1)$ befindet.

α	n	Verteilung	$\underline{T}_p^{(\underline{b}(I))}$	$\underline{T}_{np}^{(\underline{b}(I))}$	$\underline{T}_p^{(\underline{b}(II))}$	$\underline{T}_{np}^{(\underline{b}(II))}$
	5		.095	.103	.101	.098
.10	11	N(0,1)	.096	.100	.099	.096
	21		.088	.091	.085	.084
	5		.108	**.126**	.101	.098
.10	11	F(0,1,2,7)	.105	.115	.099	.096
	21		.095	.105	.085	.084
.05	5		.046	.046	.050	.050
	11	N(0,1)	.050	.054	.051	.056
	21		.047	.045	.049	.050
	5		.048	.056	.050	.050
	11	F(0,1,2,7)	.054	**.067**	.051	.056
	21		.050	.050	.049	.050
.01	5		.010	.011	.008	.009
	11	N(0,1)	.013	.012	.015	.016
	21		.019	.014	.012	.010
	5		.010	.014	.008	.009
	11	F(0,1,2,7)	.013	.011	.015	.016
	21		.019	.016	.012	.010

Tabelle 3

$\underline{T}^{mMW}$ – „Kontrolle gegen k Behandlungen"nach HOTHORN [9]
$\underline{T}^{MW}$ – Test nach MACK und WOLFE [16]
$\underline{T}^{Dmax}$ – Maximum-Test (Abschlußtest) nach DUNNETT [3]
$\underline{T}^{FW}$ – Umbrellatest gemäß FLIGNER und WOLFE [4]
$\underline{T}^{VK}$ – Umbrellatest nach VACH, KOZIOL [14], HOTHORN [10]
$\underline{T}^{Steel}$ – Maximum-Test nach STEEL [21]

zu Tabelle 4 ⟶

							parametrisch					nichtparametrisch		
μ_1	μ_2	μ_3	μ_4	μ_5	T^{mMW}	T^{MW}	T^{Dmax}	$T_p^{(\underline{b}(I))}$	$T_p^{(\underline{b}(II))}$	T^{FW}	T^{VK}	T^{Steel}	$T_{np}^{(\underline{b}(I))}$	$T_{np}^{(\underline{b}(II))}$
0	1	2	3	4	.671	.903	.884	**.945**	.922	**.120**	.740	.776	.931	**.933**
0	1	2	4	3	.668	.902	.883	**.939**	.916	.452	.237	.761	**.930**	.927
0	1	3	4	2	.872	**.973**	.885	.896	.885	.738	**.965**	.759	.887	.879
0	1	3	4	1	.934	**.980**	.871	.877	.823	.824	**.964**	.744	.862	.825
0	4	3	2	1	.887	.243	**.906**	.520	.456	.741	.090	**.800**	.548	.539
0	3	4	3	2	**.949**	.752	.923	.798	.705	.842	.680	**.851**	.786	.740
0	3	4	3	1	**.960**	.768	.881	.837	.757	**.881**	.629	.763	**.831**	.783
0	3	4	3	0	**.930**	.722	.780	.825	.753	**.826**	.512	.622	.832	.794
0	2	4	2	0	.856	.585	.828	**.952**	.903	.704	.395	.698	**.934**	.915
0	1	4	1	0	.551	.325	.840	.932	**.946**	.306	.191	.689	.908	**.954**
0	2	4	2	-1	**.781**	.508	.643	-	-	**.631**	.281	.409	-	-

Tabelle 4

Um zu verdeutlichen, welche Gütevorteile die einzelnen Tests gegenüber dem F-Test aufweisen, sind in Tabelle 5 die empirischen Effizienzen nach LEE [15], $\text{EFF}_{emp} = (\hat{\pi}-\hat{\alpha})_T/(\hat{\pi}-\hat{\alpha})_F$, für die verschiedenen Erwartungswertprofile dargestellt. Danach erhärtet sich die Feststellung, daß der Maximum-Test nach DUNNETT [3] (bzw. sein nichtparametrisches Analogon nach STEEL [21]) den *gleichmäßigsten* Gütevorteil gegenüber Ignorierung eines Umbrellaprofils (F-Test) ergibt.

μ_1	μ_2	μ_3	μ_4	μ_5	$\underline{T}^{mMW}$	$\underline{T}^{MW}$	$\underline{T}^{Dmax}$	$\underline{T}_p^{(\underline{b}(I))}$	$\underline{T}_p^{(\underline{b}(II))}$
0	1	2	3	4	(.78)	1.06	1.04	**1.11**	1.09
0	1	2	4	3	(.79)	1.06	1.04	**1.11**	1.08
0	1	3	4	2	1.03	**1.15**	1.04	1.06	1.01
0	1	4	4	1	1.10	**1.15**	1.03	1.03	.97
0	4	3	2	1	1.04	.29	**1.07**	.61	.54
0	3	4	3	2	**1.13**	.91	1.04	.99	.89
0	3	4	3	1	**1.10**	.85	.92	.97	.89
0	2	4	2	0	1.01	.69	.98	**1.12**	1.06
0	1	4	1	0	.65	.38	.99	1.10	**1.11**

Tabelle 5

Interessant ist auch der in Abbildung 1 dargestellte graphische Vergleich der Gütefunktionen von $\underline{T}^{mMW}$, $\underline{T}^{MW}$, $\underline{T}^{Dmax}$, $\underline{T}_p^{(\underline{b}(I))}$und $\underline{T}_p^{(\underline{b}(II))}$für n = const. = 5, $\alpha = 0.05$, $N(0,1)$, EW-Profil 0/1/3/4/2. Deutlich ist das ungünstige Güteverhalten bei kleinen Nichtzentralitätsparametern (der des F-Tests !) der adaptiven Umbrellatests zu erkennen.

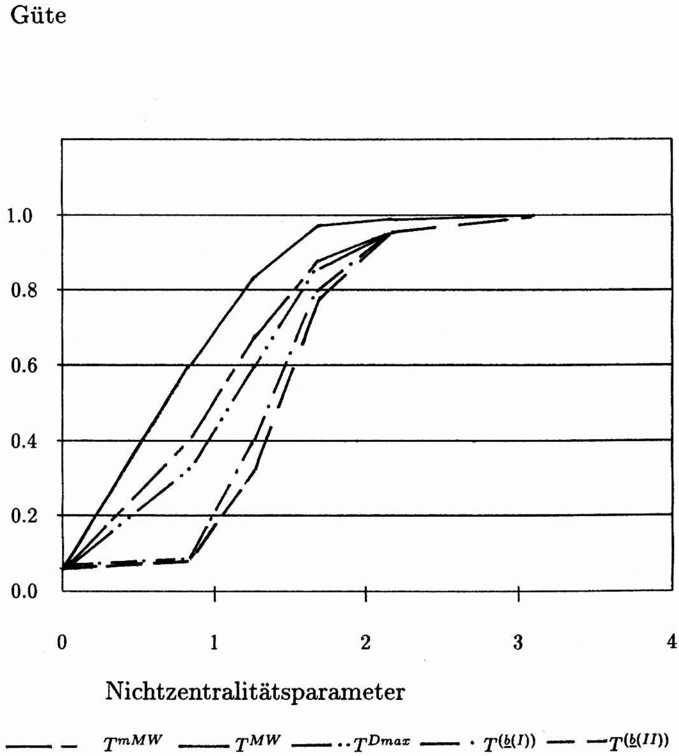

Abbildung 1

Der Einfluß verschiedener Stichprobenumfänge und Varianzen ist in Tabelle 6 ($\alpha = 0.05$, N(0,1), EW-Profil 0/1/3/4/2) veranschaulicht. Dabei ist deutlich der günstige Einfluß auf die Güte bei optimalem Stichprobenprofil, d.h. bei $n_k > n_i$, ersichtlich. Ausgeprägte Varianzinhomogenität bewirkt erwartungsgemäß einen gewissen Güteabfall, deutlicher ausgeprägt bei den parametrischen Versionen.

n_1	n_2	n_3	n_4	n_5	σ_i^2	$\underline{T}_p^{(b(I))}$	$\underline{T}_p^{(b(II))}$	$\underline{T}_{np}^{(b(I))}$	$\underline{T}_{np}^{(b(II))}$
11	11	11	11	11	1:1:1:1:1	.963	.964	.952	.949
15	13	11	9	7	1:1:1:1:1	.988	.985	.983	.984
15	13	11	9	7	1:2:4:7:10	.925	.914	.985	.983
19	9	9	9	9	1:1:1:1:1	.996	.994	.994	.995

Tabelle 6

Güteschätzungen für diverse andere Bedingungen ($\alpha = 0.01$ und 0.1; $n_i = 5, 11, 21$; F $(0, 1, 2, 7)$, Varianzinhomogenität, Unbalanciertheit, Erwartungswertprofile) sind in [12] enthalten.

Literatur

[1] **Abelson, R.P.**, and **Tukey, J.W.**: *Efficient utilization of non-numerical information in quantitative analysis.* Ann. Math. Statist. **34**, 1347–1369 (1963)

[2] **Barlow, R.E.**, et al.: *Statistical Inference under Order Restrictions.* New York 1972

[3] **Dunnett, C.W.** : *A multiple comparison procedure for comparing several treatments with a control.* J. Amer. Statist. Assoc. **50**, 1096–1121 (1955)

[4] **Fligner, M.A.**, and **Wolfe, D.A.**: *Distribution-free tests for comparing several treatments with a control.* Statist. Neerlandica **38**, 119–127 (1982)

[5] **Gretzebach, L.** : *Simulationsuntersuchungen an Tests für Umbrella-Alternativen.* Diplomarbeit, Universität Rostock 1987

[6] **Gretzebach, L.**, **Hothorn, L.**, and **Liese, F.**: *A simulation study for comparing tests under umbrella alternatives.* Probl. angew. Statist. **24**, 55–68 (1988)

[7] **Herrendörfer, G.**, et al.: *Robustheit I.* Probl. angew. Statist. **4**, 1–204 (1980)

[8] **Hothorn, L.** : *Simulationsuntersuchungen zur Robustheit von k-Stichprobenlokationstests mit geordneter Alternativhypothese.* Probl. angew. Statist. **15**, 116–134 (1985)

[9] **Hothorn, L.** : *k-Stichprobentests und Vergleichsprozeduren in Dosis-Wirkungs-Abhängigkeiten toxikologischer Untersuchungen.* Dissertation B, Martin-Luther-Universität Halle 1987

[10] **Hothorn, L.** : *Tests bei Umbrella-Alternativen, dargestellt am Beispiel des AMES-Assay.* Gesundheit u. Umwelt **3**, 33-50 (1987)

[11] **Jonckheere, A.R.** : *A ditribution-free k-sample test against ordered alternatives.* Biometrika **41**, 133–145 (1954)

[12] **Knospe, A.** : *Simulationsuntersuchungen an Tests für Umbrella-Alternativen.* Diplomarbeit, Universität Rostock 1988

[13] **Kruskal, W.H.**, and **Wallis, W.A.**: *Use of ranks in one-criterion variance analysis.* J. Amer. Statist. Assoc. **47**, 583–621 (1952)

[14] **Koziol, J.** : *A simple nonparametric test of trend.* Bometrical J. **20**, 503–509 (1978)

[15] **Lee, Y.J.** : *A distribution-free test for stochastic ordering.* J. Amer. Statist. Assoc. **71**, 722–727 (1976)

[16] **Mack, G.A.**, and **Wolfe, A.D.**: *k-sample rank test for umbrella alternatives.* J. Amer. Statist. Assoc. **76**, 175 f. (1981)

[17] **Pitman, F.J.G.** : *Lecture notes on nonparametric statistical inference.* Columbia University, New York 1948 (unpubl.)

[18] **Rothe, G.** : *Mehrstichproben-Rangtests bei einseitigen Alternativen zur statistischen Auswertung von Mutagenitätsexperimenten.* Forschungsbericht, Universität Dortmund 1985, Nr. 3

[19] **Rothe, G.** : *Linear trend test versus global test: a comparison.* Statist. Neerlandica **40**, 1-11 (1986)

[20] **Schuhmacher, M.** : *Statistische Analyse des AMES-Test.* Forschungsbericht, Universität Dortmund 1985, Nr. 18

[21] **Steel, R.G.D.** : *A multiple comparison rank sum test treatment versus control.* Biometrics **15**, 560–572 (1959)

[22] **Vach, W. :** zitiert in [18], S. 9

[23] **Van der Laan, P.**, and **Verdooren, L.R.**: *Classical analysis of variance methods and non-parametric counterparts.* Biometrical J. **29**, 635–666 (1987)

eingegangen: 15. Oktober 1990

Verfasser

Dr. habil. L. Hothorn
Arzneimittelwerk Dresden
Zentralstelle für Produktentoxikologie Graupa
Bonnewitzer Str. 34
O–8304 Graupa
Germany

Prof. Dr. habil. F. Liese
Universität Rostock
Fachbereich Mathematik
Universitätsplatz 1
O–2500 Rostock
Germany

Rostock. Math. Kolloq. 45, 75–76 (1991)

Subject Classification (AMS)
03C99; 01A80

INGO KÖLBL

Theoretisch-konzeptionelle Untersuchungen zu Fragen einer Koordination des Mathematikunterrichts mit Inhalten anderer Unterrichtsfächer, dargestellt am Beispiel der Methode der mathematischen Modellierung von außermathematischen Sachverhalten

Autorreferat der Dissertation B

Die Notwendigkeit einer Koordination des Mathematikunterrichts mit anderen Unterrichtsfächern ergibt sich aus den Entwicklungstendenzen der Mathematisierung in anderen Wissenschaften, aus Zielstellungen und Anforderungen der Gesellschaft an die Schule bei der Vermittlung einer Allgemeinbildung und aus der Gestaltung des Unterrichtsprozesses. Die stofflichen Beziehungen zwischen verschiedenen Unterrichtsfächern sind Widerspiegelungen der Verbindungen der diesen Fächern zugrunde liegenden Fachwissenschaften und sind der Strukturierung der Inhalte in den einzelnen Fachlehrgängen untergeordnet. Im Mathematikunterricht realisieren sich diese Verbindungen i. allg. über Anwendungen der Mathematik auf außermathematische Sachverhalte.

Grundlage für eine methodische Gestaltung von Anwendungen der Mathematik auf außermathematische Sachverhalte mit Hilfe der Methode der mathematischen Modellierung bilden Abbildungen zwischen Strukturen des Anwendungsbereichs der Mathematik und mathematischen Strukturen. Diese Abbildungen drücken Wechselbeziehungen zwischen verschiedenen Erkenntnisebenen aus. Die Schüler müssen zwei Problemlösungsprozesse bewältigen: die Bearbeitung eines Problems in einer Realebene und die Lösung eines daraus erstellten mathematischen Problems in der Ebene der Mathematik.

Das methodische Vorgehen innerhalb eines solchen mathematischen Modellierungsprozesses außermathematischer Sachverhalte erfolgt für die Schüler in zwei Stufen. In einer ersten wird eine für die weitere Bearbeitung des Problems günstige Struktur des Sachverhalts herausgestellt, die, auf eine mathematische Struktur abgebildet, in einer zweiten Stufe als mathematisches Modell weiter bearbeitet wird. Das hieraus folgende mathematische Ergebnis wird auf die Struktur des realen Sachverhalts wieder abgebildet und daraus das Ergebnis des Ausgangsproblems formuliert.

Im mathematisch-naturwissenschaftlichen Unterricht ist durch ein koordiniertes Vorgehen zu erreichen, daß die Schüler die Methode der mathematischen Modellierung und den Begriff „Mathematisches Modell“ inhaltlich erfassen. Dabei erkennen sie, daß ein mathematisches Modell eine mathematische Struktur ist, die aus mathematischen Objekten (Zahlen, Größen, Variablen u. a.) und den zwischen ihnen existierenden Relationen (Operationen, Ordnungsrelationen, Abhängigkeiten u. a.) besteht, die eine Struktur des Anwendungsbereiches widerspiegelt und so zu vertreten vermag, daß durch die Arbeit mit dem mathematischen Modell neue Informationen über den Anwendungsbereich gewonnen werden.

Anhand von Beispielen werden in der Arbeit Zielstellungen zur Entwicklung von Schülertätigkeiten bei mathematischen Modellierungsprozessen betrachtet, und es wird auf Differenzierungsmaßnahmen hingewiesen.

eingereicht: 10. Mai 1989 **verteidigt:** 5. Januar 1990

Gutachter: Prof. Dr. W. Walsch (Halle)
Prof. Dr. K.-H. Weber (Berlin)
Prof. Dr. W. Karsten (Rostock)
Prof. Dr. G. Hellfeldt (Rostock)

Verfasser:

Dr. sc. paed. I. Kölbl
Universität Rostock
Fachbereich Mathematik
Universitätsplatz 1
O-2500 Rostock
Germany

Rostock. Math. Kolloq. 45, 77–78 (1991) Subject Classification (AMS) 76E25; 35P30

PETER MÖLLER

Stabile Oberflächenkonfigurationen magnetischer Flüssigkeiten

Autorreferat der Dissertation A

Unter dem Einfluß eines homogenen, stationären, vertikalen Magnetfeldes bildet eine magnetische Flüssigkeit bei Überschreiten einer gewissen kritischen Feldstärke stabile hexagonale Oberflächenkonfigurationen aus. Die Arbeit befaßt sich mit der mathematischen Beschreibung dieses Phänomens.
Es sei $\Gamma : z = \gamma(x, y)$ die Trennfläche zwischen einer inkompressiblen magnetischen Flüssigkeit und einem Vakuum. Flüssigkeits- und Vakuumgebiet setzen wir als unbeschränkt voraus. Berücksichtigt werden die Einflüsse des Schwerefeldes und der Oberflächenspannung. Die Funktion $\gamma(x, y)$ sei bezüglich eines Gitters G periodisch. Unsere Untersuchungen beruhen auf dem Prinzip minimaler potentieller Energie des Systems. Mit der genannten Problemstellung haben sich bereits mehrere Arbeiten von Prof. K. Beyer befaßt. Dort wurde der Fall konstanter Gitterlänge $l = l_{cr}$ behandelt. Daran anknüpfend untersuchen wir, inwieweit die Lösung der gestellten Aufgabe beeinflußt wird, wenn man neben $\gamma(x, y)$ auch l als variabel ansieht.
Eine Formel für die potentielle Energie des Magnetfeldes wird durch Behandlung der Maxwellgleichungen im Sinne der schwachen Lösungstheorie bereitgestellt. Hierbei finden Hilfsmittel aus der Theorie der BEPPO-LEVI-Funktionen Anwendung. Zur Beschreibung des Energiefunktionals ist ein freies Randwertproblem zu betrachten, welches sich in ein Variationsproblem überführen läßt. Die bislang variablen Integrationsgebiete werden einer Festrandtransformation unterworfen. Damit ist ein Funktional $F(f, \lambda; \varepsilon, \kappa)$ zu untersuchen, welches analytisch von der Transformierten f der Funktion $\gamma(x, y)$, der Variablen $\lambda = l_{cr}^{-1} l$ sowie den Feldstärke- bzw. Suszeptibilitätsparametern ε und κ abhängt. Glieder niederer Ordnung sind konkret berechenbar, diejenigen höherer Ordnung dürfen vernachlässigt werden. Die Transformierte f variiert über periodischen Sobolevräumen H^s der Differenzierbarkeitsklasse $s \geq 3$. Dabei interessieren nur holoedrieinvariante Lösungsflächen aus den Teilräumen H^s_g.

Der Berücksichtigung der Translationsinvarianz der potentiellen Energie ist ein gesondertes Kapitel gewidmet.

Anschließend werden die kritischen Stellen des Funktionals $F(f,\lambda;\varepsilon,\kappa)$ ermittelt. Der durch $F_1(f,\lambda;\varepsilon,\kappa) = \frac{\partial}{\partial f}F(f,\lambda;\varepsilon,\kappa)$ definierte Operator bildet $H^s_g \times I\!R^+$ stetig in H^{s-2} ab. Die horizontale Trennfläche $f = 0$ liefert im unterkritischen Feldstärkebereich $\varepsilon < 0$ den einzigen Gleichgewichtszustand.

Der Operator $\frac{\partial}{\partial f}F_1(f,\lambda;\varepsilon,\kappa)$ besitzt für den kritischen Parameterwert $\varepsilon = 0$ einen eindimensionalen Nullraum N zum Eigenwert $\lambda = 1$ über H^s_g. Die kritischen Stellen von F werden in einer Umgebung von $(f,\lambda;\varepsilon,\kappa) = (0,0;1,\kappa')$ unter Anwendung der Methode von LJAPUNOV-SCHMIDT ermittelt. Es ergeben sich zwei Lösungszweige $I^\pm$ für „kleine" Parameter κ, also für $\kappa' = 0$, und einer für $\kappa' \neq 0$.

Stabilität bzw. Instabilität der gefundenen Lösungszweige sind bereits durch das Definitheitsverhalten der Form $\frac{\partial^2}{\partial f^2}F(f,\lambda;\varepsilon,\kappa)(.,.)$ über N bestimmt.

Wir gelangen schließlich zu Ergebnissen, die denjenigen aus dem Fall mit konstanter Gitterlänge analog sind: Stabiles Verhalten weist nur der Lösungszweig I^+ in einem gewissen Parametergebiet auf, die anderen sind instabil.

eingereicht: Juni 1989 **verteidigt:** 15. Dezember 1989

Gutachter: Prof. Dr. K. Beyer (Leipzig)
Prof. Dr. G. Wildenhain (Rostock)
Doz. Dr. E. Miersemann (Leipzig)

Verfasser:

Dr. Peter Möller

Universität Rostock

Fachbereich Maschinenbau und Schiffstechnik
Institut für Technische Mechanik

A.-Einstein-Straße 2

O-2500 Rostock
Germany

Rostock. Math. Kolloq. 45, 79–80 (1991)

Subject Classification (AMS)
20C20; 20C05

GUNTER TIEDT

Gruppen mit extremalen Blöcken

Autorreferat der Dissertation A

Die Arbeit behandelt Gruppen G und Gruppenalgebren FG endlicher Gruppen über Körpern der Charakteristik $p > 0$. Insbesondere werden Beziehungen zwischen der Struktur der Gruppenalgebra FG und der Einbettung der p-Sylowgruppe P in die Gruppe G untersucht. Erste Resultate in dieser Richtung stammen von CHILLAG/HERZOG (1979), HARADA (1968), HERZOG (1976) und NINOMIYA (1982).

Der Körper F sei für jede Gruppe G ein fest gewählter Zerfällungskörper. Unter einem Block b von G verstehen wir ein primitives orthogonales Idempotent von FG. Diesen Blöcken werden in üblicher Art und Weise Defektgruppen $\delta(b)$ zugeordnet. Die natürliche Zahl a heißt Defekt von b, wenn $|\delta(b)| = p^a$ gilt. In der Arbeit werden überwiegend Gruppen betrachtet, in denen jeder Block vollen Defekt hat (pFD-Gruppen) bzw. in denen jeder Block vollen oder trivialen Defekt hat ($pFZD$-Gruppen). Es wird untersucht, inwieweit diese Gruppen p-Normalteiler bzw. die pTI-Eigenschaft haben. Unter anderem gilt der folgende

Satz *1. Genau dann ist G eine pFD-Gruppe mit pCI-Eigenschaft, wenn G p-abgeschlossen ist.*

2. Die Gruppe G ist genau dann eine pTI-Gruppe, wenn folgende Bedingungen erfüllt sind:

a) G ist eine $pFZD$-Gruppe, und

b) G besitzt die pCI-Eigenschaft.

Diese Ergebnisse waren bisher nur für $p = 2$ und in einer spezielleren Form von M. HERZOG bekannt.

In §3 werden auflösbare Gruppen behandelt und die Einbettung der p-Sylowgruppe untersucht. Aus der pFD-Eigenschaft folgt dann, daß G p-abgeschlossen ist, oder für die minimale Erzeugendenzahl $d(O_p(G))$ von $O_p(G)$ gilt

$$d(O_p(G)) = \begin{cases} p-1, & \text{wenn } p \text{ eine Fermat-Primzahl ist,} \\ p & \text{sonst.} \end{cases}$$

Um zu zeigen, daß sich das Ergebnis nicht verbessern läßt, werden Beispiele für Gruppen der p-Länge 2 mit $|O_p(G)| = p^{p-1}$ bzw. $|O_p(G)| = p^p$ angegeben. Für $pFZD$-Gruppen erreicht man ein etwas besseres Resultat. Es zeigt sich, daß $pFZD$-Gruppen, welche keine pFD-Gruppen sind, immer die pTI-Eigenschaft haben. Außerdem ist die p-Sylowgruppe in diesem Fall zyklisch oder verallgemeinerte Quaternionengruppe.

Ferner werden auflösbare, nicht p-abgeschlossene Gruppen, für welche $|O_p(G)| = p^{p-1}$ bzw. $|O_p(G)| = p^p$ ist, untersucht. Wir bezeichnen sie als p-minimale Gruppen. Es finden sich Angaben über die Klasse der p-Sylowgruppe, die p-Länge und Stufe solcher Gruppen.

Der letzte Teil der Arbeit beschäftigt sich mit der Struktur der Gruppenalgebra FG. Es zeigt sich, daß bei pFD-Gruppen stets alle Blöcke ihren Träger in $O_{p'}(G)$ haben. Wie einfache Beispiele belegen, ist diese Aussage nicht auf Blöcke vom Defekt Null zu verallgemeinern. Hier beweisen wir für eine Reihe von nicht p-nilpotenten $pFZD$-Gruppen, daß sie Blöcke außerhalb von $FO_{p'}(G)$ besitzen.

eingereicht: 22. Dezember 1988 **verteidigt:** 28. Juni 1989

Gutachter: Prof. Dr. G. Pazderski (Rostock)
Prof. Dr. K. Rosenbaum (Erfurt)
Doz. Dr. L. Prohaska (Rostock)

Verfasser:

Dr. G. Tiedt
Universität Rostock
Fachbereich Mathematik
Universitätsplatz 1
O-2500 Rostock
Germany

Hinweise für Verfasser

Um die redaktionelle Bearbeitung und die Herstellung der Druckvorlage zu erleichtern, wären wir den Verfassern dankbar, sich betreffs der Form der Manuskripte an den in **Rostock. Math. Kolloq.** (ab Heft **43**) veröffentlichten Beiträgen zu orientieren.
Insbesondere beachte man:

1. Manuskripte sollten grundsätzlich **maschinengeschrieben** (Schreibmaschine, Drucker) **in deutscher oder englischer Sprache** abgefaßt sein.
2. Zur inhaltlichen Einordnung der Arbeit sind **1–2 Klassifizierungsnummern** (entsprechend der "1980 Mathematics Subject Classification" der Mathematical Reviews) anzugeben.
3. **Textbreite/Texthöhe** des Manuskripts sollten sich an den Maßen **160mm/230mm** orientieren.
4. Der Manuskripttext ist **eineinhalbzeilig, linksbündig**, wenn möglich, links-und rechtsbündig zu schreiben. Beim Auftreten von Formeln im laufenden Text ist der Zeilenabstand entsprechend zu vergrößern.
5. Der Platz für **Abbildungen** sollte beim Schreiben ausgespart werden; die Abbildungen selbst sind in der dem ausgesparten Platz entsprechenden Größe gesondert beizufügen.
6. **Literaturzitate** sind im Text durch laufende Nummern (vgl. [3], [4] ; [7, 8, 10]) zu kennzeichnen und am Schluß der Arbeit unter der Zwischenüberschrift **Literatur** bzw. **References** zusammenzustellen. Hierbei ist die durch die nachfolgenden Beispiele veranschaulichte Form einzuhalten (die Zeitschriftenabkürzungen erfolgen nach Mathematical Reviews).

   [3] **Zariski, O.**, and **Samuel, P.**: *Commutative Algebra.* Princeton 1958

   [4] **Steinitz, E.**: *Algebraische Theorie der Körper.* J. Reine Angew. Math. **137**, 167–309 (1920)

   [8] **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 erfolgen in Originalsprache ; bei kyrillischen Buchstaben sollte die bibliothekarische Transkription bzw. eine Übersetzung lt. Mathematical Reviews verwendet werden.
7. Die aktuelle, vollständige Adresse des Verfassers sollte enthalten: Titel / Vornamen Name / Institution / Strukturеinheit / Straße Hausnummer / Postleitzahl Ort / Land.

Weiterhin besteht die Möglichkeit, mit dem **Satzsystem LaTeX** erstellte Manuskripte auf unter **MS–DOS** formatierten Disketten (**3.5", 0.72MB, 1.44MB; 5.25", 0.36MB, 0.72MB, 1.2MB**) einzureichen.

Schriftenreihen der Universität Rostock

– Archiv der Freunde der Naturgeschichte in Mecklenburg	ISSN 0518-3189
– Rostocker Agrarwissenschaftliche Beiträge	ISSN 0138-3299
– Rostocker Betriebswirtschaftliche Manuskripte	ISSN 0232-3066
– Rostocker Mathematisches Kolloquium	ISSN 0138-3248
– Rostocker Philosophische Manuskripte	ISSN 0557-3599
– Rostocker Physikalische Manuskripte	ISSN 0138-3140
– Rostocker Wissenschaftshistorische Manuskripte	ISSN 0138-3191
— Lateinamerika/Semesterberichte der Sektion Lateinamerikawissenschaften	ISSN 0458-7944
– Erziehungswissenschaftliche Beiträge	ISSN 0138-2373
– Beiträge zur Geschichte der Universität Rostock	ISSN 0232-539X
– Rostocker Beiträge zur Hoch- und Fachschulpädagogik	ISSN 0233-0539
– Rostocker Informatik-Berichte	ISSN 0233-0784
– Studien zur Geschichte der deutsch-polnischen Beziehungen	ISSN 0233-0687
– Rostocker Forschungen zur Sprach- und Literaturwissenschaft	ISSN 0233-0644
– Migrationsforschung	ISSN 0863-1735
– Manuskripte zur Rostocker Universitätsgeschichte	ISSN 0863-1727
– Rostocker Universitätsreden	
– Agrargeschichte	ISSN 0863-2170

Bezugsmöglichkeiten

- Bestellungen sind zu richten an die Universität Rostock, Abt. Wissenschaftspublizistik, Vogelsang 13/14, O-2500 Rostock.

Ferner sind die Hefte im Rahmen des Schriftentausches über die Universität Rostock, Universitätsbibliothek, Tauschstelle, Universitätsplatz 5, O-2500 Rostock, zu beziehen.