Drews

ISSN 0138 - 3248

Rostocker Mathematisches Kolloquium

Heft 28

WILHELM-PIECK-UNIVERSITÄT
ROSTOCK

ROSTOCKER MATHEMATISCHES KOLLOQUIUM

Heft 28

Wilhelm-Pieck-Universität Rostock
Sektion Mathematik

1985

Herausgeber: Der Rektor der Wilhelm-Pieck-Universität Rostock

Wissenschaftliche Leitung: Prof. Dr. Wolfgang Engel (Sektionsdirektor)
Prof. Dr. Gerhard Maeß

Redaktionelle Bearbeitung: Dr. Klaus-Dieter Drews

Herstellung der Druckvorlage: Dorothea Meyer

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

Redaktionsschluß: 30. 09. 1985

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

Zitat-Kurztitel: Rostock. Math. Kolloq.

Abt. Wissenschaftspublizistik,
Wilhelm-Pieck-Universität Rostock,
DDR-2500 Rostock, Vogelsang 13/14, Telefon 369 577
Genehmigungs-Nr.: C 83/1985
Druck: ODR W II-15-14·0,50

https://doi.org/10.18453/rosdok_id00003841

Inhalt

Rostock. Math. Kolloq. 28, 5 - 17 (1985) 03G20
06A15

Walter Harnau

Ein verallgemeinerter Relationenbegriff für die Algebra der mehrwertigen Logik, Teil I (Grundlagen)

Im Jahre 1969 zeigten Bodnartschuk, Kalushnin, Kotow und Romow in /1/, daß eine Galoisbeziehung zwischen den Funktionen- und Relationenalgebren über einer endlichen Menge besteht. Das Ziel dieses Zyklus von Artikeln besteht darin, die Resultate von /1/ dahingehend zu verallgemeinern, daß in geeigneter Form Relationenpaare und Operationen darüber erklärt werden, um damit auch eine Galoisbeziehung zwischen den iterativen Algebren und den Relationenpaaralgebren über einer endlichen Menge zu erhalten.
Andere Autoren (z. B. Hikita und Nozaki (/4/, /9/), Gössel und Pöschel (/2/)) arbeiteten bereits mit Relationenpaaren, mitunter auch in einer etwas allgemeineren Art als der hier betrachteten, ohne jedoch eine Theorie dafür zu entwickeln, was hier vorgenommen werden soll. Die Resultate dieses Zyklus sind Bestandteil der B-Dissertation /3/ des Autors.
Der Teil I stellt zum Großteil bereits bekannte Resultate zusammen und soll das Lesen dieser Artikelserie auch ohne zusätzliche Literatur ermöglichen. Dabei halten wir uns in den Bezeichnungen weitestgehend an die der deutschsprachigen Monographie /10/ von Pöschel und Kalushnin, in der auch einige der in diesem Teil zusammengestellten Resultate ausführlicher dargestellt sind.

1. Grundbegriffe

Es sei $E_k := \{0,1,\ldots, k-1\}$, wobei k eine natürliche Zahl mit $k \geq 2$ sei. Für jede positive ganze Zahl n bezeichne

$$P_k^{(n)} := \{f : E_k^n \to E_k\}$$

die Menge aller n-stelligen Funktionen über E_k.

$$P_k := \bigcup_{n>0} P_k^{(n)}$$

ist die Menge der Funktionen der k-wertigen Logik. Nullstellige Funktionen (d. h. Konstanten) betrachten wir als konstante einstellige Funktionen. Für $A \subseteq P_k$ sei $A^{(n)}$ die Menge aller n-stelligen Funktionen aus A. Der Funktionswert von $f \in P_k^{(n)}$ auf dem n-Tupel $\underline{x} = (x_1, x_2, \ldots, x_n) \in E_k^n$ wird mit $f(x_1, x_2, \ldots, x_n)$ oder $f(\underline{x})$ bezeichnet.

Eine besondere Rolle werden die Projektionen $e_i^n : E_k^n \to E_k$ spielen, die jedem n-Tupel $(x_1, x_2, \ldots, x_n)$ aus E_k^n die i-te Komponente x_i zuordnen, wobei n eine beliebige positive ganze Zahl und $i \in \{1,2,\ldots, n\}$ ist. Mit J_k bezeichnen wir die Menge der Projektionen von P_k.

Eine Funktion $f \in P_k^{(n)}$ heißt wesentlich an der i-ten Stelle (in der i-ten Variablen) $(1 \leq i \leq n)$, wenn $a_1, a_2, \ldots, a_n, b \in E_k$ derart existieren, daß

$$f(a_1, a_2, \ldots, a_n) \neq f(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n)$$

gilt. Anderenfalls heißt f fiktiv an der i-ten Stelle.

Auf P_k werden mit $\zeta, \tau, \Delta, \nabla$, o die folgenden Operationen bezeichnet (/8/):

Es sei $f \in P_k^{(n)}$ und $g \in P_k^{(m)}$; dann sind $\zeta f \in P_k^{(n)}$, $\tau f \in P_k^{(n)}$ $\Delta f \in P_k^{(\max(n-1,1))}$, $\nabla f \in P_k^{(n+1)}$ und $g \circ f \in P_k^{(n+m-1)}$ mit $x_1, x_2, \ldots, x_{n+m-1}, x_{n+1} \in E_k$ definiert durch

$$(\zeta f)(x_1, \ldots, x_n) := f(x_2, \ldots, x_n, x_1),$$

$$(\tau f)(x_1, \ldots, x_n) := f(x_2, x_1, x_3, \ldots, x_n), \qquad \text{für } n \geq 2$$

$$(\Delta f)(x_1, \ldots, x_{n-1}) := f(x_1, x_1, x_2, \ldots, x_{n-1}),$$

und $\zeta f := \tau f := \Delta f := f$ für $n = 1$,

$$(\nabla f)(x_1, \ldots, x_{n+1}) := f(x_2, x_3, \ldots, x_{n+1}),$$

$$(g \circ f)(x_1, \ldots, x_{n+m-1}) := f(g(x_1, \ldots, x_m), x_{m+1}, \ldots, x_{n+m-1}).$$

Im folgenden betrachten wir die beiden universalen Algebren $\underline{P}_k = [P_k; \zeta, \tau, \Delta, \nabla, o]$ und $\hat{\underline{P}}_k = [P_k; e_1^2, \zeta, \tau, \Delta, o]$ vom Typ (1,1,1,1,2) bzw. (0,1,1,1,2). (Über Grundbegriffe der universellen Algebra, die hier undefiniert benutzt werden, informiere man sich z. B. in /7/.) Erstere bezeichnen wir als volle iterative Algebra über E_k und letztere als volle Funktionenalgebra über E_k. Eine Funktion $f \in P_k$, die sich aus gegebenen Funktionen $g_1, \ldots, g_t$ höchstens durch Anwendung der Operationen $\zeta, \tau, \Delta, \nabla$, o bzw. $e_1^2, \zeta, \tau, \Delta$, o erzeugen läßt, heißt Superposition bzw. $*$-Superposition über $\{g_1, \ldots, g_t\}$. Jede Teilalgebra F von $\underline{P}_k$ bzw. $\hat{\underline{P}}_k$ heißt iterative Algebra bzw. Funktionenalgebra über E_k. Da keine Verwechselungen zu befürchten sind, unterscheiden wir nicht zwischen den iterativen Algebren bzw. Funktionenalgebren und ihren Trägermengen.

Die von einer Menge $F \subseteq P_k$ erzeugte iterative Algebra bzw. Funktionenalgebra ist die kleinste F enthaltende iterative Algebra bzw. Funktionenalgebra, die kurz mit $[F]_k$ bzw. $\langle F \rangle_k$ bezeichnet wird. Wenn keine Mißverständnisse zu befürchten sind, kann hier auch noch der Index k weggelassen werden.

Mit $\sum_k$ bzw. $\sum_k^*$ bezeichnen wir die Menge der Teilalgebren von $\underline{P}_k$ bzw. $\hat{\underline{P}}_k$. Dann sind $(\sum_k, \subseteq)$ und $(\sum_k^*, \subseteq)$ bez. der Inklusion vollständige Verbände. (Über Grundbegriffe der Verbandstheorie lese man z. B. /13/.)

Offensichtlich sind die Teilalgebren von $\underline{P}_k$ genau die (superpositions-)abgeschlossenen Klassen von P_k im Sinne von Jablonski (/5/).

Wir wollen jetzt den Zusammenhang zwischen den iterativen Algebren und Funktionenalgebren über E_k herleiten. Dabei sei mit $\emptyset$ im weiteren stets die leere Menge bezeichnet.

Offenbar gilt das

Lemma 1: $J_k = [J_k] = \langle J_k \rangle$, $[\emptyset] = \emptyset$.

Lemma 2: $\langle \emptyset \rangle_k = J_k$.

Beweis: Da e_1^2 eine nullstellige Operation ist, kann e_1^2 aus $\emptyset$ erzeugt werden. Es ist $e_1^1 = \Delta e_1^2$, $e_2^2 = \tau e_1^2$, $e_i^n = e_1^{n-1} \circ e_2^2$ für $n > 2$ und $1 \leq i \leq n-1$, und für $n > 2$ ist $e_n^n = e_1^{n-1} \circ e_2^2$. Damit ist $\langle \emptyset \rangle_k \supseteq J_k$ nachgewiesen. Aus dem Lemma 1 folgt dann die Behauptung.

Lemma 3: $[e_1^1]_k = J_k$.

Beweis: Es ist $\nabla e_1^1 = e_2^2$ und $\tau e_2^2 = e_1^2$. Weiter verläuft der Beweis wie in Lemma 2.

Lemma 4: Wenn F eine iterative Algebra über E_k ist, so ist $F \cap J_k \in \{\emptyset, J_k\}$.

Beweis: Wenn $F \cap J_k \neq \emptyset$ ist, so gibt es ein e_i^n in $F \cap J_k$. Da F und J_k Elemente von $\sum_k$ sind und $(\sum_k, \subseteq)$ ein Verband ist, ist auch $F \cap J_k \in \sum_k$. Somit ist weiter

$$\underbrace{\Delta\Delta\ldots\Delta}_{(n-1)\text{mal}} e_i^n = e_1^1 \in F \cap J_k.$$

Aus Lemma 3 folgt $J_k \subseteq F \cap J_k$, womit das Lemma bewiesen ist.

Satz 1: Für alle natürlichen Zahlen k mit $k \geq 2$ gilt $\sum_k^* \subset \sum_k$, und aus $F \in \sum_k \setminus \sum_k^*$ folgt $F \cap J_k = \emptyset$.

Beweis: Wegen $\emptyset \in \sum_k \setminus \sum_k^*$ (Lemmata 1 und 2) ist $\sum_k \neq \sum_k^*$. Es sei $F \in \sum_k^*$. Dann gilt $e_2^2 \in J_k \cap F$ (Lemma 2) und $\nabla f = e_2^2 \circ f$. Demnach ist F gegenüber ζ, τ, Δ, ∇, $\circ$ abgeschlossen, woraus $F \in \sum_k$ und damit $\sum_k^* \subset \sum_k$ folgt. Den zweiten Teil des Satzes beweisen wir indirekt und nehmen dazu an, daß es ein $F \in \sum_k \setminus \sum_k^*$ mit $F \cap J_k \neq \emptyset$ gibt. Nach Lemma 4 ist dann $J_k \subseteq F$. Folglich ist F gegenüber ζ, τ, Δ, $\circ$ und der nullstelli-

gen Operation e_1^2 abgeschlossen und damit eine Funktionenalgebra über E_k, was einen Widerspruch zur Annahme, daß F keine Funktionenalgebra ist, bedeutet und den Beweis des Satzes abschließt.

Folgerung 1: $\sum_k^*$ besteht aus genau den iterativen Algebren F über E_k, für die $J_k \subseteq F$ gilt.

2. Die klassische Relationenalgebra

Die Idee, gewisse superpositionsabgeschlossene Klassen (iterative Algebren) über E_k durch Relationen zu charakterisieren, geht auf Kusnezow (/6/) zurück und wurde vor allem von Jablonski (/5/) und letztlich von Rosenberg (/11/ und /12/) im Jahre 1965 erfolgreich zur Lösung des "Vollständigkeitsproblems für P_k" ausgenutzt.

Bodnartschuk, Kalushnin, Kotow und Romow definierten 1969 in /1/ eine volle Relationenalgebra über E_k derart, daß deren Teilalgebren (Relationenalgebren) in Galoisbeziehung zu den Funktionenalgebren über E_k stehen. Dieses Resultat soll hier kurz dargestellt und später verallgemeinert werden.
Eine m-stellige Relation ϱ über E_k ist eine Teilmenge von E_k^m, d. h. eine Menge von m-Tupeln über E_k. Diese m-Tupel denken wir uns als Spalten $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$ und geben mitunter die Relation ϱ in Form einer Matrix an,

$$\varrho = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1r} \\ x_{21} & x_{22} & \cdots & x_{2r} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mr} \end{pmatrix},$$

deren Spalten die Elemente von ϱ sind. Deshalb werden wir auch von den Zeilen und Spalten einer Relation sprechen. Die Zeilen

(auch Komponenten oder Koordinaten von ϱ genannt) sind die in irgendeiner Reihenfolge geordneten i-ten Komponenten der Spalten (i = 1,2,...,m). Wir schreiben mitunter auch für die Spalte

$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$, um Platz einzusparen, $[x_1,x_2,\ldots,x_m]$. Nach Definition ist auch die leere Menge ($\emptyset$) eine Relation.

Für jede positive ganze Zahl m bezeichne $R_k^{(m)}$ die Menge aller m-stelligen Relationen über E_k und $R_k := \bigcup_{m>0} R_k^{(m)}$ die Menge aller endlichstelligen Relationen über E_k.

Es seien n und m positive ganze Zahlen, $f \in P_k^{(n)}$ und $\varrho \in R_k^{(m)}$. Wir sagen, daß die Funktion f die Relation ϱ bewahrt, wenn für alle $\underline{a}_1,\underline{a}_2,\ldots,\underline{a}_n \in \varrho$ mit $\underline{a}_i = [a_{1i},a_{2i},\ldots,a_{mi}]$ für i = 1,2,...,n gilt:

$$f(\underline{a}_1,\underline{a}_2,\ldots,\underline{a}_n) := \begin{pmatrix} f(a_{11},a_{12},\ldots,a_{1n}) \\ f(a_{21},a_{22},\ldots,a_{2n}) \\ \vdots \quad \vdots \quad\quad \vdots \\ f(a_{m1},a_{m2},\ldots,a_{mn}) \end{pmatrix} \in \varrho.$$

Wir vereinbaren noch, daß die Relation $\emptyset \in R_k$ von jeder Funktion aus P_k bewahrt wird.

Für $F \subseteq P_k$ und $Q \subseteq R_k$ sei Inv_kF (oder kurz InvF) die Menge aller Relationen aus R_k, die von jeder Funktion aus F bewahrt werden (die Menge der Invarianten von F), und Pol_kQ (oder kurz PolQ) sei die Menge aller Funktionen aus P_k, die jede Relation aus Q bewahren (die Menge der Polymorphismen von Q).

Es gelten die beiden folgenden Lemmata.

Lemma 1: Für alle $Q \subseteq R_k$ ist $Pol_kQ = \bigcap_{\varrho \in Q} Pol_k\varrho$.

Lemma 2: Für alle $Q \subseteq R_k$ ist Pol_kQ eine Funktionenalgebra über E_k.

Wir geben nun einige ein- und zweistellige Operationen für Relationen aus R_k an, bezüglich derer $Inv_k F$ für eine beliebige Teilmenge F von P_k abgeschlossen ist, d. h., wenn o_1 eine beliebige einstellige und o_2 eine beliebige zweistellige dieser Operationen bezeichnet und $\varrho, \mu \in Inv_k F$ gilt, so ist auch $\{o_1\varrho\,,\ \varrho\ o_2\mu\} \subseteq Inv_k F$.

Es sei $\varrho \in R_k^{(m)}$ und $\mu \in R_k^{(n)}$. Dann seien $\zeta\varrho \in R_k^{(m)}$, $\tau\varrho \in R_k^{(m)}$, $\Delta\varrho \in R_k^{(\max(m-1,1))}$, $\varrho \circ \mu \in R_k^{(m+n-2)}$, $\nabla\varrho \in R_k^{(m+1)}$, $pr\,\varrho \in R_k^{(m-1)}$, $\varrho \times \mu \in R_k^{(m+n)}$ und $\sim\varrho \in R_k^{(m+1)}$ die wie folgt definierten Relationen:

$$\zeta\varrho := \{[x_1,x_2,\ldots,x_m] : [x_2,x_3,\ldots,x_m,x_1] \in \varrho\},$$
(zyklisches Vertauschen der Zeilen)

$$\tau\varrho := \{[x_1,x_2,\ldots,x_m] : [x_2,x_1,x_3,x_4,\ldots,x_m] \in \varrho\},$$
(Vertauschen der beiden ersten Zeilen)

$$\Delta\varrho := \{[x_1,x_2,\ldots,x_{m-1}] : [x_1,x_1,x_2,x_3,\ldots,x_{m-1}] \in \varrho\},$$
(Identifizieren der beiden ersten Koordinaten)

$$\varrho \circ \mu := \{[x_1,x_2,\ldots,x_{m+n-2}]: \text{Es gibt ein } u \in E_k \text{ mit } [x_1,x_2,\ldots,x_{m-1},u] \in \varrho \text{ und } [u,x_m,x_{m+1},\ldots,x_{m+n-2}] \in \mu\},$$
(Komposition, Relationenprodukt, Faltung)

$$\nabla\varrho := \{[x_1,x_2,\ldots,x_{m+1}] : [x_2,x_3,\ldots,x_{m+1}] \in \varrho\},$$
(Hinzufügen einer fiktiven Koordinate)

$$pr\,\varrho := \{[x_1,x_2,\ldots,x_{m-1}] : \text{Es gibt ein } u \in E_k \text{ mit } [u,x_1,x_2,\ldots,x_{m-1}] \in \varrho\},$$
(Streichen der ersten Zeile)

$$\varrho \times \mu := \{[x_1,x_2,\ldots,x_{m+n}] : [x_1,x_2,\ldots,x_m] \in \varrho \text{ und } [x_{m+1},x_{m+2},\ldots,x_{m+n}] \in \mu\},$$
(kartesisches Produkt)

$$\sim\varrho := \{[x_1,x_1,x_2,\ldots,x_m] : [x_1,x_2,\ldots,x_m]\in\varrho\},$$

(Verdoppeln der ersten Zeile).

Für eine einstellige Relation $\varrho\in R_k$ oder für $\varrho = \emptyset$ sei $\mathrm{pr}\,\varrho := \emptyset$ und $\zeta\varrho := \tau\varrho := \Delta\varrho := \varrho$. Weiterhin sei $\nabla\emptyset := E_k$ und $\sim\emptyset := \emptyset$. Falls $\emptyset\in\{\varrho,\mu\}$ gilt, sei $\varrho\circ\mu := \varrho\times\mu := \emptyset$, und falls ϱ und μ einstellige Relationen sind, $\varrho\circ\mu := \emptyset$. Abschließend definieren wir noch (übereinstimmend mit /10/)

$$\delta^{\{1;2,3\}}(E_k) := E_k\times(\sim E_k).$$

Es sei $\varrho\in R_k^{(m)}$ und π eine beliebige Permutation der natürlichen Zahlen $1,2,\ldots,m$. Wir definieren
$\pi\varrho := \{[x_1,x_2,\ldots,x_m] : [x_{\pi(1)},x_{\pi(2)},\ldots,x_{\pi(m)}]\in\varrho\}$ und
$\pi\varrho := \emptyset$, falls $\varrho = \emptyset$ gilt. Dann läßt sich bekanntlich $\pi\varrho$ durch geeignete sukzessive Anwendung von ζ und τ auf ϱ erzeugen.
Im folgenden betrachten wir die beiden universalen Algebren

$$\hat{R}_k = [R_k;\delta^{\{1;2,3\}}(E_k),\ \zeta,\ \tau,\ \Delta,\ \circ] \text{ und}$$

$$\underline{R}_k = [R_k;\emptyset,\ \zeta,\ \tau,\ \Delta,\ \nabla,\ \mathrm{pr},\ \sim,\ \times]$$

vom Typ (0,1,1,1,2) bzw. (0,1,1,1,1,1,1,2).
Es sei Q eine Teilalgebra von $\hat{R}_k$ (der vollen Relationenalgebra). Dann ist nach /10/ Q auch bezüglich der Operation ∇, pr, $\sim$ und $\times$ abgeschlossen und $\emptyset\in Q$; folglich ist Q eine Teilalgebra von $\underline{R}_k$.

Lemma 3: Wenn Q eine Teilalgebra von $\underline{R}_k$ ist, so ist Q auch eine Teilalgebra von $\hat{R}_k$.

Beweis: Es genügt zu zeigen, daß die Relation $\delta^{\{1;2,3\}}(E_k)$ in Q enthalten und Q bezüglich der Operation $\circ$ abgeschlossen ist. Q enthält stets die leere Menge. Aus $\nabla\emptyset = E_k$ folgt

$$\delta^{\{1;2,3\}}(E_k) = (\nabla\emptyset)\times(\sim(\nabla\emptyset)).$$

Es seien ϱ und μ zwei beliebige Relationen aus Q. Falls $\emptyset\in\{\varrho,\mu\}$ ist oder ϱ und μ einstellige Relationen sind, so ist

$\varrho \circ \mu = \emptyset$ und in Q enthalten. Wir können also im weiteren voraussetzen, daß ϱ und μ nichtleere Relationen sind mit $\varrho \in R_k^{(m)}$, $\mu \in R_k^{(n)}$ und $m + n > 2$. In Abhängigkeit von m sei π_m folgende Permutation über $\{1,2,\ldots,m+n\}$:

$$\pi_1 = (1),$$
$$\pi_2 = (132),$$
$$\pi_m = \begin{pmatrix} 1 & 2 & \ldots & m-1 & m & m+1 & m+2 & m+3 & \ldots & m+n \\ 3 & 4 & \ldots & m+1 & 1 & 2 & m+2 & m+3 & \ldots & m+n \end{pmatrix} \text{ für } m > 2.$$

Dann ist $\varrho \circ \mu = pr(\Delta(\pi_m(\varrho \times \mu)))$, womit der Beweis des Lemmas abgeschlossen ist.

Damit haben wir festgestellt, daß die Teilalgebren von $\hat{R}_k$ mit denen von $\underline{R}_k$ übereinstimmen. Aus beweistechnischen Gründen werden wir hier vorrangig mit $\underline{R}_k$ arbeiten.

Satz 1: Es gelten die folgenden Aussagen:

(1) Wenn $Q \subseteq R_k$ und $F \in P_k$ ist, so ist $Pol_k Q$ eine Funktionen- und $Inv_k F$ eine Relationenalgebra über E_k.

(2) Wenn Q eine Relationen- und F eine Funktionenalgebra über E_k ist, so gibt es ein $F' \subseteq P_k$ und ein $Q' \subseteq R_k$ mit $Inv_k F' = Q$ und $Pol_k Q' = F$.

(3) Für $Q,Q' \subseteq R_k$ und $F,F' \subseteq P_k$ gilt stets

$Q \subseteq Inv_k Pol_k Q$, $F \subseteq Pol_k Inv_k F$,

wenn $Q' \subseteq Q$ und $F' \subseteq F$, so $Pol_k Q' \supseteq Pol_k Q$ und

$Inv_k F' \supseteq Inv_k F$,

$Pol_k Inv_k Pol_k Q = Pol_k Q$, $Inv_k Pol_k Inv_k F = Inv_k F$.

Mit diesem Satz aus /1/ ist durch Pol und Inv eine Galoiskorrespondenz zwischen den Teilmengen von P_k sowie denen von R_k gegeben (die Funktionenalgebren und die Relationenalgebren über E_k stehen in Galoisbeziehung zueinander).

Unser Ziel besteht nun darin, einen entsprechenden Satz für die iterativen Algebren über E_k zu finden. Dazu muß in geeigneter Weise der Relationenbegriff verallgemeinert werden. Das wird im zweiten Teil dieser Arbeit geschehen.

Der dritte Teil wird den Beweis des entsprechenden Satzes beinhalten. Ehe wir dazu kommen können, bedarf es noch einiger Vorbetrachtungen, die diesen ersten Teil abschließen sollen.

3. Eine Matrizen-Relationenalgebra

Es sei M eine Matrix vom Format (m,r) mit Elementen aus E_k, wobei m und r beliebige natürliche Zahlen sind. Es ist also auch m = o oder r = o zugelassen, wobei dann M zur leeren Menge entartet. Wir nennen M Relationenmatrix, wenn die Spalten von M paarweise voneinander verschieden sind oder wenn $M = \emptyset$ gilt. Mit ϱ_M bezeichnen wir die Menge der Spalten der Relationenmatrix M. Für beliebige natürliche Zahlen m und r sei $\mathfrak{M}_k^{(m,r)}$ die Menge der Relationenmatrizen vom Format (m,r) über E_k, $\mathfrak{M}_k^{(m)} := \bigcup_{r \geq o} \mathfrak{M}_k^{(m,r)}$ und $\mathfrak{M}_k := \bigcup_{m \geq o} \mathfrak{M}_k^{(m)}$. Wir sagen, daß eine Funktion $f \in P_k$ die Relationenmatrix M über E_k bewahrt, wenn f die Relation ϱ_M bewahrt (vgl. § 2). Für $F \subseteq P_k$ bzw. $Q \subseteq \mathfrak{M}_k$ sei $Minv_k F$ die Menge aller Relationenmatrizen aus $\mathfrak{M}_k$, die von jeder Funktion aus F bewahrt werden, bzw. $Mpol_k Q$ die Menge aller Funktionen aus P_k, die jede Relationenmatrix aus Q bewahren.
Analog zu § 2 gelten die beiden folgenden Lemmata.

Lemma 1: Für alle $Q \subseteq \mathfrak{M}_k$ ist $Mpol_k Q = \bigcap_{M \in Q} Mpol_k\{M\}$.

Lemma 2: Für alle $Q \subseteq \mathfrak{M}_k$ ist $Mpol_k Q$ eine Funktionenalgebra über E_k.

Wir geben jetzt einige ein- und zweistellige Operationen für Relationenmatrizen aus $\mathfrak{M}_k$ an, bezüglich derer $Minv_k F$ für eine beliebige Teilmenge F von P_k abgeschlossen ist. (Nach § 2 ist dies leicht einzusehen.)
Es sei E_k die Matrix (0 1 ... k-1), $M \in \mathfrak{M}_k^{(m,r)}$, $M' \in \mathfrak{M}_k^{(n)}$ und

$M = \begin{pmatrix} \underline{z}_1 \\ \underline{z}_2 \\ \vdots \\ \underline{z}_m \end{pmatrix} = [\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_m] = (\underline{s}_1, \underline{s}_2, \ldots, \underline{s}_r)$, wobei die $\underline{z}_i$ ($i = 1,2,\ldots,m$) die Zeilen und die $\underline{s}_j$ ($j = 1,2,\ldots,r$) die Spalten der Matrix M sind.

Dann seien $\zeta M \in \mathfrak{M}_k^{(m)}$, $\tau M \in \mathfrak{M}_k^{(m)}$, $\Delta M \in \mathfrak{M}_k^{(\max(m-1,1))}$, $\nabla M \in \mathfrak{M}_k^{(m+1)}$, $prM \in \mathfrak{M}_k^{(m-1)}$, $\sim M \in \mathfrak{M}_k^{(m+1)}$, $M \times M' \in \mathfrak{M}_k^{(m+n)}$, $\overline{\zeta} M \in \mathfrak{M}_k^{(m)}$ und $\overline{\tau} M \in \mathfrak{M}_k^{(m)}$ die wie folgt definierten Relationenmatrizen:

(1) $\zeta M := [\underline{z}_m, \underline{z}_1, \underline{z}_2, \ldots, \underline{z}_{m-1}]$;

(2) $\tau M := [\underline{z}_2, \underline{z}_1, \underline{z}_3, \underline{z}_4, \ldots, \underline{z}_m]$;

(3) $\sim M := [\underline{z}_1, \underline{z}_1, \underline{z}_2, \underline{z}_3, \ldots, \underline{z}_m]$;

(4) $\overline{\zeta} M := (\underline{s}_r, \underline{s}_1, \underline{s}_2, \ldots, \underline{s}_{r-1})$;

(5) $\tau M := (\underline{s}_2, \underline{s}_1, \underline{s}_3, \underline{s}_4, \ldots, \underline{s}_r)$;

(6) $M \times M'$ ist das Kroneckerprodukt von M mit M';

(7) $\nabla M := (E_k) \times M$;

(8) ΔM entsteht aus M, indem zunächst diejenigen Spalten von M gestrichen werden, die in ihren beiden ersten Koordinaten nicht übereinstimmen, und danach noch die erste Zeile der so entstandenen Matrix gestrichen wird ($\Delta M := \emptyset$, falls bereits nach dem ersten Schritt $\emptyset$ entstanden ist);

(9) prM entsteht aus M, indem zunächst die erste Zeile von M gestrichen wird, und danach in der so entstandenen Matrix alle Spalten, die mehrfach auftreten, mit Ausnahme ihres ersten Auftretens gestrichen werden.

Falls $M \in \mathfrak{M}_k^{(1)}$ oder $M = \emptyset$ gilt, so ist $\zeta M := \tau M := \Delta M := M$ und $prM := \emptyset$. Weiterhin sei $\nabla\emptyset := (E_k)$ und $\sim\emptyset := \emptyset$. Falls $\emptyset \in \{M, M'\}$ ist, wird $M \times M' := \emptyset$ gesetzt. Falls $M \in \mathfrak{M}_k^{(m,1)}$ oder $M = \emptyset$ gilt, sei $\overline{\zeta} M := \overline{\tau} M := M$.

Offensichtlich ergeben die derart definierten Operationen stets Relationenmatrizen.
Wenn wir die universale Algebra (die volle Relationenmatrizenalgebra über E_k)

$$\underline{\mathfrak{M}}_k = [\mathfrak{M}_k;\ \emptyset,\ \zeta,\ \tau,\ \overline{\zeta},\ \overline{\tau},\ \Delta,\ \nabla,\ \mathrm{pr},\ \sim,\ \times]$$

vom Typ (0,1,1,1,1,1,1,1,1,2) betrachten, deren Teilalgebren wir Relationenmatrizenalgebren über E_k nennen, so besteht offenbar in Analogie zu Satz 1 von § 2 der

Satz 1: Es gelten die folgenden Aussagen:

(1) Wenn $Q \subseteq \mathfrak{M}_k$ und $F \subseteq P_k$ ist, so ist $\mathrm{Mpol}_k Q$ eine Funktionen- und $\mathrm{Minv}_k F$ eine Relationenmatrizenalgebra über E_k.

(2) Wenn Q eine Relationenmatrizen- und F eine Funktionenalgebra über E_k ist, so gibt es ein $F' \subseteq P_k$ und ein $Q' \subseteq \mathfrak{M}_k$ mit $\mathrm{Minv}_k F' = Q$ und $\mathrm{Mpol}_k Q' = F$.

(3) Für $Q,Q' \subseteq \mathfrak{M}_k$ und $F,F' \subseteq P_k$ gilt stets

$Q \subseteq \mathrm{Minv}_k \mathrm{Mpol}_k Q$, $F \subseteq \mathrm{Mpol}_k \mathrm{Minv}_k F$,

wenn $Q' \subseteq Q$ und $F' \subseteq F$, so $\mathrm{Mpol}_k Q' \supseteq \mathrm{Mpol}_k Q$ und

$\mathrm{Minv}_k F' \supseteq \mathrm{Minv}_k F$,

$\mathrm{Mpol}_k \mathrm{Minv}_k \mathrm{Mpol}_k Q = \mathrm{Mpol}_k Q$, $\mathrm{Minv}_k \mathrm{Mpol}_k \mathrm{Minv}_k F = \mathrm{Minv}_k F$.

Mit diesem Satz ist durch Mpol_k und Minv_k eine Galoiskorrespondenz zwischen den Teilmengen von P_k sowie denen von $\mathfrak{M}_k$ gegeben (die Funktionenalgebren von $\hat{P}_k$ und die Relationenmatrizenalgebren von $\underline{\mathfrak{M}}_k$ stehen in Galoisbeziehung zueinander).

Literatur

/1/ Боднарчук, В. Г., Калужнин, Л. А., Котов, В. Н., и Ромов, Б. А.: Теория Галуа для алгебр Поста I, II. Кибернетика 3, 1 - 10 (1969); 5, 1 - 9 (1969)

/2/ Gössel, M., und Pöschel, R.: Invariant Relations for Automata - A Proposal. Elektron. Informationsverarb. Kybernet. 16, 147 - 169 (1980)

/3/ Harnau, W.: Ein verallgemeinerter Relationen- und ein modifizierter Superpositionsbegriff für die Algebra der mehrwertigen Logik. Dissertation (B), Wilhelm-Pieck-Universität Rostock 1983

/4/ Hikita, T.: Completeness Properties of k-valued Functions with Delays - Inclusions among Closed Spectra. Preprint Metropolitan Univ. Tokyo 1980

/5/ Яблонский, С. В.: Функциональные построения в к-значной логике. Труды Мат. Инст. Стеклова 51, 5 - 142 (1985)

/6/ Кузнецов, А. В.: Структуры с замыканием и критерии функциональной полноты. Успехи Мат. Наук 16(98), 201 - 202 (1961)

/7/ Lugowski, H.: Grundzüge der universellen Algebra. Leipzig 1976

/8/ Мальцев, А. И.: Итеративные алгебры и многообразия Поста. Алгебра и Логика 5, 2, 5 - 24 (1966)

/9/ Nozaki, A.: Some New Results on Delayed Logic. Preprint ICU Tokyo 1981

/10/ Pöschel, R., und Kalužnin, L. A.: Funktionen- und Relationenalgebren. Berlin 1979

/11/ Rosenberg, I. G.: La structure des fonctions de plusieurs variables sur un ensemble fini. C. R. Acad. Sci. Paris 260, 3817 - 3819 (1965)

/12/ Rosenberg, I. G.: Über die Funktionale Vollständigkeit in den mehrwertigen Logiken. Rozpravy Československé Akad. Věd Řada Mat. Přírod. Věd 80, 3 - 93 (1970)

/13/ Skornjakow, L. A.: Elemente der Verbandstheorie. Berlin 1973

eingegangen: 05. 11. 1984

Anschrift des Verfassers:

Dr. sc. nat. W. Harnau
Pädagogische Hochschule "K. F. W. Wander"
Sektion Mathematik
Wigardstr. 17, PSF 365
DDR-8060 Dresden

Rostock. Math. Kolloq. 28, 18 - 32 (1985)

05C80
60C05

Claudia Kaiser
Karl Weber

Degrees and Domination Number of Random Graphs in the n-Cube

Preliminaries

The n-cube E^n is the graph consisting of the 2^n vertices $\underline{a} = (a_1,\ldots,a_n)$, $a_i \in \{0,1\}$, and the $n2^{n-1}$ edges between vertices differing in exactly one coordinate. A spanning subgraph g of E^n has the same vertex set as E^n. An induced subgraph f of E^n (also called a Boolean function) with the vertex set $A \subseteq E^n$ contains exactly those edges of E^n between vertices of A. (Note that by E^n or f are not only denoted the graphs but also its vertex sets, g stands also for the edge set of g.) Choosing the edges of g (the vertices of f) at random we arrive at a random spanning subgraph (a random Boolean function). When the edges of g (the vertices of f) are chosen independently and with the same probability p, then the probabilities are defined as $P(g) = p^{|g|}q^{n2^{n-1}-|g|}$ ($P(f) = p^{|f|}q^{2^n-|f|}$), where $q = 1 - p$. We say almost all g (f) have a certain property Q or the property Q is almost surely if $P(g \text{ has } Q) \to 1$ ($P(f \text{ has } Q) \to 1$) as $n \to \infty$.

All limits, asymptotics etc. are understood as $n \to \infty$. For $\alpha = \alpha(n)$ and $\beta = \beta(n)$ we write $\alpha \sim \beta$ if $\alpha/\beta \to 1$ (α and β are asymptotically equal), $\alpha \asymp \beta$ if $\alpha = O(\beta)$ and $\beta = O(\alpha)$ (α and β are of the same order), $\alpha \lesssim \beta$ if $\alpha \leq \beta(1 + o(1))$, $\alpha \lesssim \beta$ if $\alpha = O(\beta)$ and sometimes $\alpha << \beta$ if $\alpha = o(\beta)$. Everywhere φ denotes a sequence $\varphi(n)$ tending to infinity - arbitrarily slowly unless otherwise specified - as $n \to \infty$. The natural and the binary logarithm are denoted by ln and log, respectively. For any real x, $\lfloor x \rfloor$ denotes the greatest integer not greater than x. The distance $\varrho(\underline{a},\underline{b})$ between two vertices $\underline{a},\underline{b} \in E^n$ is the customary Hamming distance, i.e. the number of coordinates in which $\underline{a}$ and $\underline{b}$ differ.

Let X be a non-negative integer-valued random variable, EX the expectation and D^2X the variance of X. Then the following bound is immediate from the inequality of Markov:

$$P(X \leq \varphi \cdot EX) \geq 1 - 1/\varphi \to 1. \tag{1}$$

In particular, (1) implies

$$P(X = 0) \to 1 \text{ if } EX \to 0. \tag{2}$$

From the inequality of Chebyshev we deduce

$$P(X \sim EX) \to 1, \tag{3}$$

i.e. there is an $\varepsilon = \varepsilon(n) \to 0$ with $P(|X - EX| < \varepsilon EX) \to 1$, if

$$D^2X = o((EX)^2). \tag{4}$$

Recall that $D^2X = EX^2 - (EX)^2 = EX + E(X)_2 - (EX)^2$, where $E(X)_2 = E(X(X-1))$ is the second factorial moment of X.
Given the random variables $X_0, X_1, X_2, \ldots$ and the integer sequences $\alpha = \alpha(n)$ and $\beta = \beta(n)$, $0 \leq \alpha \leq \beta$. Then

$$\sum_{i=\alpha}^{\beta} \frac{D^2X_i}{(EX_i)^2} \to 0 \tag{5}$$

implies that

$$P(X_i \sim EX_i) \to 1 \text{ for all } i = \alpha, \alpha+1, \ldots, \beta, \tag{6}$$

or more precisely, there is an $\varepsilon = \varepsilon(n) \to 0$ such that

$$P\Big(\bigcap_{i=\alpha}^{\beta} \{|X_i - EX_i| \leq \varepsilon EX_i\}\Big) \to 1. \tag{7}$$

Indeed, by the inequality of Chebyshev

$$P\Big(\bigcup_{i=\alpha}^{\beta} \{|X_i - EX_i| \geq \varepsilon EX_i\}\Big) \leq \sum_{i=\alpha}^{\beta} P(|X_i - EX_i| \geq \varepsilon EX_i) \leq \sum_{i=\alpha}^{\beta} \frac{D^2X_i}{\varepsilon^2 (EX_i)^2},$$

and assuming (5) we may find an $\varepsilon \to 0$ so that the last sum tends to zero, too. Now (7) follows immediately.
The well-known theorems on the approximation of the binomial distribution by the normal distribution are usually formulated for a fixed probability p. That both the local and the central (De Moivre - Laplace) limit theorem also hold for

$p = p(n) \to 0$ or 1, respectively, can easily be verified assuming that $pqn \to \infty$. That means, setting $b(i; n,p) = \binom{n}{i} p^i q^{n-i}$ and assuming that $pqn \to \infty$, x, x_1, $x_2 = o((pqn)^{1/6})$, $x_1 < x_2$, we have

$$b(pn + x\sqrt{pqn}\ ;\ n,p) \sim e^{-x^2/2}/\sqrt{2\pi pqn} \tag{8}$$

and

$$\sum_{pn + x_1\sqrt{pqn} \le i \le pn + x_2\sqrt{pqn}} b(i; n,p) \sim \Phi(x_2) - \Phi(x_1), \tag{9}$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2}\, du$ is the distribution function of the normalized normal distribution (compare /3/, pp. 133 - 137; the proofs can without trouble be carried over to the case $pqn \to \infty$ and x, x_1, $x_2 = o((pqn)^{1/6})$). By (9) for $x = O(1)$ it is evident that

$$\sum_{i \ge pn + x\sqrt{pqn}} b(i; n,p) \sim 1 - \Phi(x). \tag{10}$$

But for $x \to \infty$ the relation (10) is the statement of the theorem on large deviations. This theorem is especially usefull, since for $x \to \infty$ it holds simultanously

$$1 - \Phi(x) \sim e^{-x^2/2}/\sqrt{2\pi}\,x \tag{11}$$

(cf. /3/, p. 131). In order to show that (10) remains true for $pqn \to \infty$ and $x = o((pqn)^{1/6})$ we can also follow Feller in the general line (cf. /3/, pp. 144 - 145) but we have a little more trouble with the details. Assuming $0 < x_1 < x_2 = o((pqn)^{1/6})$ by (9)

$$\sum_{pn + x_1\sqrt{pqn} \le i \le pn + x_2\sqrt{pqn}} b(i; n,p) \sim \Phi(x_2) - \Phi(x_1)$$

$$= (1 - \Phi(x_1)) - (1 - \Phi(x_2)) \sim 1 - \Phi(x_1) \text{ if } x_1,\ x_2,\ x_2 - x_1 \to \infty:$$

by (11) we have $1 - \Phi(x_2) \sim e^{-x_2^2/2}/\sqrt{2\pi}\,x_2 = o(e^{-x_1^2/2}/\sqrt{2\pi}\,x_1)$. Thus we derived the lower bound of (10). Furthermore, it holds

$\sum_{i \geq pn + x_2\sqrt{pqn}} b(i; n,p) = o(1 - \Phi(x_1))$. Indeed, using the well-known estimate $\sum_{i \geq t} b(i; n,p) \leq \frac{tq}{t-pn}\, b(t; n,p)$, $t > pn$, and setting $t = pn + x_2\sqrt{pqn}$, by (8) and (11) we obtain that $\sum_{i \geq t} b(i; n,p)$

$\leq \frac{tq}{t-pn}\, b(t; n,p) \lesssim \frac{tq}{t-pn} \frac{e^{-x_2^2/2}}{\sqrt{2\pi\, pqn}} = \frac{t}{pn}(1 - \Phi(x_2)) \sim 1 - \Phi(x_2)$.

But, that $1 - \Phi(x_2)) = o(1 - \Phi(x_1))$ we already have seen above. Note that from (10) the following bound can easily be derived

$$\sum_{i \leq pn - x\sqrt{pqn}} b(i; n,p) \sim 1 - \Phi(x), \tag{12}$$

assuming that $pqn \to \infty$ and $x = o((pqn)^{1/6})$. Indeed, by (10):

$$1 - \Phi(x) \sim \sum_{j \geq qn + x\sqrt{pqn}} b(j; n,q) = \sum_{n-j \leq pn - x\sqrt{pqn}} b(j; n,q)$$

$$= \sum_{i \leq pn - x\sqrt{pqn}} b(n-i; n,q) = \sum_{i \leq pn - x\sqrt{pqn}} b(i; n,p).$$

1. Vertex-degrees

We begin with random spanning subgraphs. Let $v'_{\underline{a}}(g)$ denote the degree of $\underline{a} \in E^n$ in the graph g, $\delta'(g) = \min_{\underline{a} \in E^n} v'_{\underline{a}}(g)$ (minimum degree), $\Delta'(g) = \max_{\underline{a} \in E^n} v'_{\underline{a}}(g)$ (maximum degree) and $X_i(g)$ = number of vertices of degree i. Then

$$EX_i = 2^n b(i; n,p) \tag{13}$$

is evident.

Lemma 1:

$$D^2 X_i \leq \left(\frac{1}{EX_i} + \frac{n}{2^n \min\{p,q\}}\right)(EX_i)^2, \quad i = 0, \ldots, n. \tag{14}$$

Proof: For $\underline{a} \in E^n$ define the random variables $X_i^{\underline{a}}(g) = 1$ if $v'_{\underline{a}}(g) = i$ and $X_i^{\underline{a}}(g) = 0$ otherwise. Then $E(X_i)_2 = \sum E(X_i^{\underline{a}} X_i^{\underline{b}})$ taken over all ordered pairs of different vertices $\underline{a}, \underline{b} \in E^n$. Put $\sum_1 = \sum_{\varrho(\underline{a},\underline{b})=1} E(X_i^{\underline{a}} X_i^{\underline{b}})$ and $\sum_2 = \sum_{\varrho(\underline{a},\underline{b})\geq 2} E(X_i^{\underline{a}} X_i^{\underline{b}})$. Then for $i = 0$ we have $\sum_1 = n2^n q^{2n-1} = \frac{n}{q2^n} (EX_0)^2$ and analogously for $i = n$: $\sum_1 = \frac{n}{p2^n} (EX_n)^2$ and for $i \neq 0,n$:

$$\sum_1 = n2^n (p(b(i-1;n-1,p))^2 + q(b(i;n-1,p))^2)$$

$$= n2^n (b(i;n,p))^2 ((\tfrac{i}{n})^2/p + (\tfrac{n-i}{n})^2/q) = \frac{n}{2^n \min\{p,q\}} (EX_i)^2.$$

If $\varrho(\underline{a},\underline{b}) \geq 2$ then $E(X_i^{\underline{a}} X_i^{\underline{b}}) = (EX_i^{\underline{a}})(EX_i^{\underline{b}})$ such that $\sum_2 \leq (EX_i)^2$. In accordance with $D^2X_i = EX_i + E(X_i)_2 - (EX_i)^2$ it follows (14). q.e.d.

Now for $p = 1/2$ the following theorem can be obtained easily.

Theorem 2: Let $p = 1/2$. Then almost all g have asymptotically $\binom{n}{i}$ vertices of degrees $i = 1, \ldots, n-1$ (in the sense of (7)). The minimum degree is 0 or 1, and the maximum degree is n-1 or n. More precisely it holds $P(\delta' = 0) = P(\Delta' = n) \sim 1 - 1/e$ and $P(\delta' = 1) = P(\Delta' = n-1) \sim 1/e$.

Proof: Assuming $p = q = 1/2$, we know by (13) and (14) that $EX_i = \binom{n}{i}$ and $D^2X_i \leq (\frac{1}{\binom{n}{i}} + \frac{2n}{2^n})(EX_i)^2$. Thus

$\sum_{i=1}^{n-1} \frac{D^2X_i}{(EX_i)^2} \to 0$, so that the first part of the theorem is clear by (5) - (7).

Further we know (see /2/ or /9/, Remark after Theorem 2.3) that the number X_0 of isolated vertices is asymptotically Poisson

distributed with expectation $\lambda = 1$, and our assertion on δ' follows from $P(\delta' = 0) = P(X_0 \geq 1)$.[1]

Finally we have for the complement $\bar{g}$ of g ($\bar{g}$ contains exactly those edges of E^n which are not in g) $P(\bar{g}) = P(g)$ what yields the assertion concerning Δ'.

q.e.d.

Using the customary notation $H(x) = -x \log x - (1-x)\log(1-x)$ for the entropy function we obtain the following

Theorem 3: Let $p > 1/2$ be fixed. Then for almost all g we have $\Delta'(g) = n$ and

$$\left| \delta'(g) - \left(c_p n + \frac{\log n}{2\log \frac{p(1-c_p)}{c_p q}}\right) \right| < \varphi, \tag{15}$$

where c_p is the root of $h(x) = 1+H(x)+x \log p + (1-x)\log q$, $0 < x < 1$ (cf. Fig. 1).

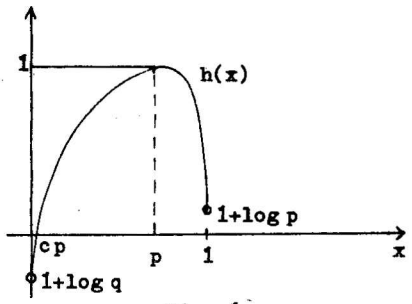


Fig. 1

Proof: $EX_i = 2^n b(i; n,p)$ is concave relative to i and takes on its maximum value for $i = \lfloor pn \rfloor$ or $i = \lfloor pn \rfloor + 1$. Clearly $EX_i \to \infty$ for $i = n$ what together with (14) and (3), (4) implies the assertion on Δ'. Put $i = \alpha n$, $1/n \leq \alpha \leq p$. Then by Stirling's formula $\binom{n}{i} \sim 2^{nH(\alpha)} / \sqrt{2\pi n\alpha(1-\alpha)}$ and

$$\log EX_i = n(1+H(\alpha) + \alpha \log p + (1-\alpha)\log q) - \tfrac{1}{2}\log n + \tfrac{1}{2}\log(1/\alpha) + O(1). \tag{16}$$

---

1 In /5/ the authors determine for given fixed degree $i=1,2,\ldots$ and certain degree sequences $i = i(n) \to \infty$ such probabilities p for which the random variable X_i is asymptotically Poisson or normally distributed, respectively.

Denoting $h(\alpha) = 1 + H(\alpha) + \alpha \log p + (1-\alpha)\log q$ and setting

$$\alpha = c_p + \frac{\log n + \delta}{2n \log \frac{p(1-c_p)}{c_p q}},$$

where δ is a real constant and c_p is the root of $h(x)$, $h(\alpha) = \frac{\log n + \delta}{2n} + O((\frac{\log n}{n})^2)$ is easily calculated. That means, in accordance with (16) we have $\log EX_i \to \infty (-\infty)$ if we replace the constant δ by $\varphi(-\varphi)$. Thus we derived (15) too.

q.e.d.

The following table is taken from /4/. It shows how the root c_p depends on p.

p	lower bound for c_p	upper bound for c_p
0,55	0,020.812.3	0,020.812.4
0,60	0,051.305.3	0,051.305.4
0,65	0,089.463.0	0,089.463.1
0,70	0,135.243.4	0,135.243.5
0,75	0,189.289.6	0,189.289.7
0,80	0,252.980.2	0,252.980.3
0,85	0,328.926.7	0,328.926.8
0,90	0,422.509.7	0,422.509.8
0,95	0,548.181.8	0,548.181.9

That in fact $c_p \to 1$ if $p \to 1$ we obtain, for example, from the estimate $c_p \geq p/2$ which is true for $p \geq 14/15$. (The function $y(p) = h(p/2)$ is strictly decreasing and $y(14/15) < 0$.) It is not difficult to derive non-trivial lower and upper bounds for c_p in the general case. But they are not very comfortable, and since we make no further use we omit them.

Corollary 4: Let $p < 1/2$ be fixed. Then for almost all g we have $\delta'(g) = 0$ and

$$\left|\Delta'(g) - (c_p^* n - \frac{\log n}{2 \log \frac{qc_p^*}{(1-c_p^*)p}})\right| < \varphi, \qquad (17)$$

where $c_p^* = 1 - c_{1-p}$ (c_{1-p} is defined as in Theorem 3).

Proof: Let $p > 1/2$ be fixed. Then by Theorem 3 $\Delta'(g) = n$ and $\delta'(g)$ satisfies (15) almost surely. Thus choosing the edges of $\bar{g}$ with probability $q = 1-p$ for almost all $\bar{g}$ we have $\delta'(\bar{g}) = 0$ and $\Delta'(\bar{g}) = n - \delta'(g) = n - (c_p n + \log n / 2 \log(p(1-c_p)/c_p q) \pm \varphi)$

$$= (1 - c_{1-q})n - \log n / 2 \log(p(1 - c_{1-q})/c_{1-q} q)) \pm \varphi$$

if $q < 1/2$.
Changing p and q and writing g for $\bar{g}$ we obtain just (17).

q.e.d.

Now let us consider the case $p \to 0$.

Theorem 5: Let $p \to 0$ such that $p \geq 2^{-o(n)}$.[2] Then for almost all g we have $\delta'(g) = 0$ and $\Delta'(g) = o(n)$ or more precisely

$$\Delta'(g) \sim n/\log(1/p). \qquad (18)$$

Proof: The assertion on δ' is evident. In order to prove (18) let $i = \delta\, n/\log(1/p)$, δ a real constant. Then by (13) and $\log\binom{n}{i} = i \log n - i \log i + O(1)$ it follows that

$$\log EX_i = n(1 - \delta + \frac{\log\log(1/p)}{\log(1/p)} - \log(1/q) + O(\frac{1}{\log(1/p)}). \qquad (19)$$

Hence $\log EX_i \to -\infty$ if $\delta > 1$ and $\log EX_i \to +\infty$ if $\delta \leq 1$. (Note that $\log(1/q) = \log(\frac{1}{1-p}) = p \log e + O(p^2)$ and $1/\log(1/p) >> p$, and $p \geq 2^{-o(n)}$ implies $n/\log(1/p) \to \infty$.) Clearly, in the last case $D^2X_i = o((EX_i)^2)$ is also satisfied (cf. (14)) what completes the proof of (18). q.e.d.

---

2 It is easy to check that if $\log(1/p) \asymp n$ then all degrees are bounded by a constant almost surely.

A more sophistic analysis of (19) shows that Δ' is in fact considerably larger than $\lfloor n/\log(1/p) \rfloor$ but we are content with the main term of the asymptotic expansion of Δ'.

Corollary 6: Let $p \to 1$ so that $q \geq 2^{-o(n)}$. Then for almost all g we have $\Delta'(g) = n$ and

$$\delta'(g) = n - (1+o(1))n/\log(1/q). \tag{20}$$

All we have to do for the proof is to repeat the above deduction from $p > 1/2$ to $p < 1/2$.
Finally we show that, assuming $pn \to \infty$, asymptotically all 2^n vertices have a degrees close to the average degree pn almost surely.

Theorem 7: Let $pn \to \infty$. Then asymptotically 2^n vertices have a degree asymptotically equal to pn almost surely. (More precisely there are $\varepsilon_1, \varepsilon_2 \to 0$ such that for almost all g more than $2^n(1-\varepsilon_1)$ vertices $\underline{a}$ have a degree $v'_{\underline{a}}(g)$ satisfying $|v'_{\underline{a}}(g) - pn| < \varepsilon_2 pn$.)

Proof: For every fixed vertex $\underline{a}$ the random variable $v'_{\underline{a}}$ is binomially distributed with parameters $\bar{v}' = Ev'_{\underline{a}} = pn$ and $D^2v' = pqn$ (we omit the index $\underline{a}$ since both the expectation and the variance are independent of $\underline{a}$). Let $\varepsilon \to 0$ be given and define $b(g) = |\{\underline{a}: |v'_{\underline{a}}(g) - \bar{v}'| \geq \varepsilon\bar{v}'\}|$. Then for the expectation $\bar{b} = Eb$ by the inequality of Chebyshev we find

$$\bar{b} = 2^n P(\text{for a fixed vertex } \underline{a} \mid |v'_{\underline{a}}(g) - \bar{v}'| \geq \varepsilon\bar{v}' \text{ is fulfiled})$$

$$\leq 2^n \frac{D^2v'}{\varepsilon^2\bar{v}'^2} = 2^n \frac{q}{\varepsilon^2 pn}.$$

Thus by (1) $b(g) \leq \frac{\varphi q}{\varepsilon^2 pn} 2^n$ almost surely. Since $pn \to \infty$ we can find $\varepsilon_1, \varepsilon_2 \to 0$ such that $1/\varepsilon_2^2 pn \to 0$ too and $\varepsilon_1 >> 1/\varepsilon_2^2 pn$.

q.e.d.

We remark that this situation is quite different from that for traditional random graphs (i.e. random spanning subgraphs of

the complete graph K_n). In this case not only asymptotically all but even all n vertices have a degree which is asymptotically equal to the average degree pn (cf. /1/).
Now let us formulate the corresponding results for random Boolean functions. Recall that f has almost surely asymptotically $p2^n$ vertices (instead of the 2^n vertices of random spanning subgraphs). Further in this case the concept of the complementary graph can not be used such that we have to verify the corresponding statements to Corollaries 4 and 6 in another way. Let $v_{\underline{a}}(f)$ be the degree of the vertex $\underline{a} \in f$ in the graph f and $X_i(f)$ the number of vertices $\underline{a} \in f$ with degree i. Minimum and maximum degree are denoted by δ and Δ, respectively. Then instead of (13) we have

$$EX_i = p2^n\, b(i;\, n,p), \tag{21}$$

and Lemma 1 has to be replaced by

Lemma 8:

$$D^2X_i \leq (\frac{1}{EX_i} + \frac{2n^2}{2^n \min\{p^2,q^2\}})(EX_i)^2, \quad i = 0, \ldots, n. \tag{22}$$

We omit the calculations since they are similar to those in the proof of Lemma 1. The most important difference is that here $E(X_i^{\underline{a}} X_i^{\underline{b}}) = (EX_i^{\underline{a}})(EX_i^{\underline{b}})$ is true only for $\varrho(\underline{a},\underline{b}) \geq 3$. Let us list the results on degrees of random Boolean functions.

1. If $p = 1/2$ then almost all f contain asymptotically $\binom{n}{i}/2$ vertices of degree $i = 1, \ldots, n-1$. The minimum degree is 0 or 1 and the maximum degree is n-1 or n. More precisely it holds $P(\delta = 0) \sim P(\Delta = n) \sim 1 - 1/\sqrt{e}$ and $P(\delta = 1) \sim P(\Delta = n-1) \sim 1/\sqrt{e}$ (cf. Theorem 2).

The number of isolated vertices is Poisson distributed with expectation $\lambda = 1/2$ (cf. /8/, Theorem 3). The same limit distribution is easily obtained for the number of vertices with degree n. But we are not able to show that $P(\delta = 0) = P(\Delta = n)$ and $P(\delta = 1) = P(\Delta = n-1)$ (compare the proof of Theorem 2, where that is easily obtained by the consideration of complementary graphs).

2. Theorem 3 as well as Corollary 4 remain true for almost all f.

The summand log p has no influence to formula (16) and consequently Theorem 3 follows. The corollary we have to prove independently of the concept of complementary graph. But it may easily be derived making use of the summetry of the function $h(x) = 1 + H(x) + x \log p + (1-x)\log q$.

Assuming $p < 1/2$ we may apply the calculations from the proof of Theorem 3 to $h(1-\alpha) = 1 + H(\alpha) + \alpha \log q + (1-\alpha)\log p$ (here q plays the role of p): For

$$\alpha = c_q + \frac{\log n + \delta}{2 \log \frac{q(1-c_q)}{c_q p}}$$

we have $h(1-\alpha) = \frac{\log n + \delta}{2n} + O((\log n / n)^2)$, i.e. $h(\alpha)$ takes on this value for $\alpha = (1-c_q) - (\log n + \delta)/(2\log(q(1-c_q)/c_q p))$ $= c_p^* - (\log n + \delta)/(2\log(qc_p^*/(1-c_p^*)p))$.

3. Theorem 5 and Corollary 6 remain true for almost all f.

In (19) we have to add the term log p but nevertheless it holds $\log EX_i \to -\infty$ for $\delta > 1$. (20) is shown as (18) (again without using the complementary graphs).

4. Let $pn \to \infty$. Then for almost all f asymptotically all $|f| (\sim p2^n)$ vertices of f have a degree which is asymptotically equal to pn (i.e. there are $\varepsilon_1, \varepsilon_2 \to 0$ such that more than $|f| - \varepsilon_1 p 2^n$ vertices $\underline{a} \in f$ satisfy $|v_{\underline{a}}(f) - pn| < \varepsilon_2 pn$) (cf. Theorem 7).

Assuming $\underline{a} \in f$ the random variable $v_{\underline{a}}$ is binomially distributed, and thus $\bar{v} = pn$ and $D^2 v = pqn$ are the conditional expectation and variance, respectively. Now we define for a given $\varepsilon \to 0$ the random variable b(f) as the number of $\underline{a} \in f$ with $|v_{\underline{a}}(f) - \bar{v}| \geq \varepsilon \bar{v}$, and similar to the above consideration we obtain

$\overline{b} = 2^n$ P(the fixed vertex $\underline{a}$ is in f and $|v_{\underline{a}}(f)-\overline{v}| \geq \varepsilon\overline{v}$)

$= p2^n$ P($|v_{\underline{a}}(f) - \overline{v}| \geq \varepsilon\overline{v}$ under the condition that $\underline{a} \in f$)

$\leq p2^n \frac{D^2 v}{\varepsilon^2 \overline{v}^2} = p2^n \frac{q}{\varepsilon^2 pn}$. Arguing as above our assertion follows.

2. Domination number

A subset A of the vertex set V(G) of a graph G is called dominating if all vertices of V(G)-A are adjacent with at least one vertex of A. The domination number dom (G) is the minimum cardinality of a dominating set. The domination number may be interpreted as a vertex covering number too: dom(G) is the minimum number of stars in G covering the vertex set of G. (Estimates for the domination number of traditional random graphs are contained in /10/.)

Let us begin again with random spanning subgraphs. After that we formulate the corresponding results for random Boolean functions. The domination numbers are denoted by dom'(g) and dom(f), respectively. We start with two trivial lower bounds. The bound

$$\text{dom}'(g) \geq 2^n/(\Delta'(g) + 1) \tag{23}$$

is evident since every vertex is dominated by itself. If there are exactly r vertices $\underline{a}$ with $v'_{\underline{a}}(g) > j$, then clearly

$$\text{dom}'(g) \geq \frac{2^n - r(\Delta'(g) + 1)}{j}. \tag{24}$$

To construct dominating sets of small cardinality we use a greedy algorithm: In each step we choose such a vertex which dominates the most of non-dominated vertices up to this step. By the greedy lemma of A. A. Saposhenko (cf. /6/, Lemma 1 or /7/, Lemma 5) we obtain the following upper bound.

Lemma 9: Assume that there are $(1-\varepsilon)2^n$ vertices $\underline{a}$ with $v'_{\underline{a}}(g) \geq j$. Then the prescribed greedy algorithm yields a dominating set of g containing no more than

$$\frac{2^n}{j}(\ln j + 1) + \varepsilon 2^n + 1 \tag{25}$$

vertices.

Theorem 10: For the domination number $\mathrm{dom}'(g)$ of almost all g the following estimates are satisfied:

(i) $\mathrm{dom}'(g) \sim 2^n$ if $p << 1/n$,

(ii) $2^n \log(1/p)/n \lesssim \mathrm{dom}'(g) \lesssim 2^n \ln pn / pn$

if $1/n << p \lesssim (\ln n)^3/n$ and

(iii) $2^n/pn \lesssim \mathrm{dom}'(g) \lesssim 2^n \ln pn / pn$ if $p >> (\ln n)^3/n$.

From (iii) we deduce for $p \asymp 1$: $\mathrm{dom}'(g) \lesssim 2^n \ln n / pn$ and, in particular, for $p \to 1$: $2^n/n \lesssim \mathrm{dom}'(g) \lesssim 2^n \ln n / n$.

Proof: If $p << 1/n$ then the expectation of the number of edges is $E|g| = n2^{n-1} p = o(2^n)$, i.e. by (1) asymptotically all 2^n vertices of g are isolated almost surely, and (i) is clear. The lower bound in (ii) follows from (23) and Theorem 5. The upper bound in (ii) and (iii) follows from Lemma 9, since assuming $pn \to \infty$ almost surely more than $(1 - \frac{\varphi q}{pn})2^n$ vertices $\underline{a}$ have a degree $v'_{\underline{a}}(g) \sim pn$ (cf. the proof of Theorem 7).

Now let us show how the lower bound in (ii) can be improved for $p >> (\ln n)^3/n$ using (24) and (10). We may assume that $q \asymp 1$ since for $p \to 1$ the lower bound $\mathrm{dom}'(g) \gtrsim 2^n/n$ from (iii) is trivial. Let $Y_j(g) = \sum_{i \geq j} X_i(g)$. By (13) we have $EY_j = \sum_{i \geq j} EX_i$ $= 2^n \sum_{i \geq j} b(i; n,p)$ and setting $j = pn + x\sqrt{pqn}$, $x \to \infty$ and $x = o((pqn)^{1/6})$ it follows from (10) and (11) that $EY_j \sim 2^n e^{-x^2/2}/\sqrt{2\pi}x$. If $p \asymp 1$ then by setting $x = \sqrt{2 \ln n}$ we obtain $EY_j = o(2^n/n)$. Thus by (1) there are almost surely only $o(2^n/n)$ vertices $\underline{a}$ with $v'_{\underline{a}}(g) \geq pn + \sqrt{2(\ln n)pqn}$ and consequently (24) yields

$$\mathrm{dom}'(g) \geq \frac{2^n - o(2^n/n)(n+1)}{pn + \sqrt{2(\ln n)pqn}} \gtrsim 2^n/pn.$$

If $p \to 0$ then (in accordance with $\Delta' \sim n/\log(1/p)$ in this case- cf. (18)) we set $x = \sqrt{2 \ln(n/\log(1/p))}$, and with $EY_j = o(2^n \log(1/p)/n)$ we obtain $dom'(g) \geq 2^n/pn$ again by (24). But in order to ensure that $\sqrt{\ln(n/\log(1/p))} = o((pqn)^{1/6})$ we have to suppose that $p \gg (\ln n)^3/n$. This completes the proof of our theorem.

q.e.d.

Of course, if $p \to 1$ not too fast (e.g. $p \leq 1 - \varphi(\ln n)^3/n$) then the upper bound in (ii) and (iii), respectively, may also be derived from the theorem of large deviations: By (12) and (11) there are almost surely only $o(2^n/pn)$ vertices a of degree $v'_a(g) \leq pn - \sqrt{2 \ln(pn) pqn}$, and (24) yields the upper bound as claimed. But for $p \to 1$ this upper bound is trivial by (25) and $d'(g) \sim n$ (cf. Corollary 6).

Finally we formulate the corresponding results for random Boolean functions.

Theorem 11: For the domination number $dom(f)$ of almost all f we have the following estimates:

(i) $dom(f) \sim p2^n$ if $p \ll 1/n$,

(ii) $p2^n \log(1/p)/n \lesssim dom(f) \lesssim 2^n \ln(pn)/n$ if $1/n \ll p \lesssim (\ln n)^3/n$ and

(iii) $2^n/n \lesssim dom(f) \lesssim 2^n \ln(pn)/n$ if $p \gg (\ln n)^3/n$.

The theorem can be proven as Theorem 9.
In particular, if $p \asymp 1$ (iii) implies $2^n/n \lesssim dom(f) \lesssim 2^n \ln n/n$ almost surely.

References

/1/ Bollobás, B.: Degree sequences of random graphs. Discrete Math. 33, 1 - 19 (1981)

/2/ Erdös, P., and Spencer, J.: Evolution of the n-cube. Comput. Math. Appl. 5, 33 - 39 (1979)

/3/ Feller, W.: An Introduction to Probability Theory and its Applications. Vol. 1. New York 1950

/4/ Kaiser, C.: Properties of random spanning subgraphs of the n-cube. Diplomarbeit, Wilhelm-Pieck-Universität Rostock 1983 (in German)

/5/ Palka, Z., and Ruciński, A.: Vertex-degrees in a random subgraph of a regular graph. Preprint 1983

/6/ Saposhenko, A. A.: On the complexity of disjunctive normal forms obtained by a greedy algorithm (in Russian). Diskret. Analiz. 21, 62 - 71 (1972)

/7/ Weber, K.: The length of random Boolean functions. Elektron. Informationsverarb. Kybernet. 18, 659 - 668 (1982)

/8/ -: Subcubes of random Boolean functions. Elektron. Informationsverarb. Kybernet. 19, 365 - 374 (1983)

/9/ -: Subcube coverings of random spanning subgraphs of the n-cube. Math. Nachrichten 120, 327 - 345 (1985)

/10/ -: Domination number for almost every graph. Rostock. Math. Kolloq. 16, 31 - 43 (1981)

received: 28. 01. 1985

Authors addresses:

Dipl.-Lehrer C. Kaiser
Lilienthalstraße 12
DDR-8900 Görlitz

Dr. K. Weber
Wilhelm-Pieck-Universität
Rostock
Sektion Mathematik
Universitätsplatz 1
DDR-2500 Rostock

Rostock. Math. Kolloq. 28, 33 - 45 (1985)

03G20
08A40

Dietlinde Lau

Klassen quasilinearer Funktionen von P_3

Unter quasilinearen Funktionen aus P_3 versteht man n-stellige Funktionen f^n (n=1,2,...) der dreiwertigen Logik, die entweder nur von einer Variablen wesentlich abhängen oder für die gewisse einstellige Funktionen f_0, f_1, ..., f_n aus P_3^1 mit

$$f(x_1,\ldots,x_n) = f_0(f_1(x_1) + \ldots + f_n(x_n) \bmod 2)$$

existieren. Bezeichnet wird die Menge all dieser Funktionen hier mit $\mathfrak{L}$. Gewisse Eigenschaften von $\mathfrak{L}$ wurden bereits in mehreren Arbeiten ermittelt. Drei Beispiele: In /1/ wies G. A. Burle nach, daß $\mathfrak{L}$ eine submaximale Klasse von P_3 ist.

I. A. Mal'cev zeigte in /7/, daß $\mathfrak{L}$ genau abzählbar unendlich viele Teilklassen besitzt. In /2/ gaben er und J. Demetrovics eine nicht weiter verfeinerbare endliche Kette von P_3 zu einer Teilklasse von $\mathfrak{L}$ an, die nicht endlich erzeugt ist.
In der vorliegenden Arbeit sollen nun auf der Grundlage der in /5/ ermittelten Unterhalbgruppen von $(P_3^1;*)$ sämtliche Teilklassen von $\mathfrak{L}$ charakterisiert werden. Dazu übernehmen wir aus /4/ die Bezeichnungen für die einstelligen Funktionen aus P_3 und aus /5/ die Beschreibungen der Unterhalbgruppen von $(P_3^1;*)$.

Nachfolgend nicht erläuterte Begriffe und Bezeichnungen entnehme man /8/, /3/ und /4/.

1. Einige Bezeichnungen

Wir vereinbaren, anstelle von + mod 2 nachfolgend nur + zu schreiben. L bezeichne die Menge der linearen Booleschen Funktionen, und es sei

$$\mathfrak{L}_{ab} := \bigcup_{n\geq 1}\left\{f^n \in \mathfrak{L} \mid f^n : E_3^n \to \{a,b\}\right\}, \quad \{a,b\} \subset E_3.$$

Mit pr_{ab} wird nachfolgend die durch

$pr_{ab}f^n = F^n :\Leftrightarrow \forall \tilde{x} \in \{a,b\}^n: g(f(x_1,..,x_n)) = F(g(x_1),..,g(x_n))$

definierte homomorphe Abbildung von $\mathcal{L}_{ab}$ auf L ($\subseteq P_2$) bezeichnet, wobei $g\binom{a}{b} = \binom{0}{1}$, $f^n \in \mathcal{L}_{ab}$ und $F \in L$ ist.

Mit Hilfe der bekannten Teilklassen von L aus /3/ lassen sich durch $pr_{ab}^{-1} A := \{f \in \mathcal{L}_{ab} | pr_{ab} f \in A\}$, $A = [A] \subseteq L$, Teilklassen von $\mathcal{L}_{ab}$ beschreiben.

Weiter sei Z_{ab} die Bezeichnung für die Menge $Pol_3\binom{b\ c\ b}{b\ c\ a}$, $\{a,b,c\} = E_3$.

Die Abbildungen $\varphi_i : f^n \to s_i^{-1}(f(s_i(x_1),\ldots,s_i(x_n)))$ sind für $i = 1,2,\ldots,6$ offenbar Automorphismen auf $\mathcal{L}$, mit deren Hilfe (wie schon in /5/) isomorphe Teilklassen von $\mathcal{L}$ beschreibbar sind. Offensichtlich ist jede Teilklasse A von $\mathcal{L}$ in der Form

$$A = (A \cap \mathcal{L}_{01}) \cup (A \cap \mathcal{L}_{02}) \cup (A \cap \mathcal{L}_{12}) \cup (A \cap [S])$$

darstellbar. Deshalb werden in Abschnitt 2 zunächst die Teilklassen von $\mathcal{L}_{01}$ und dann in Abschnitt 3 die nicht in $\mathcal{L}_{01}$ oder $\mathcal{L}_{02}$ enthaltenen Teilklassen von $\mathcal{L}_{01} \cup \mathcal{L}_{02}$ bestimmt, mit deren Hilfe abschließend in Abschnitt 4 leicht die restlichen Teilklassen von $\mathcal{L}$ charakterisiert werden können.

2. Teilklassen von $\mathcal{L}_{01}$

Aus der Definition von $\mathcal{L}_{01}$ und den Beziehungen $j_0 = 1 + j_1 + j_2$, $j_3 = 1 + j_2$, $j_4 = 1 + j_1$, $j_5 = j_1 + j_2$ folgt unmittelbar

Lemma 1: $\mathcal{L}_{01} = \bigcup_{n \geq 1} \Big\{ f^n \in P_3 \,\Big|\, \exists a_0,\ldots,a_1,\ b_1,\ldots,b_n \in E_2 :$

$$f(\tilde{x}) = a_0 + \sum_{i=1}^{n} (a_i j_1(x_i) + b_i j_2(x_i)) \Big\}. \square$$

Mit Hilfe dieses Lemmas ist leicht einzusehen, daß die Identi-

täten

$$\mathrm{pr}_{01}^{-1} A = \bigcup_{n\geq 1} \Big\{ f^n \in \mathscr{L}_{01} \mid f(\tilde{x}) = a_o + \sum_{i=1}^{n} (a_i j_1(x_i) + b_i j_2(x_i)) \wedge (\mathrm{pr}_{01} f)(\tilde{y}) = a_o + \sum_{i=1}^{n} a_i y_i \in A \Big\} \text{ und}$$

$$\mathrm{pr}_{01}^{-1} A \cap Z_{2c} = \bigcup_{n\geq 1} \Big\{ f^n \in \mathscr{L}_{01} \mid f(\tilde{x}) = a_o + \sum_{i=1}^{n} (a_i (j_1(x_i) + c j_2(x_i))) \wedge (\mathrm{pr}_{01} f)(\tilde{y}) = a_o + \sum_{i=1}^{n} a_i y_i \in A \Big\}$$

gelten, wobei c aus E_2 ist und A eine abgeschlossene Teilmenge von L bezeichnet. Ebenfalls Teilklassen von $\mathscr{L}_{01}$ sind die Mengen

$$B_{cr} := \bigcup_{n\geq 1} \Big\{ f^n \in \mathscr{L}_{01} \mid \exists a_1, \ldots, a_n \in E_2 : f(\tilde{x}) = c + \sum_{i=1}^{n} a_i j_2(x_i) \wedge (\text{höchstens r der } a_i \text{ sind gleich } 1) \Big\},$$

$r \in \{\infty, 1, 2, \ldots\}$ und
$B_r := B_{0r} \cup B_{1r}$, $r \in \{\infty, 1, 2, \ldots\}$.

Die Mengen $B_{c\infty}$ und B_∞ haben offensichtlich keine Basen, sind also insbesondere nicht endlich erzeugbar.

Lemma 2: Sei A eine Teilklasse von $\mathscr{L}_{01}$, die eine Funktion $i(x_1) + j_2(x_2)$, $i \in \{j_1, j_5\}$, enthält. Dann gilt $A = \mathrm{pr}_{01}^{-1}(\mathrm{pr}_{01} A)$ und $A = [A' \cup \{j_1(x_1) + j_2(x_2)\}]$ für jedes A' mit $A' \subseteq A$ und $[\mathrm{pr}_{01} A'] = \mathrm{pr}_{01} A$.

Beweis: Offensichtlich ist $A \subseteq \mathrm{pr}_{01}^{-1}(\mathrm{pr}_{01} A)$.

Sei $f^n \in \mathrm{pr}_{01}^{-1}(\mathrm{pr}_{01} A)$ und $f(\tilde{x}) = a_o + \sum_{i=1}^{n} (a_i j_1(x_i) + b_i j_2(x_i))$.
Es soll gezeigt werden, daß f zu A gehört. Wegen $\mathrm{pr}_{01} f \in \mathrm{pr}_{01} A$

gibt es in A eine Funktion f' mit

$f'(\tilde{x}) = a_o + \sum_{i=1}^{n} (a_i j_1(x_i) + b_i' j_2(x_i))$. Eine Superposition über $i(x_1) + j_2(x_2)$ ist sicher die Funktion

$$q(x,x_1,\ldots,x_n) = i(x) + \sum_{i=1}^{n} (b_i+b_i')j_2(x_i).$$

Folglich gilt $f(\tilde{x}) = q(f'(\tilde{x}), x_1,\ldots,x_n) \in A$ und somit $A = pr_{01}^{-1}(pr_{01}\, A)$.

Da $f' \in [A']$ für jedes A' mit $[pr_{01}\, A'] = pr_{01}\, A$ gilt, folgt aus dem oben Gezeigten auch die Beziehung $A = [A' \cup \{j_1(x_1) + j_2(x_2)\}]$. □

Lemma 3: Sei A eine Teilklasse von $\mathcal{L}_{01}$ und $pr_{01}A \notin [L^1]$. Dann ist A entweder die Menge $pr_{01}^{-1}(pr_{01}A)$, $Z_{20} \cap pr_{01}^{-1}(pr_{01}A)$ oder $Z_{21} \cap pr_{01}^{-1}(pr_{01}A)$.

Beweis: Für $f(\tilde{x}) = a_o + \sum_{i=1}^{n} (a_i j_1(x_i) + b_i j_2(x_i))$ bezeichne Ch(f) die Menge $\{(a_i,b_i) \mid i=1,2,\ldots,n\}$. Für A lassen sich folgende drei Fälle unterscheiden:

Fall 1: Es existiert ein $a \in E_2$, so daß für jedes $f \in A$ stets $Ch(f) \subseteq \{(1,a),(0,0)\}$ gilt.

In diesem Fall bewahrt $f \in A$ die Relation $\begin{pmatrix} 0 & 1 & a \\ 0 & 1 & 2 \end{pmatrix}$, und A ist offenbar gleich $pr_{01}^{-1}(pr_{01}\, A) \cap Z_{2a}$.

Fall 2: A enthält eine Funktion f mit $(0,1) \in Ch(f)$.
Es sei o.B.d.A. $(a_1,b_1) = (0,1)$. Dann gilt

$f'(x_1,x_2) := f(x_1,x_2,\ldots,x_2) = a_o + j_2(x_1) + r(x_2)$, wobei $r \in \{c_0,j_1,j_2,j_5\}$ ist. Wenn r zu $\{j_1,j_5\}$ gehört, haben wir $f'(x_1,f'(x_2,x_2)) = j_2(x_1) + i(x_2) \in A$, $i \in \{j_1,j_5\}$, womit Lemma 3 aus Lemma 2 folgt. Gilt $r \in \{c_0,j_2\}$, so erhalten wir

$f''(x) := f'(x,f'(x,x)) \in \{j_2,j_3\}$. Da pr_{01} A keine Teilmenge von $[L^1]$ ist, gehört nach /3/ oder /6/ zu A ein Urbild h der Funktion $H(\tilde{y}) := y_1 + y_2 + y_3$. Folglich ist

$h(x_1,f''(x_2), f''(f''(x_2))) = i(x_1) + j_2(x_2)$, $i \in \{j_1,j_5\}$, eine Funktion aus A. Also folgt auch für $r \in \{c_0,j_2\}$ Lemma 3 aus Lemma 2.

Fall 3: A enthält eine Funktion f mit $\{(1,0),(1,1)\} \subseteq Ch(f)$. Es sei o.B.d.A. $(a_1,b_1) = (1,0)$ und $(a_2,b_2) = (1,1)$. Dann erfüllt die Funktion $f(x_1,x_1,x_2,\ldots,x_2)$ die Bedingung von Fall 2. □

Satz 1: Sei A eine Teilklasse von $\mathcal{L}_{01}$, die keine Teilmenge von $[\mathcal{L}_{01}^1] \cup B_\infty$ ist. Dann gilt

$A \in \{pr_{01}^{-1}(pr_{01}\ A),\ Z_{20} \cap pr_{01}^{-1}(pr_{01}\ A),\ Z_{21} \cap pr_{01}^{-1}(pr_{01}\ A)\}$.

Beweis: Wenn A keine Teilmenge von $[\mathcal{L}_{01}^1] \cup B_\infty$ ist, so gibt es in A eine Funktion f mit nichtkonstanter Projektion $pr_{01}f$, die von mindestens zwei Variablen wesentlich abhängt.

Es sei o.B.d.A. $f(\tilde{x}) = a_o + \sum_{i=1}^{n} (a_i j_1(x_i) + b_i j_2(x_i)) \in A$, wobei $a_1 = 1$ und $(a_2,b_2) \neq (0,0)$ ist. Folgende zwei Fälle sind möglich:

Fall 1: $(a_2,b_2) = (0,1)$.

Durch Identifizieren der Variablen $x_3, \ldots, x_n$ erhält man aus f die Funktion $f'(x_1,x_2,x_3) = a_o + j_1(x_1) + b_1 j_2(x_1) + j_2(x_2) + p(x_3)$ mit $p \in \{c_0,j_1,j_2,j_5\}$. Dann gilt

$f'(f'(x_1,x_2,x_2),x_1,x_2) = i(x_1) + j_2(x_2)$, $i \in \{j_1,j_5\}$.

Nach Lemma 2 ist folglich $A = pr_{01}^{-1}(pr_{01}\ A)$.

Fall 2: $a_2 = 1$.

In diesem Fall ist $pr_{01}A \notin [L^1]$, und die Behauptung des Satzes folgt aus Lemma 3. □

Satz 2: Die Teilklassen $\neq \emptyset$ von $[\mathscr{L}_{01}^1] \cup B_\infty$ sind $[\{c_0\}]$, $[\{c_1\}]$, $[\{c_0,c_1\}]$, $[J_i]$, $[J_a] \cup B_r$, $[J_b] \cup B_{0r} \cup B_{1s}$, $[J_c] \cup B_{0r}$, $[J_d] \cup B_{1r}$, wobei $\{r,s\} \subset \{\infty,2,3,\ldots\}$, $1 \leq i \leq 41$, $a \in \{27,33,36,38,39,40,41\}$, $b \in \{27,33,36,40\}$, $c \in \{4,13,16,18,23,27,28,30,33,34,36,40\}$ und $d \in \{7,14,19,21,24,27,29,31,33,35,36,40\}$ gilt.

Beweis: Die Aussagen des Satzes erhält man leicht anhand der Eigenschaften der Funktionen aus $[\mathscr{L}_{01}^1] \cup B_\infty$ und der in /5/ angegebenen Unterhalbgruppen von $(P_3^1;\circ)$. □

Lemma 4: Sei A eine Teilklasse von $\mathscr{L}_{01}$ und $pr_{01}A \not\subseteq [L^1]$. Dann existiert in A eine Funktion $h(x_1,x_2,x_3) = i(x_1) + i(x_2) + i(x_3)$, $i \in \{j_1,j_5\}$, und es gilt $A = [A^1 \cup \{h\}]$.

Beweis: Nach /3/ oder /6/ haben wir $pr_{01}A = [(pr_{01}A)^1 \cup \{pr_{01}h\}]$. Ist also $A = Z_{2a} \cap pr_{01}^{-1}(pr_{01}A)$, $a \in E_2$, so gilt offensichtlich Lemma 4. Für $A = pr_{01}^{-1}(pr_{01}A)$ gehört $j_1(x_1) + j_2(x_2)$ zu A. Nach Lemma 2 und dem oben Bemerkten gilt $A = [A^1 \cup \{h, j_1(x_1)+j_2(x_2)\}]$, wobei h ein Urbild der Funktion $H(\tilde{y}) = y_1 + y_2 + y_3$ in A bezeichnet. Wegen $\{j_1,j_5\} \subseteq A$ können wir $h(\tilde{x}) = j_1(x_1) + j_1(x_2) + j_1(x_3)$ annehmen, und es gilt $h(x_1,x_2,j_5(x_2)) = j_1(x_1) + j_2(x_2) \in [A^1 \cup \{h\}]$, d. h., es ist $A = [A^1 \cup \{h\}]$ im Fall $A = pr_{01}^{-1}(pr_{01}A)$. Da es nach Satz 1 keine weiteren als die von uns betrachteten Möglichkeiten für A gibt, gilt Lemma 4. □

Satz 3: Für eine Teilklasse A von $\mathscr{L}_{01}$ ist

$$\operatorname{ord} A = \begin{cases} 2, \text{ falls } A \not\subseteq [\mathscr{L}_{01}^1] \cup B_\infty \wedge (C \cap A^1 \neq \emptyset \vee pr_{01}A \subseteq [L^1]), \\ 3, \text{ falls } pr_{01}A \not\subseteq [L^1] \wedge C \cap A^1 = \emptyset, \\ t, \text{ falls } A \subseteq [\mathscr{L}_{01}^1] \cup B_\infty \wedge A \cap (B_t \setminus [A^1]) \neq \emptyset \\ \qquad \wedge (\forall r \geq t: A \cap (B_r \setminus [A^1]) = \emptyset) \text{ gilt.} \end{cases}$$

(ord A bezeichnet die Ordnung der Klasse A.)

Beweis: Folgende Fälle sind nach Satz 1 für eine Teilklasse A von $\mathscr{L}_{01}$ möglich:

Fall 1: $A = Z_{2a} \cap pr_{01}^{-1}(pr_{01}A)$, $a \in E_2$.

In diesem Fall ist ord A = ord $pr_{01}A$. Die Behauptungen des Satzes kann man also der Arbeit /3/ entnehmen.

Fall 2: $A = pr_{01}^{-1}(pr_{01}A)$ und $A \notin [\mathscr{L}_{01}^1] \cup B_\infty$.

Nach Lemma 4 und 2 gilt in diesem Fall

$$\text{ord } pr_{01}A \leq \text{ord } A \begin{cases} \leq 3 \text{ für } pr_{01}A \notin [L^1], \\ = 2 \text{ für } pr_{01}A \in [L^1]. \end{cases}$$

Bekanntlich haben nur die Teilklassen von L, die keine Konstanten enthalten und nicht Teilmengen von $[L^1]$ sind, die Ordnung 3. Gehört eine Konstante zu $pr_{01}A$, so ist die Funktion h aus Lemma 4 offensichtlich eine Superposition über zweistellige Funktionen aus A. Also gelten im Fall 2 die Aussagen des Satzes.

Fall 3: $A \subseteq [\mathscr{L}_{01}^1] \cup B_\infty$.

Die Ordnung von A ergibt sich in diesem Fall aus Satz 2 und der Beziehung $[\mathscr{L}_{01}^1] \cup B_r \subsetneq [\mathscr{L}_{01}^1] \cup B_{r+1}$, $2 \leq r \leq \infty$. □

Als unmittelbare Folgerung aus Satz 3 erhält man den

Satz 4: Die einzigen nicht endlich erzeugbaren Teilklassen von $\mathscr{L}_{01}$ sind

$[J_a] \cup B_\infty$, $[J_b] \cup B_{0\infty}$, $[J_c] \cup B_{1\infty}$, $[J_d] \cup B_{0\infty} \cup B_{1s}$ und

$[J_d] \cup B_{1\infty} \cup B_{0s}$, wobei

$a \in \{27,33,36,38,39,40,41\}$, $b \in \{4,13,16,18,23,27,28,30,33,34,36,40\}$, $c \in \{7,14,19,21,24,27,29,31,33,35,36,40\}$, $d \in \{27,33,36,40\}$ und $2 \leq r,s \leq \infty$ gilt. □

3. Teilklassen von $\mathscr{L}_{01} \cup \mathscr{L}_{02}$, die keine Teilklassen von $\mathscr{L}_{01}$ oder $\mathscr{L}_{02}$ sind

Ziel dieses Abschnittes ist ein notwendiges und hinreichendes Kriterium, mit dem man die Abgeschlossenheit einer Menge $A_1 \cup A_2$ ($A_1 \subseteq \mathscr{L}_{01}$, $A_2 \subseteq \mathscr{L}_{02}$) entscheiden kann.
Bezeichne $A * A'$ die Menge $\{f * g \mid f \in A \wedge g \in A'\}$. Offensichtlich gilt dann folgendes

Lemma 5: Seien A_1 und A_2 Teilklassen von $\mathscr{L}$ mit $A_1 \subseteq \mathscr{L}_{01}$ und und $A_2 \subseteq \mathscr{L}_{02}$. Dann ist $A_1 \cup A_2$ genau dann abgeschlossen, wenn $A_1 * A_2 \subseteq A_1$ und $A_2 * A_1 \subseteq A_2$ ist. □

Lemma 6: Seien A_1 und A_2 Teilklassen von $\mathscr{L}$, $A_1 \subseteq \mathscr{L}_{01}$, $A_2 \subseteq \mathscr{L}_{02}$, $(A_1 \cup A_2)^1$ eine Unterhalbgruppe von $(P_3^1; *)$ und $\{i,j\} = \{1,2\}$.
Dann gilt:

(a) $A_1^1 \subseteq \{c_0, c_1, j_1, j_4\} \Longrightarrow A_1 * A_2 \subseteq A_1$,

(b) $A_2^1 \subseteq \{c_0, c_2, u_2, u_3\} \Longrightarrow A_2 * A_1 \subseteq A_2$,

(c) $\mathrm{pr}_{0i} A_i \notin [L^1] \Longrightarrow A_1 * A_j \subseteq A_i$,

(d) $\mathrm{pr}_{01} A_1 \subseteq [L^1] \wedge j_1(x_1) + j_2(x_2) \in A_1 \wedge A_2^1 \subseteq \{c_0, c_2, u_2, u_3\} \Longrightarrow A_1 * A_2 \subseteq A_1$.

Beweis: (a) Nach Abschnitt 2 und den Voraussetzungen über A_1 ist A_1 eine Teilklasse von $Z_{20} \cap \mathscr{L}_{01} = [\{c_1, j_1(x_1) + j_1(x_2)\}]$. Wegen $j_1\binom{0}{2}$ gilt folglich $A_1 * A_2 \subseteq A_1$.

(b) A_2 ist laut Voraussetzung und Abschnitt 2 eine Teilmenge von $\varphi_2(Z_{20} \cap \mathscr{L}_{01})$. Wegen $u_2\binom{0}{1} = \binom{0}{0}$ folgt hieraus $A_2 * A_1 \subseteq A_2$.

(c) Es sei o.B.d.A. $\mathrm{pr}_{01} A_1 \subseteq [L^1]$. Dann ist wegen Lemma 4 $A_1 = [A_1^1 \cup \{h\}]$, wobei h ein beliebig wählbares Urbild der Funk-

tion $H(\tilde{y}) = y_1 + y_2 + y_3$ aus A_1 bezeichnet. Angenommen, es gilt $A_1 \ast A_2 \notin A_1$. Dann ist $A_1^* := [A_1 \cup (A_1 \ast A_2)]$ eine abgeschlossene Menge mit $A_1^* \supsetneq A_1$ und $A_1^* = [(A_1^*)^1 \cup \{h\}]$. Folglich gilt $(A_1^*)^1 \supsetneq A_1^1$, d. h., es gibt eine Funktion $g \in A_1$ und gewisse Funktionen $p_1, \ldots, p_n$ aus $(A_1 \cup A_2)^1$ mit $g'(x) := g(p_1(x), p_2(x), \ldots, p_n(x)) \notin A_1^1$. Für die Funktion g gilt jedoch auch $g(\tilde{x}) = g_o(g_1(x_1), g_2(x_2), \ldots, g_n(x_n))$ für bestimmte Funktionen $g_o \in [\{h\}]$ und $g_1, \ldots, g_n \in A_1^1$. Damit erhalten wir $g' = g_o(g_1 \ast p_1, g_2 \ast p_2, \ldots, g_n \ast p_n)$. Laut Voraussetzung ist aber $g_i \ast p_i \in (A_1 \cup A_2)^1$, $i = 1, 2, \ldots, n$, sowie A_1 abgeschlossen. Also gehört g' zu A_1 im Widerspruch zur Annahme.

(d) Sei $pr_{01}A_1 \subseteq [L^1]$, $j_1(x_1) + j_2(x_2) \in A_1$ und $A_2^1 \subseteq \{c_0, c_2, u_2, u_3\}$. Nach Abschnitt 2 ist dann $A_1 = [A_1^1 \cup \{j_1(x_1) + j_2(x_2)\}]$ und $A_2 \subseteq \varphi_2(Z_{20} \cap \mathcal{L}_{01}) = [\{c_2, u_2(x_1) \dot{+} u_2(x_2)\}]$ mit

x	y	$x \dot{+} y$
0	0	0
0	2	2
2	0	2
2	2	0 .

Da $j_1\binom{0}{2} = \binom{0}{0}$, $j_2(a \dot{+} u_2(x_1) \dot{+} u_2(x_2) \dot{+} \ldots \dot{+} u_m(x_m))$ $= j_2(a) + j_2(x_1) + \ldots + j_2(x_m)$ gilt und $(A_1 \cup A_2)^1$ bez. $\ast$ abgeschlossen ist, folgt $A_1 \ast A_2 \subseteq A_1$. □

<u>Satz 5:</u> Seien A_1 und A_2 Teilklassen von $\mathcal{L}$, $A_1 \subseteq \mathcal{L}_{01}$, $A_2 \subseteq \mathcal{L}_{02}$, $A_1 \notin [\mathcal{L}_{01}^1] \cup B_\infty$, $A_2 \notin \varphi_2([\mathcal{L}_{01}^1] \cup B_\infty)$ und $A_1^1 \subseteq \{c_0, c_1, j_1, j_4\}$ oder $A_2^1 \subseteq \{c_0, c_1, u_2, u_3\}$.

Dann ist $A_1 \cup A_2$ genau dann abgeschlossen, wenn $(A_1 \cup A_2)^1$ eine Unterhalbgruppe von $(P_3^1; \ast)$ ist.

Beweis: Ist $A_1 \cup A_2$ abgeschlossen, so muß $(A_1 \cup A_2)^1$ offensichtlich eine Unterhalbgruppe von $(P_3^1; \ast)$ sein.
Sei jetzt $(A_1 \cup A_2)^1$ eine bez. $\ast$ abgeschlossene Menge. Folgende drei Fälle sind zu untersuchen:

Fall 1: $(A_1 \cup A_2)^1 \subseteq \{c_0, c_1, c_2, j_1, j_4, u_2, u_3\}$.
Nach Lemma 6(a,b) und Lemma 5 ist in diesem Fall $A_1 \cup A_2$ abgeschlossen.

Fall 2: $A_1^1 \not\subseteq \{c_0, c_1, j_1, j_4\}$ und $A_2^1 \subseteq \{c_0, c_2, u_2, u_3\}$.
Nach Lemma 6(b) gilt $A_2 \ast A_1 \subseteq A_2$. Da A_1 keine Teilmenge von $[\mathfrak{L}_{01}^1] \cup B_\infty$ ist, haben wir $\mathrm{pr}_{01} A_1 \not\subseteq [L^1]$ oder $\mathrm{pr}_{01} A_1 \subseteq [L^1]$ und $j_1(x_1) + j_2(x_2) \in A_1$. Mit Hilfe von Lemma 6(c,d) folgt hieraus $A_1 \ast A_2 \subseteq A_1$. Wegen Lemma 5 ist also $A_1 \cup A_2$ abgeschlossen.

Fall 3: $A_1^1 \subseteq \{c_0, c_1, j_1, j_4\}$ und $A_2^1 \not\subseteq \{c_0, c_2, u_2, u_3\}$.
Wegen $\varphi_2(\{c_0, c_1, j_1, j_4\}) = \{c_0, c_2, u_2, u_3\}$ sind die den Bedingungen von Fall 3 genügenden Klassen A_1 und A_2 isomorph zu solchen, die die Bedingungen von Fall 2 erfüllen. Also gilt Satz 5 auch im Fall 3. □

Satz 6: Seien A_1 und A_2 Teilklassen von $\mathfrak{L}$, $A_1 \subseteq \mathfrak{L}_{01}$, $A_2 \subseteq \mathfrak{L}_{02}$, $A_1 \cup A_2 \not\subseteq [\mathfrak{L}^1]$, $A_1^1 \not\subseteq \{c_0, c_1, j_1, j_4\}$ und $A_2^1 \not\subseteq \{c_0, c_2, u_2, u_3\}$. Dann ist $A_1 \cup A_2$ genau dann abgeschlossen, wenn $\mathrm{pr}_{01} A_1 \not\subseteq [L^1]$, $\mathrm{pr}_{02} A_2 \not\subseteq [L^1]$ gilt und $(A_1 \cup A_2)^1$ bez. $\ast$ abgeschlossen ist.

Beweis: Ist $pr_{01}A_i \not\subseteq [L^1]$ für $i = 1, 2$ und $(A_1 \cup A_2)^1$ eine Unterhalbgruppe von $(P_3^1; \equiv)$, so folgt aus Lemma 6(c) und 5, daß $A_1 \cup A_2$ abgeschlossen ist.

Sei $A_1 \cup A_2$ eine Teilklasse von $\mathfrak{L}$.

Da laut Voraussetzung $A_1^1 \not\subseteq \{c_0, c_1, j_1, j_4\}$ und $A_2^1 \not\subseteq \{c_0, c_2, u_2, u_3\}$ gilt, gibt es in A_1^1 eine Funktion p mit $p\binom{0}{2} \in \binom{0\ 1}{1\ 0}$ und in A_2^1 eine Funktion q mit $q\binom{0}{1} \in \binom{0\ 2}{2\ 0}$. Weil $A_1 \cup A_2$ abgeschlossen ist, gilt $p \equiv A_2 \subseteq A_1$ und $q \equiv A_1 \subseteq A_2$. Offensichtlich ergibt sich hieraus und aus der Voraussetzung $A_1 \cup A_2 \not\subseteq [\mathfrak{L}^1]$, daß $A_1 \not\subseteq [\mathfrak{L}^1]$ und $A_2 \not\subseteq [\mathfrak{L}^1]$ ist. Angenommen, es gelte $pr_{01}A_1 \subseteq [L^1]$. Dann gehört zu A_1 nach Abschnitt 2 eine Funktion $t(x_1, x_2) = g(x_1) + j_2(x_2)$, $g \in A_1^1$. Also gehört auch die Funktion

$$t'(x_1, x_2) := \begin{cases} t(x_1, q(x_2)) & \text{für } g \in \{j_0, j_1, j_4, j_5\}, \\ t(q(x_1), q(x_2)) & \text{für } g \in \{j_2, j_3\} \end{cases}$$

zu A_1, wobei aber offensichtlich $pr_{01}t' \not\subseteq [L^1]$ ist im Widerspruch zur Annahme. Folglich gilt $pr_{01}A_1 \not\subseteq [L^1]$. Analog zeigt man $pr_{02}A_2 \not\subseteq [L^1]$. □

Zur vollständigen Charakterisierung sämtlicher Teilklassen von $\mathfrak{L}_{01} \cup \mathfrak{L}_{02}$, die nicht in $\mathfrak{L}_{01}$, $\mathfrak{L}_{02}$ oder $[\mathfrak{L}^1]$ enthalten sind, fehlen uns nach Satz 5 und 6 nur noch die Teilklassen von $[\mathfrak{L}_{01}^1] \cup B_\infty \cup \varphi_2(Z_{20} \cap \mathfrak{L}_{01})$ und die von $\varphi_2([\mathfrak{L}_{01}^1] \cup B_\infty) \cup (Z_{20} \cap \mathfrak{L}_{01})$. Da diese noch zu untersuchenden Mengen untereinander isomorph sind, genügt es, nur die Teilklassen einer der genannten Mengen zu bestimmen. Dies geschieht in Satz 7.

Satz 7: Die nicht in $[\mathfrak{L}^1]$, $\mathfrak{L}_{01}$ oder $\mathfrak{L}_{02}$ enthaltenen Teilklassen von $[\mathfrak{L}_{01}] \cup B_\infty \cup \varphi_2(Z_{20} \cap \mathfrak{L}_{01})$ sind $A \cup A'$, wobei

$A \in \{[\{c_0,c_1\}], [J_a] \cup B_\infty, [J_b] \mid a \in \{27,33,36,38,39,40,41\}, b \in \{3,12,25\}\}$ und $A' \in \{\varphi_2(Z_{20} \cap \mathcal{L}_{01}), \varphi_2([\{j_1(x_1)+j_1(x_2)\}])\}$ gilt,
$[J_c] \cup B_\infty \cup \varphi_2([\{j_1(x_1)+j_1(x_2)\}])$, $c \in \{4,13,16,18,23,28,30,34\}$,
$[J_d] \cup B_{1\infty} \cup \varphi_2([\{j_1(x_1)+j_1(x_2)\}])$, $d \in \{14,24\}$, sowie

$[A_1] \cup B_r$,
$[J_e] \cup B_{0r} \cup [\{c_0,u_2\}]$,
$[J_f] \cup B_{1r} \cup [\{c_0,u_2\}]$,
$[J_g] \cup B_{\alpha r} \cup B_{\beta,r-1} \cup [A_2]$,
$[J_g] \cup B_{0r} \cup B_{1s} \cup [\{c_0,u_2\}]$, wobei
$e \in \{4,13,16,18,23,27,28,30,33,34,36,40\}$,
$f \in \{7,14,19,21,24,27,29,31,33,35,36,40\}$,
$g \in \{27,33,36,40\}$, $2 \leq r,s \leq \infty$, $\{\alpha,\beta\} = E_2$ gilt, A_1 eine Unterhalbgruppe von $J \cup U_{25}$, $(A_1 \setminus \{c_0\}) \cap U_{25} \neq \emptyset$, die $\{j_2,j_3\}$ enthält sowie $A_2 \in \{[\{c_2\}], [\{c_0,c_2,u_2\}]\}$ ist.

Beweis: Der Satz folgt aus /5/ und 2. □

4. Die restlichen Teilklassen von $\mathcal{L}$

Offensichtlich folgt aus Lemma 5 das folgende

Lemma 7: Seien A_1, A_2 und A_3 Teilklassen von $\mathcal{L}$ mit $A_1 \subseteq \mathcal{L}_{01}$, $A_2 \subseteq \mathcal{L}_{02}$ und $A_3 \subseteq \mathcal{L}_{12}$. Dann ist $A_1 \cup A_2 \cup A_3$ genau dann abgeschlossen, wenn $A_1 \cup A_2$, $A_1 \cup A_3$ und $A_2 \cup A_3$ abgeschlossene Mengen sind und $(A_1 \cup A_2 \cup A_3)^1$ eine Halbgruppe ist. □

Da die Mengen $\mathcal{L}_{01} \cup \mathcal{L}_{02}$, $\mathcal{L}_{01} \cup \mathcal{L}_{12}$, $\mathcal{L}_{02} \cup \mathcal{L}_{12}$ untereinander isomorph sind, hat man durch Lemma 7 und die Abschnitte 2 und 3 eine vollständige Beschreibung der Teilklassen von $\mathcal{L}_{01} \cup \mathcal{L}_{02} \cup \mathcal{L}_{12}$. Über die jetzt noch fehlenden Teilklassen von $\mathcal{L}$ sagt der folgende Satz etwas aus.

Satz 8: Die Teilklassen von $\mathcal{L}$, die keine Teilmengen von $[\mathcal{L}^1]$ sind und die mindestens eine Permutation enthalten, sind
$A_1 \cup [\{s_1\}]$ (A_1 Teilklasse von $\mathcal{L}_{01} \cup \mathcal{L}_{02} \cup \mathcal{L}_{12}$),
$A_2 \cup [\{s_1, s_i\}]$ ($i \in \{2,3,6\}$, A_2 Teilklasse von $\mathcal{L}_{01} \cup \mathcal{L}_{02} \cup \mathcal{L}_{12}$ mit $s_i \notin A_2 \subseteq A_2$, $A_2^1 \cup \{s_1, s_i\}$ Unterhalbgruppe von $(P_3^1; \circ)$),
$\mathcal{L}_{01} \cup \mathcal{L}_{02} \cup \mathcal{L}_{12} \cup [\{s_1, s_4, s_5\}]$ und $\mathcal{L}$.

Beweis: Der Satz folgt aus /5/ und den Eigenschaften der Funktionen aus $\mathcal{L}$. □

Literatur

/1/ Бурле, Г. А.: Классы к-значной логики, содержащие все функции одной переменной. Дискретный анализ 10, 3-7 (1967)

/2/ Demetrovics, J., und Malcev, I. A.: On the depth of infinitely generated subalgebras of Post's iterative algebra P_3. Eingereicht bei Colloquium on Universal Algebra, Szeged (1983)

/3/ Jablonski, S. W., Gawrilow, G. P., und Kudrjawzew, W. B.: Boolesche Funktionen und Postsche Klassen. Berlin 1970

/4/ Lau, D.: Submaximale Klassen von P_3. Elektron. Informationsverarb. Kybernetik 18, 227 - 243 (1982)

/5/ Lau, D.: Unterhalbgruppen von $(P_3^1, \circ)$. Rostock. Math. Kolloq. 26, 55 - 62 (1984)

/6/ Lau, D.: Über abgeschlossene Mengen linearer Funktionen von P_k, Teil I, Preprint Wilhelm-Pieck-Universität Rostock (1983)

/7/ Мальцев, И. А.: Некоторые свойства клеток алгебр Поста. Дискретный анализ 23, 24 - 31 (1973)

/8/ Pöschel, R., und Kalužnin, L. A.: Funktionen- und Relationenalgebren. Berlin 1979

eingegangen: 23. 10. 1984

Anschrift des Verfassers:

Dr. D. Lau
Wilhelm-Pieck-Universität Rostock
Sektion Mathematik
Universitätsplatz 1
DDR-2500 Rostock

Rostock. Math. Kolloq. 28, 46 - 48 (1985)

05A99
06A10

Рафаил Азриелевич Розенфельд

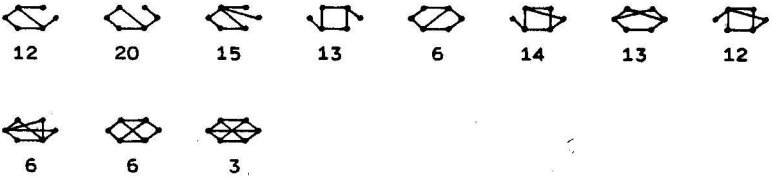
О числе частичных упорядочений на 6-множестве

Назовём множество из n элементов n-множеством ($n \in N$). Будем различать число $H^*(n)$ неизоморфных частичных порядков на n-множестве (или число неизоморфных диаграмм Хассе на n-множестве) и число $H(n)$ всех частичных порядков на n-множестве, полученных из $H^*(n)$ переименованием вершин диаграмм Хассе. Если $Z(n)$ ($Z^*(n)$) число соответствующих связных диаграмм Хассе, а $K(n)$ ($K^*(n)$) - число соответствующих несвязных диаграмм, то, очевидно

$$H(n) = Z(n) + K(n), \quad H^*(n) = Z^*(n) + K^*(n).$$

Для $1 \leq n \leq 6$ $H(n)$ и $H^*(n)$ известны /1/. Известно также $H(7) = 6129859$ (ibid), без обоснований. Обоснования не были опубликованы. Поэтому появление работы /2/, где детально подсчитано $Z(n)$ для $1 \leq n \leq 6$ было целесообразным. Как указано в /2/, определение $K(n)$ по известным $Z(i)$, $i < n$, не представляет трудностей. Однако в /2/ были допущены ошибки, что привело к неверному значению $Z(6)$, а потому и $H(6)$. Рассмотрим поэтому еще раз один возможный подход к определению $Z(6)$. Диаграммы Хассе не содержат треугольников. Но по известной теореме Турана /3/ наибольшее число рёбер в графе с n вершинами без треугольников будет $[n^2/4]$, что позволяет для 6-множества не рассматривать графы с числом рёбер больше 9. Но графов с числом рёбер не больше 9 на 6-множестве имеется всего 38 /3/, из них связных 19. Ниже приведены эти графы. Под каждым графом указано число неизоморфных диаграмм Хассе, соответствующих данному графу. Сумма всех этих чисел даёт $Z^*(6) = 238$.

16 16 20 6 9 24 7 20



Дальше надо рассмотреть все $Z^*(6)$ диаграмм и под каждой указать число всех частичных порядков, которые соответствуют данной диаграмме. Например, для графа  получаем такие диаграммы:

	$\frac{6!}{2}$	$6!$	$6!$	$\frac{6!}{2}$	$\frac{6!}{2}$	$\frac{6!}{2}$	$\frac{6!}{3}$
или	360	720	720	360	360	360	120

Это и было фактически сделано в работе /2/ , но для диаграммы


360

(см. /2/ , стр. 59 , последняя строка, четвёртая картинка) было ошибочно указано число 180, а такие диаграммы


720 720 720 360 360 180 180 360 360

были опущены вовсе. С учётом исправлений число $Z(6)$, данное в /2/, увеличится на $(360-180)+720\cdot3+360\cdot4+180\cdot2$ и будет равно 101642, а $H(6) = 130023$, что соответствует /1/. Изменится соответственно и число замкнутых максимальных классов 8-значной логики, указанное в /2/. Итоговая таблица для $1 \le n \le 6$ будет такая

n	графы без △		Z*(n)	K*(n)	H*(n)	Z(n)	K(n)	H(n)
	всего	связные						
1	1	0	1	0	1	1	0	1
2	2	1	1	1	2	2	1	3
3	3	1	3	2	5	12	7	19
4	7	3	10	6	16	146	73	219
5	14	6	44	19	63	3060	1171	4231
6	38	19	238	80	318	101642	28381	130023

Литература

/1/ Birkhoff, G.: Lattice theory. Third edition. Providence 1967

/2/ Radtke, S.: Die Anzahl aller möglichen Halbordnungsrelationen auf einer maximal sechselementigen Menge. Rostock. Math. Kolloq. 23, 55 - 61 (1983)

/3/ Harary, F.: Graph theory. Reading 1969

поступило: 4 декабря 1984

Адрес автора:

Р. А. Розенфельд
д. 20, кв. 5
454015 Челябинск 15
СССР

Rostock. Math. Kolloq. 28, 49 - 54 (1985)

05B05
62K10

Hans-Dietrich O. F. Gronau
Irina Rentner

On the decomposition of the set of all k-element subsets of a v-element set into indecomposable t-(v,k,λ) designs

1. Introduction

D. Mesner (A. Rosa /11/) asked the following question:
In which ways the set of all k-element subsets of a v-element set can be decomposed into indecomposable t-(v,k,λ) designs?
This is a generalization of older problems:
What is the maximum number of pairwise disjoint Steiner triple systems? For which v the set of all triples of a v-element set can be decomposed into Steiner triple systems? Investigations in this direction go back to the beginning of the study of designs. In 1973 Teirlinck /13/ and others conjectured that for $v \geq 9$ ($v \equiv 1, 3 \bmod 6$) the set of all triples can be partitioned into Steiner triples systems. L. Teirlinck /13/ was the first to establish that the conjecture is true for infinitely many values of v. Denniston, Schreiber and Wilson gave other constructions of such decompositions. For an excellent survey on results on this topic we refer to A. Rosa /10/. A further progress in proving the conjecture is due to Lu Jia-Xi /8/. He established that the conjecture is true for $v = pn + 2$, where p is a prime number and n is a positive integer such that $p \equiv 7 \bmod 8$ or $p \in \{5,17,19,29\}$ and $(p,n) \neq (5,1)$. Mesner's problem is more general since he askes for all decompositions into indecomposable designs, i. e. designs which cannot be decomposed further. In general a solution is hopeless since only in few cases a complete survey on indecomposable designs of given parameters is known, see Gronau /1/.
In this paper we survey results on such decompositions. Our own most important results are that on 2-(8,4,λ) designs. Also we give proofs for the 2-(7,3,λ) designs, but these are easily obtained and probably known by others.

2. Definitions and results

First we describe the problem more precisely. A t-(v,k,λ) design is a system of k-element subsets (called blocks) of a v-element set K such that every t-element subset of K occurs exactly λ times in the blocks. A t-(v,k,λ) design B is called indecomposable (or elementary) if and only if there is no subsystem B' of B which is a t-(v,k,λ') design with $0 < \lambda' < \lambda$. Let $B^*(v,k)$ denote the set of all k-element subsets of K. It is wellknown that the existence of a t-(v,k,λ) design implies that

$$\lambda \binom{v-i}{t-i} \Big/ \binom{k-i}{t-i} \qquad (1)$$

is an integer for every $i = 0,1,\ldots,t-1$.
Let λ_0 denote the smallest possible λ according to (1) for fixed v,k, and t. Then for every t-(v,k,λ) design we have $\lambda \equiv 0 \bmod \lambda_0$. Obviously, Steiner systems ($\lambda = 1$) and more generally t-(v,k,λ_0) designs are indecomposable. Note that $B^*(v,k)$ is a t-$(v,k,\binom{v-t}{k-t})$ design. Every decomposition of $B^*(v,k)$ into t-(v,k,λ) designs S_i can be characterized by the distribution of the λ's. More precisely, we associate to a decomposition

$$B^*(v,k) = S_1 \cup S_2 \cup \ldots \cup S_r,$$

where S_i is an indecomposable t-(v,k,λ_i) design $(i=1,2,\ldots,r)$, the partition $\lambda_1 + \lambda_2 + \ldots + \lambda_r$ of $\binom{v-t}{k-t}$. Note $\lambda_i \equiv 0 \bmod \lambda_0$. In the following table we list all partitions

$$\lambda_1 + \lambda_2 + \ldots + \lambda_r$$

of $\binom{v-t}{k-t}$ with $\lambda_1 \geq 1$ and $\lambda_i \equiv 0 \bmod \lambda_0$ $(i=1,2,\ldots,r)$ for small parameters v,k,t according to $2 \leq t < k \leq \frac{v}{2}$ (other cases can be reduced to these ones) and survey the results on the existence of corresponding decompositions of $B^*(v,k)$ into indecomposable t-(v,k,λ_i) designs.

Remark: The number of nonisoporphic 2-(9,3,3) designs is not clear. Harnau /6/ has 329 and Ivanov /7/ (by Computer) has

v	k	t	λ_0	$\binom{v-t}{k-t}$	Partitions	Existence	Reference
6	3	2	2	4	4	No	
					2+2	Yes	
7	3	2	1	5	5	No	Section 3
					4+1	No	
					3+2	No	
					3+1+1	Yes	
					2+2+1	No	
					2+1+1+1	No	
					1+1+1+1+1	No	
8	3	2	6	6	6	Yes	
8	4	2	3	15	15	No	Section 4
					12+3	No	
					9+6	No	
					9+3+3	Yes	
					6+6+3	Yes	
					6+3+3+3	Yes	
					3+3+3+3+3	Yes	
8	4	3	1	5	5	No	Consequence of the results on 2-(7,3,λ) designs (unique extensions of the 2-(7,3,λ) designs), see e. g. Gronau /3/
					4+1	No	
					3+2	No	
					3+1+1	Yes	
					2+2+1	No	
					2+1+1+1	No	
					1+1+1+1+1	No	
9	3	2	1	7	7	No	Harnau /6/
					6+1	No	
					5+2	No	
					5+1+1	No	
					4+3	No (?)	
					4+2+1	No (?)	
					4+1+1+1	No (?)	
					3+3+1	Yes	Rosa /12/
					3+2+2	Yes	
					3+2+1+1	Yes	
					3+1+1+1+1	Yes	
					2+2+2+1	Yes	
					2+2+1+1+1	Yes	
					2+1+1+1+1+1	Yes	
					1+1+1+1+1+1+1	Yes	
10	3	2	2	8	8	No	
					6+2	?	
					4+4	Yes	Rosa /12/
					4+2+2	Yes	
					2+2+2+2	Yes	

Table

330 designs. Harnau proved that the corresponding 329 2-(9,3,4) designs - complements of the 2-(9,3,3) designs with respect to $B^*(9,3)$ - are all indecomposable. Is the one missing design decomposable? The proof of nonexistence of decompositions into indecomposable designs containing 2-(9,3,4) designs is therefore not complete; which is marked in the table by (?).

3. Decompositions with $v = 7$, $k = 3$, $t = 2$

In Rasch /9/ and Gronau /2/ it is shown that the only indecomposable 2-$(7,3,\lambda)$ designs without repeated blocks have $\lambda = 1$ or 3. In both cases there is exactly one of such designs (up to isomorphism). Hence, any decomposition must have at least two disjoint 2-(7,3,1) designs, which form obviously a 2-(7,3,2) design. Furthermore, in /2/ it is proved that any two 2-(7,3,2) designs without repeated blocks are isomorphic and its complement with respect to $B^*(7,3)$ is isomorphic to the mentioned 2-(7,3,3) design. Here is the only decomposition:

$$(34)(576)T_1 \cup (36)(47)T_1 \cup T_2,$$

where

$$T_1 = \begin{pmatrix} 1112233 \\ 2464545 \\ 3576776 \end{pmatrix} \quad \text{and} \quad T_2 = \begin{pmatrix} 11111111122222233345 \\ 22233444633445544656 \\ 35756567746676757767 \end{pmatrix}.$$

4. Decompositions with $v = 8$, $k = 4$, $t = 2$

We know by Gronau /4/ that there is no indecomposable 2-(8,4,12) and 2-(8,4,15) design. In Gronau, Reimer /5/ all indecomposable 2-$(8,4,\lambda)$ designs with $\lambda = 6$ and 9 are determined. It turned out that there is only one indecomposable 2-(8,4,9) design and its complement is decomposable, i. e. $9+6$ has no decomposition. Surprisingly, the only indecomposable 2-(8,4,9) design is the only 3-(8,4,3) design. For the remaining partitions we give examples of decompositions:

$9+3+3$: $(4567)B_1 \cup (476)B_1 \cup C$,

$6+6+3$: $(354)(687)B_4 \cup (12)(36)(587)C_3 \cup C_1$,

$6+3+3+3$: $(587)B_4 \cup (12645387)B_4 \cup (17258364)B_3 \cup C_{16}$,

$3+3+3+3+3$: $(345687)B_4 \cup (365847)B_4 \cup (127)(36)B_4$
$\cup (15723468)B_4 \cup (37)(58)B_4$.

Here the B's and C's denote designs from /4/ and /5/, respectively, namely

$$B_1 = \begin{pmatrix} 11111112222335 \\ 22233443344446 \\ 35756565656577 \\ 46878878778688 \end{pmatrix}, \qquad B_3 = \begin{pmatrix} 11111112222333 \\ 22234453445445 \\ 33675676576566 \\ 45886787788887 \end{pmatrix},$$

$$B_4 = \begin{pmatrix} 11111112222333 \\ 22234453445445 \\ 33665777566566 \\ 45786888788877 \end{pmatrix},$$

$$C_1 = \begin{pmatrix} 11111111111111122222222333445 \\ 22222233333444333333455444566 \\ 34456755567556445665675677677 \\ 46767867888788587778888678788 \end{pmatrix},$$

$$C_3 = \begin{pmatrix} 11111111111111122222222333445 \\ 22222233333445333334444556556 \\ 34556744456676445575665676777 \\ 65678857887887677888788688788 \end{pmatrix},$$

$$C_{16} = \begin{pmatrix} 11111111111111122222222333444 \\ 22222233333445333334554565556 \\ 34466744557566444555676676777 \\ 65878857788678678687788788888 \end{pmatrix},$$

$$C = \begin{pmatrix} 11111111111111111111112222222222223333334445 \\ 22222222233333344445633333344445644445555 6 \\ 33344555744556655676744556755666756676767677 \\ 46867678857787868887858677878788678888788 \end{pmatrix}.$$

The indecomposability of these designs is proved in /4/ and /5/.

References

/1/ Gronau, H.-D. O. F.: A survey of results on the number of t-(v,k,λ) designs. Ann. Discrete Math., to appear

/2/ Gronau, H.-D. O. F.: Einige Bemerkungen zur Arbeit "Über die Anzahl elementarer BUB in eingeschränkten (v,k)-Familien" von D. Rasch und G. Herrendörfer, Rostock. Math. Kolloq. 9, 27 - 34 (1978)

/3/ Gronau, H.-D. O. F.: Über (2p-1)-(4p,2p,λ)-Blockpläne. Rostock. Math. Kolloq. 11, 67 - 74 (1979)

/4/ Gronau, H.-D. O. F.: Über nichtisomorphe elementare blockwiederholungsfreie 2-(8,4,λ)-Blockpläne I. Rostock. Math. Kolloq. 11, 59 - 66 (1979)

/5/ Gronau, H.-D. O. F., und Reimer, R.: Über nichtisomorphe elementare blockwiederholungsfreie 2-(8,4,λ)-Blockpläne III. Rostock. Math. Kolloq. 17, 37 - 47 (1981)

/6/ Harnau, W.: Die Anzahl paarweise nichtisomorpher, elementarer, wiederholungsfreier 2-(9,3,λ)-Blockpläne, Teil II. Rostock. Math. Kolloq. 13, 43 - 47 (1980)

/7/ Иванов, А. В.: Конструктивное перечисление систем инцидентности III. Rostock. Math. Kolloq. 24, 43-62 (1983)

/8/ Lu Jia-Xi: On large sets of disjoint Steiner triple systems I, II, III. J. Combin. Theory Ser. A 34, 140 - 146, 147 - 155, 156 - 182 (1983)

/9/ Rasch, D., und Herrendörfer, G.: Über die Anzahl elementarer BUB in eingeschränkten (v,k)-Familien. Rostock. Math. Kolloq. 8, 71 - 82 (1978)

/10/ Rosa, A.: Intersection properties of Steiner systems. Ann. Discrete Math. 7, 115 - 128 (1980)

/11/ Rosa, A.: personal communication. Eger (Hungary) 1981

/12/ Rosa, A.: information by letter. February 1984

/13/ Teirlinck, L.: On the maximum number of disjoint Steiner triple systems. Discrete Math. 6, 299 - 300 (1973)

received: November 9, 1984

Authors addresses:

Prof. Dr. H.-D. O. F. Gronau
Ernst-Moritz-Arndt-Universität
Greifswald Sektion Mathemaik
Jahnstraße 15 a
DDR-2200 Greifswald

Dipl.-Lehrer I. Rentner
Spezialschule für Mathematik
"Heinrich Hertz"
Frankfurter Allee 14 a
DDR-1035 Berlin

Rostock. Math. Kolloq. 28, 55 - 60 (1985)

13B25
26A42
44A10

Peter Schatte

Partialbruchzerlegung im Reellen mittels Polynomkongruenzen

Die Grenzwertmethode ist ein einfaches Mittel zur Partialbruchzerlegung einer echt gebrochenen rationalen Funktion $f(x) = Z(x)/N(x)$, wenn der Nenner $N(x)$ von $f(x)$ nur lineare Faktoren enthält. Im folgenden soll eine Ergänzung dieser Methode für den Fall beschrieben werden, daß auch quadratische Faktoren auftreten. Dadurch wird es möglich, die Partialbruchzerlegung von $f(x)$ mit $O(n^2)$ ausschließlich reellen Operationen durchzuführen, wo n der Nennergrad von $f(x)$ ist. Wir bemerken, daß der Aufwand zur Zerlegung des Nenners in lineare und quadratische Faktoren auch etwa quadratisch mit n wächst.

1. Die Grenzwertmethode

Bei der Anwendung der Grenzwertmethode wird vorausgesetzt, daß a eine r-fache Nullstelle von $N(x)$ ist. Dann gilt bekanntlich die Darstellung (vgl. etwa /1/, S. 26 und /3/, S. 31)

$$f(x) = \frac{Z(x)}{N(x)} = \frac{A}{(x-a)^r} + f_1(x), \qquad (1)$$

worin A eine Konstante und $f_1(x) = U(x)(x-a)/N(x)$ wieder eine echt gebrochene rationale Funktion mit kleinerem Nennergrad ist. Durch erneute Anwendung von (1) auf $f_1(x)$ usw. gelangt man Schritt für Schritt zur Partialbruchzerlegung von $f(x)$. Die Konstante A kann sehr einfach bestimmt werden, indem man (1) mit $(x-a)^r$ multipliziert und anschließend $x = a$ setzt. Die Funktion $f_1(x)$ ergibt sich dann durch Subtraktion und Kürzen (vgl. etwa /3/, S. 36 und /4/). Zur praktischen Durchführung der Grenzwertmethode kann man vorteilhaft das Horner-Schema nutzen, das ja die Zerlegung $P(x) = Q(x)(x-a) + P(a)$ liefert (vgl. etwa /2/, S. 104).

Beispiel 1: Es sei

$$f(x) = (x^3-13x^2+69)/(x^4+2x^3-11x^2-12x+36).$$

Da 2 ein Teiler von 36 ist, versuchen wir, ob x = 2 eine Nullstelle des Nenners ist. Wir teilen solange durch x-2, bis ein von Null verschiedener Rest entsteht. Durch dreimalige Anwendung des Horner-Schemas erhalten wir

$$x^4 + 2x^3 - 11x^2 - 12x + 36 = \{(x+8)(x-2) + 25\}(x-2)^2.$$

Wir können also den Ansatz (1) für a = 2, r = 2 aufschreiben. Wenn wir nun (1) gleich mit dem Nenner N(x) multiplizieren und die linke Seite erneut mit dem Horner-Schema zerlegen, so ergibt sich

$$(x^2-11x-22)(x-2) + 25 = A\{(x+8)(x-2) + 25\} + U(x)(x-2). \qquad (2)$$

Hieraus erhalten wir A = 1, $U(x) = x^2 - 12x - 30$ und $f_1(x) = (x^2 - 12x - 30)/(x^3 + 4x^2 - 3x - 18)$. Wir behandeln nun $f_1(x)$ in derselben Weise wie zuvor f(x) und erhalten für den Ansatz (1) bei r = 1 die Gleichung

$$(x-10)(x-2) - 50 = A\{(x+8)(x-2) + 25\} + U(x)(x-2). \qquad (3)$$

Die rechten Seiten von (2) und (3) sind identisch. Wir erhalten A = -2, U(x) = 3x+6 und $f_2(x) = (3x+6)/(x^2+6x+9)$. In derselben Weise fahren wir mit der nächsten Nennernullstelle x = -3 fort:

$$x^2 + 6x + 9 = (x+3)^2,$$

und der Ansatz (1) lautet jetzt für a = -3, r = 2 nach Multiplikation mit dem Nenner

$$3(x+3) - 3 = A + U(x)(x+3),$$

woraus sich A = -3, U(x) = 3 ergibt. Da $f_3(x) = 3/(x+3)$ bereits ein Partialbruch ist, bekommen wir das Ergebnis

$$f(x) = \frac{1}{(x-2)^2} - \frac{2}{x-2} - \frac{3}{(x+3)^2} + \frac{3}{x+3}.$$

Leider versagt die Grenzwertmethode, wenn komplexe Nennernullstellen auftreten, man aber im Reellen rechnen will. In diesem Fall kann man nach der Grenzwertmethode zunächst alle linearen Faktoren des Nenners beseitigen. Verbleibt nur ein Paar konjugiert komplexer Nullstellen, so ist schließlich die Partialbruchzerlegung von $f^*(x) = Z^*(x)(x^2 + px + q)^{-r}$ durchzuführen. Dies ist leicht möglich durch (r-1)-malige Division mit Rest

oder durch Koeffizientenvergleich, bei dem ein gestaffeltes Gleichungssystem entsteht.

2. Die Methode der Polynomkongruenzen

Wir setzen nun voraus, daß $N(x)$ den quadratischen Faktor $x^2 + px + q$ mit $4q > p^2$ genau r-mal ($r \geq 1$ ganz) enthält, also $N(x) = (x^2 + px + q)^r S(x)$. Wir ersetzen (1) durch den Ansatz

$$f(x) = \frac{Z(x)}{N(x)} = \frac{Ax+B}{(x^2 + px + q)^r} + f_1(x), \tag{4}$$

worin A und B Konstanten sind und $f_1(x) = U(x)(x^2 + px + q)/N(x)$ wieder echt gebrochen sein soll. Dann ist (4) äquivalent mit

$$Z(x) = (Ax + B)S(x) + U(x)(x^2 + px + q). \tag{5}$$

Die Konstanten A und B werden aus (5) durch Betrachtung von Kongruenzen modulo $x^2 + px + q$ bestimmt. Es gelte

$$\begin{aligned} Z(x) &\equiv fx + g \mod (x^2 + px + q), \\ S(x) &\equiv bx + d \mod (x^2 + px + q), \end{aligned} \tag{6}$$

wobei b und d nicht gleichzeitig verschwinden können, da $S(x)$ nicht durch $x^2 + px + q$ teilbar ist. Es folgt

$$(Ax+B)S(x) \equiv \{(d-bp)A+b\}x + \{-bqA+dB\} \mod (x^2+px+q). \tag{7}$$

Aus (5) bis (7) erhält man das Gleichungssystem

$$\begin{aligned} (d - bp)A + bB &= f, \\ -bqA + dB &= g \end{aligned} \tag{8}$$

für A und B mit der Determinante

$$D = (d - b\tfrac{p}{2})^2 + b^2(q - \tfrac{p^2}{4}) > 0.$$

Damit sind A und B eindeutig bestimmbar. Wir haben nachgewiesen, daß der Ansatz (4) gültig und eindeutig ist und daß A und B in der angegebenen Weise mittels Polynomkongruenzen ermittelt werden können. Man rechnet leicht nach, daß $f_1(x)$ echt gebrochen ist. Durch wiederholte Anwendung des Ansatzes (4) kann man nacheinander alle quadratischen Faktoren des Nenners von $f(x)$ abspalten. Die eindeutige Lösbarkeit von (8) folgt übrigens auch

aus der Äquivalenz von (8) und (4) und aus der allgemein bekannten Gültigkeit und Eindeutigkeit von (4) (vgl. etwa /1/, S. 27 oder /3/, S. 32).
Man erkennt natürlich, daß bei der Methode der Polynomkongruenzen das Rechnen mit komplexen Zahlen in verschleierter Form genutzt wird. Es wird ja im Körper der reellen Zahlen gerechnet, dem eine Lösung x der Gleichung $x^2 + px + q$ adjungiert wurde.

3. Die praktische Durchführung

Die Methode der Polynomkongruenzen kann völlig analog zur Grenzwertmethode realisiert werden. Man hat lediglich das einzeilige Horner-Schema durch das zweizeilige Horner-Schema zu ersetzen, um die Division durch $x^2 + px + q$ mit Rest auszuführen (vgl. etwa /2/, S. 108).

Beispiel 2: Es sei

$$f(x) = \frac{x^5+x^3+16x}{x^8+10x^7+53x^6+178x^5+416x^4+682x^3+781x^2+570x+225}.$$

Etwa nach dem Bairstow-Verfahren (vgl. /2/, S. 109) sei $x^2 + 2x + 3$ als irreduzibler Faktor des Nenners ermittelt. Wir teilen solange durch $x^2 + 2x + 3$, bis ein von Null verschiedener Rest entsteht. Durch dreimalige Anwendung des Horner-Schemas erhält man

$$\begin{aligned} &x^8+10x^7+53x^6+178x^5+416x^4+682x^3+781x^2+570x+225 \\ &\quad = \left\{(x^2+4x+8)(x^2+2x+3) + 2x+1\right\}(x^2+2x+3)^2. \end{aligned} \tag{9}$$

Wir müssen also den Ansatz (4) für r = 2 aufschreiben. Wenn wir auf der linken Seite von (5) erneut das zweizeilige Horner-Schema anwenden und auf der rechten Seite die Identität

$$(Ax + B)(2x + 1) = 2A(x^2 + 2x + 3) + (-3A + 2B)x + (-6A + B)$$

berücksichtigen, so folgt

$$\begin{aligned} &(x^3-2x^2+2x+2)(x^2+2x+3) + 6x-6 \\ &= \left\{(Ax+B)(x^2+4x+8)+2A+U(x)\right\}(x^2+2x+3)+(-3A+2B)x+(-6A+B). \end{aligned} \tag{10}$$

Durch Betrachtung modulo $x^2 + 2x + 3$ erhalten wir A = 2, B = 6, $U(x) = -(x^3 + 16x^2 + 38x + 50)$ und

$$f_1(x) = -(x^3 + 16x^2 + 38x + 50)/(x^6 + 8x^5 + \ldots + 75).$$

Wir verfahren nun mit $f_1(x)$ wie zuvor mit $f(x)$ und erhalten für den Ansatz (4) bei $r = 1$ die Gleichung

$$\begin{aligned} &-(x+14)(x^2+2x+3) - 7x-8 \\ &= \left\{(Ax+B)(x^2+4x+8)+2A+U(x)\right\}(x^2+2x+3)+(-3A+2B)x+(-6A+B). \end{aligned} \tag{11}$$

Die rechten Seiten von (10) und (11) stimmen wieder überein. Es folgt $A = 1$, $B = -2$, $U(x) = -(x^3 + 2x^2 + x)$ und

$$f_2(x) = -(x^3 + 2x^2 + x)/(x^4 + 6x^3 + 19x^2 + 30x + 25).$$

Ein weiterer irreduzibler Faktor des Nenners ist nun $x^2 + 3x + 5$. Wir rechnen

$$x^4 + 6x^3 + 19x^2 + 30x + 25 = (x^2 + 3x + 5)^2, \tag{12}$$

und die Gleichung (5) lautet

$$(-x+1)(x^2 + 3x + 5) + x - 5 = Ax + B + U(x)(x^2 + 3x + 5), \tag{13}$$

woraus sich $A = 1$, $B = -5$, $U(x) = -x + 1$ und

$$f_3(x) = (-x+1)/(x^2 + 3x + 5) \text{ ergibt.}$$

Somit erhalten wir zusammenfassend

$$f(x) = \frac{2x+6}{(x^2+2x+3)^2} + \frac{x-2}{x^2+2x+3} + \frac{x-5}{(x^2+3x+5)^2} + \frac{-x+1}{x^2+3x+5}.$$

4. Aufwandsbetrachtung

Wir nehmen an, daß der Nenner von $f(x)$ genau k quadratische Faktoren mit den Vielfachheiten $v_1, v_2, \ldots, v_k$ enthält. Bei der Feststellung dieser Tatsache fallen die Beziehungen (9) und (12) mit an, gleichgültig, ob mit Polynomkongruenzen oder dem sonst üblichen Koeffizientenvergleich weiter gearbeitet werden soll. Bei der Methode der Polynomkongruenzen sind dann

$$d = \sum_{i=1}^{k} v_i - 1 = \frac{n}{2} - 1$$

zusätzliche Divisionen auszuführen, wobei n der Nennergrad von $f(x)$ ist. Im Beispiel 2 sind dies die Divisionen links in (10),

(11) und (13). Liegt die Zerlegung des Nenners bereits vor, so sind noch k-1 weitere Divisionen auszuführen, nämlich im Beispiel 2 die Division (9). In jedem Fall beträgt der Aufwand für die Methode der Polynomkongruenzen $O(n^2)$ Operationen. Beim Koeffizientenvergleich (oder bei der Einsetzmethode) ist ein System von n linearen Gleichungen mit n Unbekannten aufzulösen, was mit wesentlich höherem Aufwand verbunden ist. Im Beispiel 2 stehen 3 Divisionen mit Rest gegen die Auflösung eines Gleichungssystems der Ordnung 8. Die Methode der Polynomkongruenzen scheint also schon vorteilhaft zu sein, wenn mehr als ein Paar konjugiert komplexer Nullstellen vorhanden ist. Tritt höchstens ein solches Paar auf, so kann man die Grenzwertmethode benutzen.

Man kann also bei der Partialbruchzerlegung völlig auf die Auflösung großer Gleichungssysteme verzichten. Besonders effektiv wird die Partialbruchzerlegung, wenn sie gemeinsam mit der Faktorzerlegung des Nenners durchgeführt wird.

Literatur

/1/ Berg, L.: Einführung in die Operatorenrechnung. Berlin 1965

/2/ Kiesewetter, H., und Maeß, G.: Elementare Methoden der numerischen Mathematik. Berlin 1974

/3/ Mangoldt, H. v., und Knopp, K.: Einführung in die höhere Mathematik, Bd. III. Leipzig 1957

/4/ Straight, H. J., and Dowds, R.: An alternate method for finding the partial fraction decomposition of a rational function. Amer. Math. Monthly 91, 365 - 367 (1984)

eingegangen: 16. 01. 1985

Anschrift des Verfassers:

Doz. Dr. sc. P. Schatte
Bergakademie Freiberg
Sektion Mathematik
Bernhard-von Cotta-Str. 2
DDR-9200 Freiberg

Rostock. Math. Kolloq. 28, 61 - 65 (1985)

94B35
94B60

Anna Racsmány

Perfekte Codes für einen verallgemeinerten Hamming-Abstand

In der Codierungstheorie werden Unterschiede in der Struktur der Kanäle durch Heranziehen von verschiedenen Abstandsbegriffen berücksichtigt. Die am häufigsten untersuchten Abstände sind der Hamming- und der Lee-Abstand. In der vorliegenden Arbeit betrachten wir Codes bezüglich eines von G. O. H. Katona eingeführten allgemeinen Abstandes. Dieser Abstand ist eine Verallgemeinerung des Hamming-Abstandes und bringt die Bedingung zum Ausdruck, daß in dem Kanal bei der Transmission der Zeichen 1,2,...(mod q) häufig kleinere Sprünge vorkommen, größere aber, die als Fehler angesehen werden, nur selten. Im folgenden wollen wir einige codetheoretische Fragen für diesen Abstandsbegriff untersuchen. Insbesondere werden wir die Frage über die Existenz von perfekten Codes bezüglich dieses Abstands auf die Frage über die Existenz von perfekten Codes für die Hamming-Metrik zurückführen.

Betrachten wir die n-dimensionalen Vektoren über dem Alphabet $\{0,1,\ldots,q-1\}$. Es sei $a \leq \frac{q+1}{2}$ eine natürliche Zahl. Wir definieren für je zwei Vektoren $\underline{x}^1$ und $\underline{x}^2$ den "Abstand"

$$d(\underline{x}^1,\underline{x}^2) = \sum_{i=1}^{n} d(x_i^1,x_i^2), \text{ mit}$$

$$d(x_i^1,x_i^2) = \begin{cases} 0, & \text{wenn } \min(|x_i^1-x_i^2|,\ q-|x_i^1-x_i^2|) < a \\ 1, & \text{wenn } \min(|x_i^1-x_i^2|,\ q-|x_i^1-x_i^2|) \geq a. \end{cases}$$

Offensichtlich ist $d(\underline{x}^1,\underline{x}^2)$ im Falle $a = 1$ der Hamming-Abstand von $\underline{x}^1$ und $\underline{x}^2$. Leider genügt $d(\underline{x}^1,\underline{x}^2)$ für $a > 1$ nicht der Dreiecksungleichung. Ist z. B. $n = 3$, $q = 4$, $a = 2$, $\underline{x}^1 = (0,0,0)$, $\underline{x}^2 = (1,1,1)$, $\underline{x}^3 = (2,2,2)$, so gilt $d(\underline{x}^1,\underline{x}^2) + d(\underline{x}^2,\underline{x}^3) < d(\underline{x}^3,\underline{x}^1)$.

Vorerst untersuchen wir die um $\underline{x}$ mit dem Radius e geschlagene "Kugel" $K(\underline{x},e)$, d. h. die Menge derjenigen Punkte $\underline{y}$, für die $d(\underline{x},\underline{y}) \leq e$ ist. Wir merken an, daß $K(\underline{x},o)$ nicht leer ist.

Hilfssatz 1: Zwei Kugeln $K(\underline{x}^1,e)$ und $K(\underline{x}^2,e)$ sind dann und nur dann disjunkt, wenn es unter den Koordinaten von $\underline{x}^1 - \underline{x}^2$ mindestens $2e + 1$ solche gibt, deren Betrag mindestens gleich $2a - 1$ ist.

Beweis: Wir zeigen, daß $K(\underline{x}^1,e)$ und $K(\underline{x}^2,e)$ dann und nur dann nicht disjunkt sind, wenn unter den Koordinaten von $\underline{x}^1 - \underline{x}^2$ der Betrag von mindestens $n - 2e$ höchstens $2a - 2$ ist.

Möge es einen Punkt $\underline{s}$ mit $\underline{s} \in K(\underline{x}^1,e) \cap K(\underline{x}^2,e)$ geben. Dann gibt es zwei Punkte $\underline{y}^1$ und $\underline{y}^2$, so daß $\underline{x}^1 + \underline{y}^1 = \underline{s}$ und $\underline{x}^2 + \underline{y}^2 = \underline{s}$ ist. Wegen $\underline{s} \in K(\underline{x}^i,e)$ hat $\underline{y}^i$ höchstens e solche Koordinaten, deren Beträge größer sind als $a - 1$ $(i=1,2)$. Deshalb hat $\underline{y}^1 - \underline{y}^2$ höchstens $2e$ solche Koordinaten, deren Beträge größer sind als $2a - 2$. Wegen $\underline{x}^2 - \underline{x}^1 = \underline{y}^1 - \underline{y}^2$ gilt aber dasselbe für $\underline{x}^1 - \underline{x}^2$. Wir müssen noch zeigen, daß es einen Vektor $\underline{s}$ mit $\underline{s} \in K(\underline{x}^1,e) \cap K(\underline{x}^2,e)$ gibt, wenn unter den Beträgen der Koordinaten von $\underline{x}^1 - \underline{x}^2$ höchstens $2e$ größer sind als $2a - 2$. Offensichtlich läßt sich $\underline{x}^1 - \underline{x}^2$ als Differenz von solchen Vektoren $\underline{y}^1$ und $\underline{y}^2$ darstellen, in welchen höchstens e Koordinaten größer sind als $a - 1$. Nun gilt für den Vektor $\underline{s} = \underline{x}^1 + \underline{y}^1 = \underline{x}^2 + \underline{y}^2$ einerseits $\underline{s} = \underline{x}^1 + \underline{y}^1 \in K(\underline{x}^1,e)$, andererseits $\underline{s} = \underline{x}^2 + \underline{y}^2 \in K(\underline{x}^2,e)$.

Hilfssatz 2: Die Anzahl der Punkte in $K(\underline{x},e)$ ist

$$|K(\underline{x},e)| = (2a-1)^{n-e} \sum_{i=0}^{e} \binom{n}{i} (q-(2a-1))^i (2a-1)^{e-i}.$$

Beweis: Wir erhalten durch einfaches Abzählen

$$|K(x,o)| = (2a-1)^n,$$
$$|K(x,1)| - |K(x,o)| = n(q-(2a-1))(2a-1)^{n-1},$$
$$\vdots$$
$$|K(x,i)| - |K(x,i-1)| = \binom{n}{i}(q-(2a-1))^i(2a-1)^{n-i},$$
$$\vdots$$
$$|K(x,e)| - |K(x,e-1)| = \binom{n}{e}(q-(2a-1))^e(2a-1)^{n-e}.$$

Addieren wir diese Ausdrücke, so ergibt sich die gewünschte Relation.

Satz 1: Ein e-fehlerkorrigierender Code kann nur dann perfekt sein, wenn $2a-1 \mid q$ gilt.

Beweis: Ist der Code perfekt, so ist $|C| = \frac{q^n}{|K(\underline{x},e)|}$ eine ganze Zahl, nämlich die Anzahl $|C|$ der Codewörter.
Nach Hilfssatz 2 haben wir

$$|C| = \frac{q^n}{(2a-1)^{n-e}\sum_{i=o}^{e}\binom{n}{i}(q-(2a-1))^i(2a-1)^{e-i}},$$

und hieraus folgt $2a-1 \mid q^n$. Es sei p der größte gemeinsame Teiler von $2a-1$ und q. Aus

$$|C| = \frac{p^n\left(\frac{q}{p}\right)^n}{p^{n-e}\left(\frac{2a-1}{p}\right)^{n-e}\sum_{i=o}^{e}\binom{n}{i}p^i\left(\frac{q}{p}-\frac{2a-1}{p}\right)^i p^{e-i}\left(\frac{2a-1}{p}\right)^{e-i}}$$

$$= \frac{\left(\frac{q}{p}\right)^n}{\left(\frac{2a-1}{p}\right)^{n-e}\sum_{i=o}^{e}\binom{n}{i}\left(\frac{q}{p}-\frac{2a-1}{p}\right)^i\left(\frac{2a-1}{p}\right)^{e-i}}$$

folgt $\frac{2a-1}{p} \mid \left(\frac{q}{p}\right)^n$. Da aber die Zahlen $\frac{2a-1}{p}$ und $\frac{q}{p}$ relativ prim sind, gilt $\frac{2a-1}{p} = 1$, woraus sich $2a-1 \mid q$ ergibt.

Es bedeute C(n,a,q,e) einen Code mit den betreffenden Parametern.

Satz 2: In dem Code $C(n,a,q,e)$ mögen q und $2a-1$ einen gemeinsamen Teiler $s>1$ haben, d. h. $q = s\cdot\bar{q}$ und $2a-1 = s(2\bar{a}-1)$. Dann läßt sich ein Code $C(n,\bar{a},\bar{q},e)$ mit derselben Anzahl von Codewörtern wie $C(n,a,q,e)$ konstruieren.

Beweis: Wir ordnen jedem Vektor $\underline{c} = (c_1,\ldots,c_n)$ des Codes $C(n,a,q,e)$ denjenigen Vektor $\underline{\bar{c}} = (\bar{c}_1,\ldots,\bar{c}_n)$ zu, für den $c_i = \bar{c}_i s + z$ mit $0 \leq z < s$ ist. Dann gilt offensichtlich $0 \leq \bar{c}_i \leq \frac{c_i}{s} < \frac{q}{s} = \bar{q}$. Sind $\underline{c}^1$ und $\underline{c}^2$ zwei Vektoren aus dem Code $C(n,a,q,e)$, so besteht für mindestens $2e+1$ Koordinaten die Ungleichung $|c_i^1 - c_i^2| \geq 2a-1$. Wegen der Definition von $\bar{c}_i$ haben wir

$$c_i^1 = \bar{c}_i^1 s + z^1, \quad 0 \leq z^1 < s,$$

$$c_i^2 = \bar{c}_i^2 s + z^2, \quad 0 \leq z^2 < s.$$

Deshalb gilt

$$|c_i^1 - c_i^2| = |(\bar{c}_i^1 - \bar{c}_i^2)s + z^1 - z^2| \geq 2a-1 = s(2\bar{a}-1).$$

Hieraus folgt wegen $z^1 - z^2 < s$, daß $|\bar{c}_i^1 - \bar{c}_i^2| \geq 2\bar{a}-1$ ist. Aus $\underline{c}^1 \neq \underline{c}^2$ folgt also $\underline{\bar{c}}^1 \neq \underline{\bar{c}}^2$ und wir haben für mindestens $2e+1$ Koordinaten $|\bar{c}_i^1 - \bar{c}_i^2| \geq 2\bar{a}-1$. Dies bedeutet, daß $\underline{\bar{c}}^1$ und $\underline{\bar{c}}^2$ Codewörter eines Codes $C(n,\bar{a},\bar{q},e)$ sind.

Korollar 1: Wenn in einem Code $C(n,a,q,e)$ die Beziehung $2a-1 \mid q$ gilt, dann gibt es einen Code der Länge n über einem Alphabet von $\frac{q}{2a-1}$ Buchstaben, der e Fehler korrigiert, in der Hamming-Metrik und dieselbe Anzahl von Codewörtern hat wie $C(n,a,q,e)$.

Korollar 2: Ein Code $C(n,a,q,e)$ ist dann und nur dann perfekt, wenn ein perfekter Code der Länge n über einem Alphabet von $\frac{q}{2a-1}$ Buchstaben existiert, der e Fehler korrigiert in der Hamming-Metrik.

Die Frage, für welche Parameter es einen perfekten Code für die Hamming-Metrik gibt, ist noch offen. Nach der sogenannten "Perfekte-Codes-Vermutung" gibt es perfekte Codes, außer gewissen trivialen Codes, nur für die Parameter eines Hamming-Codes und eines binären oder ternären Golay-Codes. Diese Vermutung wurde von Best /1/ "fast" bewiesen, indem er zeigte, daß außer den bekannten perfekten Codes weitere e-fehlerkorrigierende Codes nur für e = 1,2,6 und 8 existieren können. Für den Fall, daß die Anzahl der Buchstaben im Alphabet eine Primzahlpotenz ist, wurde die Perfekte-Codes-Vermutung von van Lint /2/, Tietäväinen /4/ sowie V. A. Zinov'ev und V. K. Leont'ev /5/ bewiesen (s. auch /3/ Ch. 6. Theorem 33).

Literatur

/1/ Best, M. R.: Perfect Codes Hardly Exist. IEEE Trans. Inform. Theory, IT-29, 349 - 351 (1983)

/2/ van Lint, J. H.: On the nonexistence of certain perfect codes. In: Atkin, A. O. L., and Birch, B. J. (Eds.): Computers in number theory. Proc. Atlas Symp. No. 2, Oxford 1969, 277 - 282, London 1971

/3/ MacWilliams, F. J., and Sloane, N. J. A.: The Theory of Error-Correcting Codes. North-Holland 1977

/4/ Tietäväinen, A.: On the nonexistence of perfect codes over finite fields. SIAM J. Appl. Math. 24, 88 - 96 (1973)

/5/ Zinov'ev, V. A., i Leont'ev, V. K.: Ne suščestvovanie soveršennych kodov nad poljami Galois. Probl. Upravl. Teorie Inform. 2, 123 - 132 (1973)

eingegangen: 07. 01. 1985

Anschrift des Verfassers:

Dr. A. Racsmány
Karl-Marx-Universität Budapest
Institut für Mathematik
Dimitrov tér 8
Levelaim 1828 Bp. 5 Pf. 489
1093 Budapest, VR Ungarn

Rostock. Math. Kolloq. 28, 66 - 74 (1985) 06D05
20M14

Lothar Berg

Distributive Verbände als relativ invertierbare Halbgruppen[1]

Der Begriff der relativ invertierbaren Halbgruppe wurde von K. Markwardt /3/ eingeführt und in /1/, /4/ weiter untersucht. Ein Beispiel hierzu ist der Boolesche Verband mit

$$a + b = a \vee b,\ a - b = a \setminus (a \wedge b), \tag{1}$$

wobei auf der rechten Seite der zweiten Gleichung das Komplement von $a \wedge b$ bez. a steht. Im folgenden soll dieses Beispiel auf beliebige endliche distributive Verbände übertragen und anschließend in zweifacher Hinsicht verallgemeinert werden.

Definition: Eine (teilweise) geordnete Menge $M(\leq,+,-)$ heißt eine relativ invertierbare Halbgruppe, wenn sie bezüglich der "Addition +" eine kommutative Halbgruppe mit Nullelement o bildet und wenn sie bezüglich der "Subtraktion -" das folgende Axiom erfüllt: Für beliebige $a,b,x \in M$ ist

$$a \leq b + x \text{ äquivalent zu } a - b \leq x. \tag{2}$$

Nach /3/ ergibt sich als Folgerung, daß die Addition bezüglich beider Summanden monoton wachsend ist, während die Subtraktion $a - b$ bezüglich a monoton wachsend und bezüglich b monoton fallend ist. Für beliebige a,b,c gilt außerdem $a - o = a$ sowie

$$a - (b+c) = (a-b) - c. \tag{3}$$

Nach /1/ ist die Subtraktion in M durch

$$a - b = \min (x \mid a \leq b + x) \tag{4}$$

eindeutig bestimmt, und es ist leicht zu sehen, daß auch umgekehrt gilt:

$$a + b = \max (x \mid x - b \leq a). \tag{5}$$

---

1 Herrn Dr. K. Markwardt sei für kritische Hinweise zur Arbeit vielmals gedankt.

Bis auf weiteres sei M zusätzlich ein Verband mit o als kleinstem Element. Hierüber kann man aus K. Markwardt /3/ folgende Aussagen entnehmen.

Satz 1: Für beliebige $a,b,c \in M$ gilt

$$a - a = o - a = o, \tag{6}$$

$$a + (b \wedge c) = (a+b) \wedge (a+c), \tag{7}$$

$$a - (b \wedge c) = (a-b) \vee (a-c), \tag{8}$$

$$(a \vee b) - c = (a-c) \vee (b-c). \tag{9}$$

Ist die Addition in M wie in (1) gleich dem Verbandssupremum, so besagt die Gleichung (7), daß der Verband dann distributiv ist. Für endliche Verbände, auf die wir uns der Einfachheit halber beschränken wollen, gilt hiervon auch die Umkehrung.

Satz 2: Jeder endliche distributive Verband ist mit dem Verbandssupremum als Addition eine relativ invertierbare Halbgruppe.

Beweis: Aus $a \leq b + x$ und $a \leq b + y$ folgt wegen (7)

$$a \leq (b+x) \wedge (b+y) = b + (x \wedge y),$$

so daß das Minimum in (4) existiert und die Behauptung nach Satz 1 von /1/ bewiesen ist.

Um jetzt für die Subtraktion eine explizite Darstellung anzugeben, berufen wir uns nach H. Gericke /2/ auf den folgenden

Strukturssatz: Ein endlicher Verband ist genau dann distributiv, wenn jedes Element a eine (bis auf die Reihenfolge) eindeutig bestimmte direkt unverkürzbare Zerlegung

$$a = a_{n_1} \vee a_{n_2} \vee \ldots \vee a_{n_k} \tag{10}$$

in irreduzible Elemente besitzt.

Es seien $a_o = o$, a_1, ..., a_r die irreduziblen Elemente des distributiven Verbandes. Dann gilt mit dem Verbandssupremum als Addition nach (4)

$$a_i - a_j = \begin{cases} o & \text{für } a_i \leq a_j, \\ a_i & \text{sonst.} \end{cases}$$

Ist jetzt $b = a_{m_1} \vee \ldots \vee a_{m_l}$ die unverkürzbare Zerlegung eines zweiten Elementes, so ist $a - b$ offenbar das Supremum von o und denjenigen Elementen a_{n_i} von (10), für die es kein a_{m_j} aus der Zerlegung von b mit $a_{n_i} \leq a_{m_j}$ gibt. Da $a \wedge b$ das Supremum aller $a_{n_i} \wedge a_{m_j}$ ist, können wir für die Differenz "-" wieder die zweite der Formeln (1) verwenden, wenn wir dort die Differenz " \ " als Streichung der den unverkürzbaren Zerlegungen von a und $a \wedge b$ gemeinsamen Elemente interpretieren, während die übrigen Elemente der Zerlegung von a stehenbleiben.

Erste Verallgemeinerung: Es sei M_o' ein beliebiger endlicher distributiver Verband, der mit den Verknüpfungen (1) zugleich eine relativ invertierbare Halbgruppe ist, und es sei G eine beliebige endliche abelsche Gruppe der Ordnung k+1 mit den Elementen $u_o = o, u_1, \ldots, u_k$. Dabei mögen $u_1, \ldots, u_l$ die l erzeugenden Elemente dieser Gruppe sein. Wir konstruieren jetzt eine neue relativ invertierbare Halbgruppe M mit k+1 Komponenten $M_o, M_1, \ldots, M_k$ im Sinne von /1/. Zu diesem Zweck greifen wir aus M_o' genau l Elemente $c_1, \ldots, c_l$ heraus, die nicht untereinander verschieden zu sein brauchen, und führen die distributiven Teilverbände $M_i' = \{a | a \geq c_i\}$ von M_o' mit der Filtereigenschaft $M_o' + M_i' = M_i'$ ein.

Jedes Element u_i der Gruppe G besitzt bekanntlich eine Darstellung der Form $u_i = \beta_{i1} u_1 + \ldots + \beta_{il} u_l$ mit nichtnegativen ganzen Zahlen β_{ij}, die kleiner als die Ordnung von u_j gewählt werden können. Mit diesen Zahlen lassen sich einerseits die vorhergehenden Filter einschließlich M_o' in der Form

$$M_i' = \{a | a \geq \beta_{i1} c_1 + \ldots + \beta_{il} c_l\} \qquad (11)$$

schreiben und andererseits für $1 < i \leq k$ durch (11) neue Filter M_i' einführen, so daß allgemein die Eigenschaft $M_i' + M_j' = M_t'$ für $u_i + u_j = u_t$ vorliegt. Die gesuchten Komponenten M_i definieren

wir schließlich als Mengen der Paare (a,u_i) mit $a \in M_i'$ sowie mit folgenden Relationen und Verknüpfungen. Die Ordnung in M_i wird durch die Ordnung in M_i' festgelegt, die Addition wird komponentenweise erklärt, und die Subtraktion lautet nach (4)

$$(a,u_i) - (b,u_j) = (c,u_i-u_j), \tag{12}$$

wobei c dasjenige Element von M_s' mit $u_s = u_i - u_j$ ist, das sich aus a - b ergibt, wenn in der unverkürzbaren Zerlegung dieses Elementes die irreduziblen Elemente a_p durch die kleinsten Elemente b_q mit $a_p \leq b_q$ ersetzt werden, die in M_s' irreduzibel sind. Bezeichnen wir die Linearkombination der c_j in (11) auch für i = 0 sowie für $1 < i$ mit c_i, so läßt sich das Element c in (12) auch in der Form $c = (a-b) + c_s$ schreiben. Es ist unmittelbar klar, daß hierdurch eine relativ invertierbare Halbgruppe M entsteht. Besteht der Durchschnitt von M_0' und G genau aus dem Nullelement o, so können wir die Schreibweise $(a,u_i) = a + u_i$ verwenden. Dann ist $M_0 = M_0'$, die der zugehörigen Komponenten M_i minimalen Elemente lauten $c_i + u_i$, und die Subtraktion (12) ist ebenfalls komponentenweise ausführbar:

$$(a+u_i) - (b+u_j) = (a-b) + (u_i-u_j).$$

Bei der vorhergehenden Konstruktion können die Filter M_i' für $i > 1$ offenbar auch durch umfassendere Filter ersetzt werden, wenn nur im Fall $u_i + u_j = u_t$ die Eigenschaft $M_i' + M_j' \subset M_t'$ für alle i,j erhalten bleibt. Bei Ersetzung der Filterdefinition durch eine geeignete Idealdefinition ist die Konstruktion darüber hinaus auch durchführbar, wenn M_0' kein Verband, sondern eine beliebige relativ invertierbare Halbgruppe mit einem größten Element ist, wodurch Satz 4 von /1/ verallgemeinert wird. Weiterhin ist unter den Zusatzvoraussetzungen von /2/ und /3/ auch eine Übertragung auf den unendlichen Fall möglich.

Beispiel: Es sei M_0 der distributive Verband mit den nichttrivialen irreduziblen Elementen a_1, a_2, a_3 und $a_1 \leq a_2$ als einziger Ordnungsrelation zwischen diesen Elementen. Das maximale Element $m = a_2 \vee a_3$ ist wegen der Eindeutigkeit der Zerlegung (10)

von $b = a_1 \vee a_3$ verschieden, so daß der Ordnungsgraph mit dem minimalen Element $o = a_1 \wedge a_3$ die in Abb. 1 angegebene Gestalt besitzt.

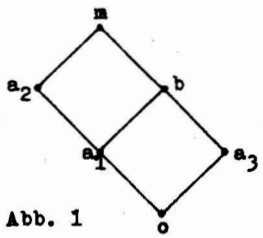


Abb. 1

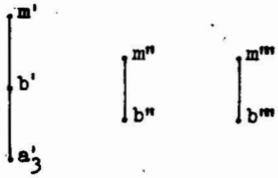


Abb. 2

Wir wählen G als Kleinsche Vierergruppe und wählen weiterhin

$M_1' = \{a_3, b, m\}$, $M_2' = \{b, m\}$,

so daß nach unserer Konstruktion $M_3' = M_2'$ wird. Mit den Abkürzungen

$a' = a + u_1$, $a'' = a + u_2$, $a''' = a + u_3$

ergeben sich dann die in Abb. 2 angegebenen Ordnungsgraphen und nach Weglassen der bekannten Operationen mit o die Additionstabelle

+	a_1	a_2	a_3	b	m	a_3'	b'	m'	b''	m''	b'''	m'''
a_1	a_1	a_2	b	b	m	b'	b'	m'	b''	m''	b'''	m'''
a_2	a_2	a_2	m	m	m	m'	m'	m'	m''	m''	m'''	m'''
a_3	b	m	a_3	b	m	a_3'	b'	m'	b''	m''	b'''	m'''
b	b	m	b	b	m	b'	b'	m'	b''	m''	b'''	b'''
m	m	m	m	m	m	m'	m'	m'	m''	m''	m'''	m'''
a_3'	b'	m'	a_3'	b'	m'	a_3	b	m	b'''	m'''	b''	m''
b'	b'	m'	b'	b'	m'	b	b	m	b'''	m'''	b''	m''
m'	m'	m'	m'	m'	m'	m	m	m	m'''	m'''	m''	m''
b''	b''	m''	b''	b''	m''	b'''	b'''	m'''	b	m	b'	m'
m''	m''	m''	m''	m''	m''	m'''	m'''	m'''	m	m	m'	m'
b'''	b'''	m'''	b'''	b'''	m'''	b''	b''	m''	b'	m'	b	m
m'''	m'''	m'''	m'''	m'''	m'''	m''	m''	m''	m'	m'	m	m

sowie die Subtraktionstabelle

−	a_1	a_2	a_3	b	m	a_3'	b'	m'	b''	m''	b'''	m'''
a_1	o	o	a_1	o	o	b'	a_3'	a_3'	b''	b''	b'''	b'''
a_2	a_2	o	a_2	a_2	o	m'	m'	a_3'	m''	b''	m'''	b'''
a_3	a_3	a_3	o	o	o	a_3'	a_3'	a_3'	b''	b''	b'''	b'''
b	a_3	a_3	a_1	o	o	b'	a_3'	a_3'	b''	b''	b'''	b'''
m	m	a_3	a_2	a_2	o	m'	m'	a_3'	m''	b''	m'''	b'''
a_3'	a_3'	a_3'	a_3'	a_3'	a_3'	o	o	o	b'''	b'''	b''	b''
b'	a_3'	a_3'	b'	a_3'	a_3'	a_1	o	o	b'''	b'''	b''	b''
m'	m'	a_3'	m'	m'	a_3'	a_2	a_2	o	m'''	b'''	m''	b''
b''	b''	b''	b''	b''	b''	b'''	b'''	b'''	o	o	a_3'	a_3'
m''	m''	b''	m''	m''	b''	m'''	m'''	b'''	a_2	o	m'	a_3'
b'''	b'''	b'''	b'''	b'''	b'''	b''	b''	b''	a_3'	a_3'	o	o
m'''	m'''	b'''	m'''	m'''	b'''	m''	m''	b''	m'	a_3'	a_2	o

Ein Beispiel für einen umfassenderen Filter an Stelle von M_3' ist $\{a_1, a_2, b, m\}$, doch dieser Fall läßt sich nach einer Umnumerierung auch mit Hilfe der ursprünglichen Konstruktion begründen.

Zweite Verallgemeinerung: Es sei M eine endliche additive kommutative Halbgruppe mit Nullelement. Aus M wählen wir l Basiselemente $a_1, \ldots, a_l$ aus, so daß jedes beliebige Element $b \in M$ eine Darstellung der Form

$$b = \beta_1 a_1 + \ldots \beta_l a_l \qquad (13)$$

mit nichtnegativen ganzzahligen Koeffizienten besitzt (und umgekehrt). Wegen der Endlichkeit von M gibt es zu jedem Element a_i eine natürliche Zahl n_i, so daß die Elemente $o, a_i, 2a_i, \ldots, n_i a_i$ paarweise verschieden sind und in (13) stets $\beta_i \leq n_i$ gewählt werden kann.

Zusätzlich fordern wir:

1°. Jedes Element b besitzt eine (eindeutig bestimmte) Maximaldarstellung $b = \beta_1^+ a_1 + \ldots + \beta_l^+ a_l$ mit $\beta_i^+ \leq n_i$ und $\beta_i \leq \beta_i^+$ für jede andere Darstellung mit $\beta_i \leq n_i$.

2°. Für die Summe zweier Elemente $a = \alpha_1 a_1 + \ldots + \alpha_l a_l$ und (13) gilt im Fall $\alpha_i^+ + \beta_i^+ \leq n_i$ stets $(\alpha_i+\beta_i)^+ = \alpha_i^+ + \beta_i^+$.

Weiterhin möge in M eine monotone Halbordnung mit folgenden Eigenschaften einführbar sein:

3°. Für jedes Element a_i existiert eine minimal gewählte natürliche Zahl k_i mit $k_i a_i \geq 0$.

4°. Für zwei beliebige Elemente a und b mit $\alpha_i \leq n_i$ für alle i ist genau dann $a \leq b$, wenn es nichtnegative ganze Zahlen p_i gibt mit $\beta_i^+ = \alpha_i + p_i k_i$ für alle i.

5°. Im Fall $\beta_i > n_i$ ist $\beta_i^+ = \beta_i - q_i k_i$ die größte Zahl dieser Form mit $\beta_i^+ \leq n_i$ und einer natürlichen Zahl q_i.

Aus 3° folgt, daß es zu jedem Element a eine Zahl $k \geq 1$ gibt mit $ka \geq 0$ und daß das Nullelement das einzige Element a mit $a \leq o$ ist.

Satz 3: Mit den Eigenschaften 1° bis 5° ist M eine relativ invertierbare Halbgruppe mit der Subtraktion $a - b = c = \gamma_1 a_1 + \ldots + \gamma_l a_l$, deren Koeffizienten folgendermaßen definiert sind:

$$\gamma_i = \begin{cases} \alpha_i^+ - \beta_i^+ \text{ für } \alpha_i^+ \geq \beta_i^+ \\ \alpha_i^+ - \beta_i^+ - \left[\dfrac{\alpha_i^+ - \beta_i^+}{k_i}\right] k_i \text{ für } \beta_i^+ \geq \alpha_i^+. \end{cases} \tag{14}$$

Beweis: Für die Koeffizienten (14) gilt $\beta_i^+ + \gamma_i = \alpha_i^+ + j_i k_i$ mit $j_i \geq 0$ für alle i. Im Fall $\beta_i^+ + \gamma_i \leq n_i$ ist $(\beta_i^+ + \gamma_i)^+ = \beta_i^+ + \gamma_i + p_i k_i$ mit $p_i \geq 0$, und im Fall $\beta_i^+ + \gamma_i > n_i$ ist $(\beta_i^+ + \gamma_i)^+ = \beta_i^+ + \gamma_i - q_i k_i \geq \alpha_i^+$ und daher $j_i - q_i \geq 0$, d. h. $a \leq b + c$.

Somit bleibt nach (4) nur noch zu zeigen, daß für jedes Element $x = \xi_1 a_1 + \ldots + \xi_l a_l$ mit $a \leq b + x$ auch $c \leq x$ gilt. Die erste dieser Ungleichungen bedeutet $(\beta_i+\xi_i)^+ = \alpha_i^+ + p_i k_i$ mit $p_i \geq 0$. Wegen $(\beta_i+\xi_i)^+ = \beta_i^+ + \xi_i^+ - q_i k_i$ mit $q_i \geq 0$ folgt hieraus

$\xi_i^+ = \alpha_i^+ - \beta_i^+ + (p_i+q_i)k_i = \gamma_i + r_i k_i$ mit $r_i \geq 0$ und damit $c \leq x$.

Als Anwendung sei gezeigt, daß sich das vorhergehende Beispiel dem Satz 3 unterordnen läßt, wenn wir die vier Basiselemente

a_1, a_2, $a_4 = a_3'$, $a_5 = b''$

wählen mit $n_1 = k_1 = n_2 = k_2 = 1$ und $n_4 = k_4 = n_5 = k_5 = 2$. Das Element a_3 des Beispiels ist kein Basiselement. Die Maximaldarstellungen lauten

$a_1 = a_1$	$b = a_1 + 2a_4 + 2a_5$	$a_5 = a_1 + 2a_4 + a_5$
$a_2 = a_1 + a_2$	$m = a_1 + a_2 + 2a_4 + 2a_5$	$m'' = a_1 + a_2 + 2a_4 + a_5$
$a_3 = 2a_4$	$b' = a_1 + a_4 + 2a_5$	$b''' = a_1 + a_4 + a_5$
$a_4 = a_4$	$m' = a_1 + a_2 + a_4 + 2a_5$	$m''' = a_1 + a_2 + a_4 + a_5$,

die übrigen nichttrivialen Darstellungen sind

$b = 2a_5 = a_1+2a_4 = a_1+2a_5 = 2a_4+2a_5$

$m = a_2+2a_4 = a_2+2a_5 = a_1+a_2+2a_4 = a_1+a_2+2a_5 = a_2+2a_4+2a_5$

$b' = a_1+a_4 = a_4+2a_5$

$m' = a_2+a_4 = a_1+a_2+a_4 = a_2+a_4+2a_5$

$a_5 = a_1+a_5 = 2a_4+a_5$

$m'' = a_2+a_5 = a_1+a_2+a_5 = a_2+2a_4+a_5$

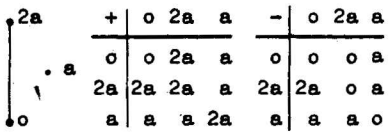
$b''' = a_4+a_5$

$m''' = a_2+a_4+a_5$.

Außer der zyklischen Gruppe dritter Ordnung gibt es vier relativ invertierbare Halbgruppen mit drei Elementen, die nachfolgend angegeben werden, wobei die beiden Fälle III dieselbe Additionstabelle besitzen. Bis auf den Fall III_2, wo ein negatives Element vorkommt, wird die Differenz nach Formel (14) gebildet. Abschließend wird noch unter IV die einzige kommutative Halb-

gruppe mit drei Elementen angegeben, die nicht relativ invertierbar ist.

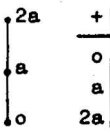
I



+	o	2a	a
o	o	2a	a
2a	2a	2a	a
a	a	a	2a

-	o	2a	a
o	o	o	a
2a	2a	o	a
a	a	a	o

II



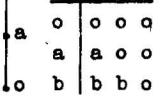
+	o	a	2a
o	o	a	2a
a	a	2a	2a
2a	2a	2a	2a

-	o	a	2a
o	o	o	o
a	a	o	o
2a	2a	a	o

III

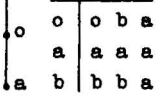
+	o	a	b
o	o	a	b
a	a	a	b
b	b	b	b

III_1



-	o	a	b
o	o	o	o
a	a	o	o
b	b	b	o

III_2



-	o	a	b
o	o	b	a
a	a	a	a
b	b	b	a

IV

+	o	a	b
o	o	a	b
a	a	o	b
b	b	b	b

Literatur

/1/ Berg, L.: Über die Struktur der endlichen relativ invertierbaren Halbgruppen. Math. Nachr. 117, 229 - 234 (1984)

/2/ Gericke, H.: Theorie der Verbände. Mannheim 1963

/3/ Markwardt, K.: Analysis in relativ invertierbaren Halbgruppen. Wiss. Z. Hochsch. Architektur und Bauwesen Weimar 26, 55 - 58 (1979)

/4/ Markwardt, K.: Maß und Integral auf einheitlicher algebraischer Grundlage. Rostock. Math. Kolloq. 29, (1986), im Druck

eingegangen: 26. 10. 1984

Anschrift des Verfassers:

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

Rostock. Math. Kolloq. 28, 75 - 86 (1985) 46F10

Brian Fisher

On defining the change of variable in distributions

In the following we let N be the neutrix, see van der Corput /1/, having domain $N' = \{1,2,\ldots, n,\ldots\}$ and range N'' the real numbers, with negligible functions linear sums of the functions $n^{\lambda}\ln^{r-1}n$, $\ln^{r}n$ for $\lambda > 0$ and $r = 1,2,\ldots$, and all functions which converge to zero as n tends to infinity.
It follows that if

$$f(n) = f_1(n) + f_2(n),$$

where $f_1(n)$ is a negligible function and the limit as n tends to infinity of $f_2(n)$ exists, then the neutrix limit as n tends to infinity of $f(n)$ exists and

$$N\text{-}\lim_{n\to\infty} f(n) = \lim_{n\to\infty} f_2(n).$$

Now let ϱ be a fixed infinitely differentiable function having the properties

(i) $\varrho(x) = 0$ for $|x| \geq 1$,

(ii) $\varrho(x) \geq 0$,

(iii) $\varrho(x) = \varrho(-x)$,

(iv) $\int_{-1}^{1} \varrho(x)dx = 1.$

We define the function δ_n by $\delta_n(x) = n\varrho(nx)$ for $n = 1,2,\ldots$.
It is obvious that $\{\delta_n\}$ is a regular sequence converging to the Dirac delta-function δ.
We now define the ordinary summable function x_+^{λ} for $\lambda > -1$ by

$$x_+^{\lambda} = \begin{cases} x^{\lambda}, & x > 0, \\ 0, & x < 0. \end{cases}$$

The distribution x_+^{λ} is then defined inductively by

$$x_+^\lambda = (\lambda+1)^{-1}(x_+^{\lambda+1}),$$

for $\lambda < -1$ and $\lambda \neq -2, -3, \ldots$ and the distributions x_-^λ and $|x|^\lambda$ are defined by

$$x_-^\lambda = (-x)_+^\lambda, \quad |x|^\lambda = x_+^\lambda + x_-^\lambda$$

for $\lambda \neq -1, -2, \ldots$.
We now give the following definition.

Definition: Let F be a distribution and let f be a summable function. We say that the distribution F(f(x)) exists and is equal to h on the open interval (a,b) if

$$\underset{n\to\infty}{\text{N-lim}} \int_{-\infty}^{\infty} F_n(f(x))\varphi(x)dx = (h(x),\varphi(x))$$

for all test functions φ with compact support contained in (a,b), where

$$F_n(x) = F(x) * \delta_n(x)$$

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

This definition was considered in /6/ for the case where f is an infinitely differentiable function.
We now prove the following theorem.

Theorem 1: The distribution $(|x|^\mu)_-^\lambda$ exists and

$$(|x|^\mu)_-^\lambda = 0$$

for $\mu > 0$, $\lambda \neq -1, -2, \ldots$ and $\lambda\mu \neq -1, -3, -5, \ldots$.

Proof: We have

$$x_-^\lambda = \frac{(-1)^s\,\Gamma(\lambda+1)}{\Gamma(\lambda+s+1)}\,\frac{d^s}{dx^s}\,x_-^{\lambda+s}$$

where s is a non-negative integer chosen so that $\lambda + s > -1$. Then $x_-^{\lambda+s}$ is a summable function and

$$(x_-^{\lambda})_n = x_-^{\lambda} * \delta_n(x)$$

$$= \frac{(-1)^s \Gamma(\lambda+1)}{\Gamma(\lambda+s+1)} \int_{-1/n}^{1/n} (t-x)_-^{\lambda+s} \delta_n^{(s)}(t)\,dt.$$

Thus

$$\frac{(-1)^s \Gamma(\lambda+s+1)}{\Gamma(\lambda+1)} (x_-^{\lambda})_n$$

$$= \begin{cases} \displaystyle\int_{-1/n}^{1/n} (t-x)^{s+\lambda} \delta_n^{(s)}(t)\,dt, & x \le -n^{-1}, \\ \displaystyle\int_{x}^{1/n} (t-x)^{s+\lambda} \delta_n^{(s)}(t)\,dt, & |x| < n^{-1}, \\ 0, & x \ge n^{-1} \end{cases}$$

and so

$$\frac{(-1)^s \Gamma(\lambda+s+1)}{\Gamma(\lambda+1)} ((|x|^{\mu})_-^{\lambda})_n$$

$$= \begin{cases} \displaystyle\int_{|x|^{\mu}}^{1/n} (t-|x|^{\mu})^{s+\lambda} \delta_n^{(s)}(t)\,dt, & |x|^{\mu} < n^{-1}, \\ 0, & |x|^{\mu} \ge n^{-1}. \end{cases}$$

The support of $((|x|^{\mu})_-^{\lambda})_n$ is therefore contained in the interval $(-n^{-1/\mu}, n^{-1/\mu})$.

Since $((|x|^{\mu})_-^{\lambda})_n$ is an even function we have

$$\int_{-n^{-1/\mu}}^{n^{-1/\mu}} ((|x|^{\mu})_-^{\lambda})_n x^i\,dx = 0 \qquad (1)$$

for i = 1,3,5, ... and

$$\frac{(-1)^s\,\Gamma(\lambda+s+1)}{\Gamma(\lambda+1)}\int_{-n^{-1/\mu}}^{n^{-1/\mu}}((|x|^\mu)_-^\lambda)_n x^i\,dx$$

$$= 2\int_0^{n^{-1/\mu}} x^i \int_{|x|^\mu}^{1/n}(t-|x|^\mu)^{s+\lambda}\delta_n^{(s)}(t)\,dt\,dx$$

$$= 2\int_0^{1/n}\delta_n^{(s)}(t)\int_0^{t^{1/\mu}} x^i(t-x^\mu)^{s+\lambda}\,dx\,dt$$

$$= \frac{2}{\mu}\int_0^{1/n} t^{s+\lambda+(i+1)/\mu}\delta_n^{(s)}(t)\int_0^1 y^{-1(i+1)/\mu}(1-y)^{s+\lambda}\,dy\,dt$$

$$= \frac{2B((i+1)/\mu, s+\lambda+1)}{\mu}\,n^{-\lambda-(i+1)/\mu}\int_0^1 u^{s+\lambda+(i+1)/\mu}\varrho^{(s)}(u)\,du \qquad (2)$$

for i = 0,2,4,..., where the substitutions $x^\mu = ty$ and $nt = u$ have been made and B denotes the beta function. It follows that

$$\underset{n\to\infty}{\text{N-lim}}\int_{-n^{-1/\mu}}^{n^{-1/\mu}}((|x|^\mu)_-^\lambda)_n x^i\,dx = 0$$

for i = 0,1,2, ..., $\mu > 0$, $\lambda \neq -1, -2, \ldots$ and $\lambda\mu \neq -1,-3,-5,\ldots$. It also follows that if k is an integer chosen so that $\lambda + (k+1)/\mu > 0$ then for arbitrary $\epsilon > 0$

$$\int_{-n^{-1/\mu}}^{n^{-1/\mu}}|((|x|^\mu)_-^\lambda)_n x^k|\,dx = O(n^{-\lambda-(k+1)/\mu}) \qquad (3)$$
$$< \epsilon$$

for large enough n.

Now let φ be an arbitrary test function with compact support. Then by Taylor's theorem

$$\varphi(x) = \sum_{i=0}^{k-1}\frac{x^i}{i!}\varphi^{(i)}(0) + \frac{x^k}{k!}\varphi^{(k)}(\xi x)$$

where $0 < \xi < 1$. It follows from what we have just proved that

$$\left| \int_{-n^{-1/\mu}}^{n^{-1/\mu}} ((|x|^{\mu})_{-}^{\lambda})_n x^k \varphi^{(k)}(\xi x)\,dx \right|$$

$$\leq \sup_x \{|\varphi^{(k)}(x)|\} \cdot \int_{-n^{-1/\mu}}^{n^{-1/\mu}} |((|x|^{\mu})_{-}^{\lambda})_n x^k|\,dx \to 0$$

as n tends to infinity and so

$$\underset{n\to\infty}{\text{N-lim}}\ ((((|x|^{\mu})_{-}^{\lambda})_n, \varphi(x)) = \underset{n\to\infty}{\text{N-lim}} \int_{-n^{-1/\mu}}^{n^{-1/\mu}} ((|x|^{\mu})_{-}^{\lambda})_n \varphi(x)\,dx$$

$$= \underset{n\to\infty}{\text{N-lim}} \sum_{i=0}^{k-1} \frac{\varphi^{(i)}(0)}{i!} \int_{-n^{-1/\mu}}^{n^{-1/\mu}} ((|x|^{\mu})_{-}^{\lambda})_n x^i\,dx$$

$$+ \lim_{n\to\infty} \frac{1}{k!} \int_{-n^{-1/\mu}}^{n^{-1/\mu}} ((|x|^{\mu})_{-}^{\lambda})_n x^k \varphi^{(k)}(\xi x)\,dx$$

$$= 0 = (0, \varphi(x)).$$

We have therefore proved that $(|x|^{\mu})_{-}^{\lambda} = 0$ for $\mu > 0$, $\lambda \neq -1, -2, \ldots$ and $\lambda\mu \neq -1, -3, -5, \ldots$. This completes the proof of the theorem.

Corollary: The distribution $(-|x|^{\mu})_{+}^{\lambda}$ exists and

$$(-|x|^{\mu})_{+}^{\lambda} = 0$$

for $\mu > 0$, $\lambda \neq -1, -2, \ldots$ and $\lambda\mu \neq -1, -3, -5, \ldots$.

Proof: The result follows on noting that $(-x)_{+}^{\lambda} = x_{-}^{\lambda}$ and so $(-|x|^{\mu})_{+}^{\lambda} = (|x|^{\mu})_{-}^{\lambda} = 0$.

Theorem 2: The distribution $(|x|^{\mu})_{-}^{\lambda}$ exists and

$$(|x|^{\mu})_{-}^{\lambda} = -\frac{\pi \operatorname{cosec}(\pi\lambda)}{\mu(-\lambda\mu-1)!}\, \delta^{(-\lambda\mu-1)}(x)$$

for $\mu > 0$, $\lambda \neq -1, -2, \ldots$ and $\lambda\mu = -1, -3, -5, \ldots$.

Proof: With the notation used in the proof of Theorem 1 we note that (1), (2) and (3) still hold. However, when i is the non-negative even integer $-\lambda\mu - 1 = k - 1$, we have from (2)

$$\frac{(-1)^s \Gamma(\lambda+s+1)}{\Gamma(\lambda+1)} \int_{-n^{-1/\mu}}^{n^{-1/\mu}} ((|x|^\mu)_-^\lambda)_n x^{k-1} dx$$

$$= \frac{2B(-\lambda, s+\lambda+1)}{\mu} \int_0^1 u^s \varrho^{(s)}(u)\, du$$

$$= \frac{2B(-\lambda, s+\lambda+1)}{\mu} \cdot \frac{1}{2}(-1)^s s!$$

$$= \frac{(-1)^s \Gamma(-\lambda)\, \Gamma(s+\lambda+1)}{\mu}$$

and so

$$\int_{-n^{-1/\mu}}^{n^{-1/\mu}} ((|x|^\mu)_-^\lambda)_n x^{k-1} dx = \frac{\Gamma(\lambda+1)\, \Gamma(-\lambda)}{\mu}$$

$$= -\frac{\pi \operatorname{cosec}(\pi\lambda)}{\mu} .$$

Now let φ be an arbitrary test function with compact support. Then it follows as above that

$$\underset{n\to\infty}{\text{N-lim}} \left(((|x|^\mu)_-^\lambda)_n, \varphi(x)\right) = \underset{n\to\infty}{\text{N-lim}} \int_{-n^{-1/\mu}}^{n^{-1/\mu}} ((|x|^\mu)_-^\lambda)_n \varphi(x)\, dx$$

$$= \underset{n\to\infty}{\text{N-lim}} \sum_{i=0}^{k-1} \frac{\varphi^{(i)}(0)}{i!} \int_{-n^{-1/\mu}}^{n^{-1/\mu}} ((|x|^\mu)_-^\lambda)_n x^i dx$$

$$+ \lim_{n\to\infty} \frac{1}{k!} \int_{-n^{-1/\mu}}^{n^{-1/\mu}} ((|x|^\mu)_-^\lambda)_n x^k \varphi^{(k)}(\xi x)\, dx$$

$$= -\frac{\varphi^{(k-1)}(0)}{(k-1)!} \cdot \frac{\pi \operatorname{cosec}(\pi\lambda)}{\mu}$$

$$= \frac{(-1)^k \pi \operatorname{cosec}(\pi\lambda)}{\mu(k-1)!} \left(\delta^{(k-1)}(x), \varphi(x)\right)$$

$$= - \frac{\pi \operatorname{cosec}(\pi\lambda)}{\mu(-\lambda\mu-1)!} \left(\delta^{(-\lambda\mu-1)}(x), \varphi(x)\right)$$

and so

$$(|x|^{\mu})_{-}^{\lambda} = - \frac{\pi \operatorname{cosec}(\pi\lambda)}{\mu(-\lambda\mu-1)!} \delta^{(-\lambda\mu-1)}(x)$$

for $\mu > 0$, $\lambda \neq -1, -2, \ldots$ and $\lambda\mu = -1, -3, -5, \ldots$. This completes the proof of the theorem.

Corollary: The distribution $(-|x|^{\mu})_{+}^{\lambda}$ exists and

$$(-|x|^{\mu})_{+}^{\lambda} = - \frac{\pi \operatorname{cosec}(\pi\lambda)}{\mu(-\lambda\mu-1)!} \delta^{(-\lambda\mu-1)}(x)$$

for $\mu > 0$, $\lambda \neq -1, -2, \ldots$ and $\lambda\mu = -1, -3, -5, \ldots$.

Proof: Since $(-x)_{+}^{\lambda} = x_{-}^{\lambda}$ we have

$$(-|x|^{\mu})_{+}^{\lambda} = (|x|^{\mu})_{-}^{\lambda} = - \frac{\pi \operatorname{cosec}(\pi\lambda)}{\mu(-\lambda\mu-1)!} \delta^{(-\lambda\mu-1)}(x).$$

Theorem 3: The distribution $(x_{-}^{\mu} + x_{+}^{\nu})_{-}^{\lambda}$ exists and

$$(x_{-}^{\mu} + x_{+}^{\nu})_{-}^{\lambda} = 0$$

for $\mu, \nu > 0$ and $\lambda, \lambda\mu, \lambda\nu \neq -1, -2, \ldots$,

$$(x_{-}^{\mu} + x_{+}^{\nu})_{-}^{\lambda} = - \frac{\pi \operatorname{cosec}(\pi\lambda)}{2\mu(-\lambda\mu-1)!} \delta^{(-\lambda\mu-1)}(x)$$

for $\mu, \nu > 0$, $\lambda, \lambda\nu \neq -1, -2, \ldots$ and $\lambda\mu = -1, -2, \ldots$,

$$(x_{-}^{\mu} + x_{+}^{\nu})_{-}^{\lambda} = \frac{(-1)^{\lambda\nu} \pi \operatorname{cosec}(\pi\lambda)}{2\nu(-\lambda\nu-1)!} \delta^{(-\lambda\nu-1)}(x)$$

for $\mu, \nu > 0$, $\lambda, \lambda\mu \neq -1, -2, \ldots$ and $\lambda\nu = -1, -2, \ldots$ and

$$(x_{-}^{\mu} + x_{+}^{\nu})_{-}^{\lambda} = - \frac{\pi \operatorname{cosec}(\pi\lambda)}{2\mu(-\lambda\mu-1)!} \delta^{(-\lambda\mu-1)}(x)$$

$$+ \frac{(-1)^{\lambda\nu} \pi \operatorname{cosec}(\pi\lambda)}{2\nu(-\lambda\nu-1)!} \delta^{(-\lambda\nu-1)}(x)$$

for $\mu, \nu > 0$, $\lambda \neq -1, -2, \ldots$ and $\lambda\mu, \lambda\nu = -1, -2, \ldots$.

Proof: The results follow on noting that

$$\frac{(-1)^s \Gamma(\lambda+s+1)}{\Gamma(\lambda+1)} ((x_-^\mu + x_+^\nu)_-^\lambda)_n =$$

$$= \begin{cases} \int_{|x|^\mu}^{1/n} (t - |x|^\mu)^{s+\lambda} \delta_n^{(s)}(t)\,dt, & -n^{-1/\mu} \le x \le 0, \\ \int_{x^\nu}^{1/n} (t - x^\nu)^{s+\lambda} \delta_n^{(s)}(t)\,dt, & 0 \le x \le n^{-1/\nu}, \\ 0, & x < -n^{-1/\mu}, \quad x > n^{-1/\nu} \end{cases}$$

so that

$$(-1)^s \Gamma(\lambda+s+1) \int_{-n^{-1/\mu}}^{n^{-1/\nu}} ((x_-^\mu + x_+^\nu)_-^\lambda)_n x^i dx =$$

$$= \frac{(-1)^i B((i+1)/\mu, s+\lambda+1)}{\mu} n^{-\lambda-(i+1)/\mu} \int_0^1 u^{s+\lambda+(i+1)/\mu} \varrho^{(s)}(u)\,du$$

$$+ \frac{B((i+1)/\nu, s+\lambda+1)}{\nu} n^{-\lambda-(i+1)/\nu} \int_0^1 u^{s+\lambda+(i+1)/\nu} \varrho^{(s)}(u)\,du$$

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

Corollary: The distribution $(-x_-^\mu - x_+^\nu)_+^\lambda$ exists and

$$(-x_-^\mu - x_+^\nu)_+^\lambda = 0$$

for $u, \nu > 0$ and $\lambda, \lambda\mu, \lambda\nu \neq -1, -2, \ldots$,

$$(-x_-^\mu - x_+^\nu)_+^\lambda = \frac{\pi \operatorname{cosec}(\pi\lambda)}{2\mu\,(-\lambda\mu-1)!} \delta^{(-\lambda\mu-1)}(x)$$

for $\mu, \nu > 0$, λ, $\lambda\nu \neq -1, -2, \ldots$ and $\lambda\mu = -1, -2, \ldots$,

$$(-x_-^\mu - x_+^\nu)_+^\lambda = \frac{(-1)^{\lambda\nu}\ \pi \operatorname{cosec}(\pi\lambda)}{2\nu(-\lambda\nu-1)!} \delta^{(-\lambda\nu-1)}(x)$$

for $\mu, \nu > 0$, $\lambda, \lambda\mu \neq -1, -2, \ldots$ and $\lambda\nu = -1, -2, \ldots$ and

$$(-x_-^\mu - x_+^\nu)_+^\lambda = -\frac{\pi \operatorname{cosec}(\pi\lambda)}{2\mu(-\lambda\mu-1)!}\,\delta^{(-\lambda\mu-1)}(x)$$

$$+\frac{(-1)^{\lambda\nu}\,\pi \operatorname{cosec}(\pi\lambda)}{2\nu(-\lambda\nu-1)!}\,\delta^{(-\lambda\nu-1)}(x)$$

for $\mu, \nu > 0$, $\lambda \neq -1, -2, \ldots$ and $\lambda\mu, \lambda\nu = -1, -2, \ldots$.

Theorem 4: The distributions $(x_-^\mu)_-^\lambda$ and $(x_+^\mu)_-^\lambda$ exist and

$$(x_-^\mu)_-^\lambda = (x_+^\mu)_-^\lambda = 0$$

for $\mu > 0$ and λ, $\lambda\mu = -1, -2, \ldots$ and

$$(x_-^\mu)_-^\lambda = (-1)^{\lambda\mu}(x_+^\mu)_-^\lambda = \frac{\pi \operatorname{cosec}(\pi\lambda)}{2\mu(-\lambda\mu-1)!}\,\delta^{(-\lambda\mu-1)}(x)$$

for $\mu > 0$ $\lambda \neq -1, -2, \ldots$ and $\lambda\mu = -1, -2, \ldots$.

Proof: The results follow on noting that

$$\frac{(-1)^s\,\Gamma(\lambda+s+1)}{\Gamma(\lambda+1)}\,((x_+^\mu)_-^\lambda)_n =$$

$$= \begin{cases} \displaystyle\int_0^{1/n} t^{s+\lambda}\,\delta_n^{(s)}(t)\,dt, & x \leq 0, \\ \displaystyle\int_{x^\mu}^{1/n} (t - x^\mu)^{s+\lambda}\,\delta_n^{(s)}(t)\,dt, & 0 \leq x \leq n^{-1/\mu}, \\ 0, & x > n^{-1/\mu}, \end{cases}$$

so that

$$\frac{(-1)^s\,\Gamma(\lambda+s+1)}{\Gamma(\lambda+1)} \int_0^{n^{-1/\mu}} ((x_+^\mu)_-^\lambda)_n x^i\,dx$$

$$= \frac{B((i+1)/\mu, s+\lambda+1)}{\mu}\, n^{-\lambda-(i+1)/\mu} \int_0^1 u^{s+\lambda+(i+1)/\mu}\,\varphi^{(s)}(u)\,du$$

for $i = 0,1,2,\ldots$ and for any test function φ and $a < 0$

$$\frac{(-1)^s\,\Gamma(\lambda+s+1)}{\Gamma(\lambda+1)}\int_a^0((x_+^\mu)_-^\lambda)_n\varphi(x)\,dx$$

$$= n^{-\lambda}\int_0^1 u^{s+\lambda}\varrho^{(s)}(u)\,du\cdot\int_a^0\varphi(x)\,dx.$$

Corollary: The distributions $(-x_-^\mu)_+^\lambda$ and $(-x_+^\mu)_+^\lambda$ exist and

$$(-x_-^\mu)_+^\lambda = (-x_+^\mu)_+^\lambda = 0$$

for $\mu > 0$ and $\lambda,\ \lambda\mu \neq -1, -2, \ldots$ and

$$(-x_-^\mu)_+^\lambda = (-1)^{\lambda\mu}(-x_+^\mu)_+^\lambda = \frac{\pi\,\mathrm{cosec}(\pi\lambda)}{2\mu\,(-\lambda\mu-1)!}\,\delta^{(-\lambda\mu-1)}(x)$$

for $\mu > 0$, $\lambda \neq -1, -2, \ldots$ and $\lambda\mu = -1, -2, \ldots$.

Theorem 5: Let f be a summable function which is continuous and positive for $x \neq 0$ and equal to x^μ for $x > 0$. Then the distribution $(f(x))_-^\lambda$ exists and

$$(f(x))_-^\lambda = 0$$

for $\mu > 0$ and $\lambda,\ \lambda\mu \neq -1, -2, \ldots$ and

$$(f(x))_-^\lambda = \frac{(-1)^{\lambda\mu}\,\pi\,\mathrm{cosec}(\pi\lambda)}{2\mu\,(-\lambda\mu-1)!}\,\delta^{(-\lambda\mu-1)}(x)$$

for $\mu > 0$, $\lambda \neq -1, -2, \ldots$ and $\lambda\mu = -1, -2, \ldots$.

Proof: The results follow on noting that

$$\frac{(-1)^s\,\Gamma(\lambda+s+1)}{\Gamma(\lambda+1)}\,((f(x))_-^\lambda)_n =$$

$$= \begin{cases} \displaystyle\int_{x^\mu}^{1/n}(t-x^\mu)^{s+\lambda}\delta_n^{(s)}(t)\,dt, & 0 \leq x \leq n^{-1/\mu}, \\ 0, & x > n^{-1/\mu},\ x < 0,\ f(x) > n^{-1}. \end{cases}$$

Corollary: Let f be as in the theorem. Then the distribution $(-f(x))_+^\lambda$ exists and

$$(-f(x))_+^\lambda = 0$$

for $\mu > 0$ and λ, $\lambda\mu \neq -1, -2, \ldots$ and

$$(-f(x))_+^\lambda = \frac{(-1)^{\lambda\mu}\pi\operatorname{cosec}(\pi\lambda)}{2\mu(-\lambda\mu-1)!}\,\delta^{(-\lambda\mu-1)}(x)$$

for $\mu > 0$, $\lambda \neq -1, -2, \ldots$ and $\lambda\mu = -1, -2, \ldots$.

Similarly we have

Theorem 6: Let f be a summable function which is continuous and positive for $x \geq 0$ and equal to $|x|^\mu$ for $x < 0$. Then the distribution $(f(x))_-^\lambda$ exists and

$$(f(x))_-^\lambda = 0$$

for $\mu > 0$ and λ, $\lambda\mu \neq -1, -2, \ldots$ and

$$(f(x))_-^\lambda = -\frac{\pi\operatorname{cosec}(\pi\lambda)}{2\mu(-\lambda\mu-1)!}\,\delta^{(-\lambda\mu-1)}(x)$$

for $\mu > 0$, $\lambda \neq -1, -2, \ldots$ and $\lambda\mu = -1, -2, \ldots$.

Corollary: Let f be as in the theorem. Then the distribution $(-f(x))_+^\lambda$ exists and

$$(-f(x))_+^\lambda = 0$$

for $\mu > 0$ and λ, $\lambda\mu \neq -1, -2, \ldots$ and

$$(-f(x))_+^\lambda = -\frac{\pi\operatorname{cosec}(\pi\lambda)}{2\mu(-\lambda\mu-1)!}\,\delta^{(-\lambda\mu-1)}(x)$$

for $\mu > 0$, $\lambda \neq -1, -2, \ldots$ and $\lambda\mu = -1, -2, \ldots$.

For further related results see /2/, /3/, /4/ and /5/.

References

/1/ van der Corput, J. G.: Introduction to the neutrix calculus. J. Analyse Math. 7, 291 - 398 (1959)

/2/ Fisher, B.: On defining the distribution $\delta^{(r)}(f(x))$. Rostock. Math. Kolloq. 23, 73 - 80 (1983)

/3/ Fisher, B.: On defining the distribution $\delta^{(r)}(f(x))$ for summable f. Publ. Math. Debrecen, (to appear)

/4/ Fisher, B., and Itano, M.: Some results on distributions and the change of variable. Mem. Fac. Int. Arts Sci. Hiroshima Univ. (to appear)

/5/ Fisher, B., and Kuribayashi, Y.: On defining the distribution $(x^r)_-^{-s}$. J. Fac. Educ. Tottori Univ., Nat. Sci. (to appear)

/6/ Fisher, B., and Kuribayashi, Y.: Changing the variable in distributions. Demonstratio Math. 17, 499-514 (1984)

received: November 28, 1984

Author's address:

Dr. B. Fisher
Department of Mathematics
The University
Leicester
LE 1 7RH
England

Rostock. Math. Kolloq. 28, 87 - 98 (1985) 68G10
68C25

Hoang Kiem

Geometric Transforms of Digital Images

1. Introduction

The problem of geometric transforms for digital images can be stated as follows.
A given digital image f(x,y) is processed by a spatial distortion block. The resulting image g is a function of the coordinates (u,v) which are related to the original coordinates (x,y) through a transform T,

$u = T_u(x,y)$ and

$v = T_v(x,y)$.

The realisation of this process may take place in the following two steps:
1) The coordinates (u,v) of the observed pixel (u,v,g(u,v)), into which a new image pixel ((x,y,f(x,y)) maps, are computed.

2) The gray (or color) level g(u,v) of such a pixel is computed.

The analytic description of the transform T may be obtained either a priori (i.e. based on some general assumptions) or by means of reference points (the control grid points).
It should be noted that the coordinates (u,v) resulting from step (1) do not fall, in general, on the sampling grid. Therefore, the gray levels of the grid points must be interpolated. However, a considerable amount of computation is needed to this approach - the indirect approach. We also shall present algorithms for solving this problem with the direct approach, i.e. by the direct mapping of grid points from one grid to the other.

2. The indirect approach for geometric transforms

2.1. The global model

Suppose that there is a spatial relationship between the two reference frames given by

$u = T_u(x,y)$ and

$v = T_v(x,y)$.

We shall approximate T_u and T_v by polynomials in x and y of the form:

$$u = \sum_{i=o}^{N} \sum_{j=o}^{N} k_{ij}^{u} x^i y^j \text{ and}$$

$$v = \sum_{i=o}^{N} \sum_{j=o}^{N} k_{ij}^{v} x^i y^j$$

where k_{ij}^{u}, k_{ij}^{v} are the constant polynomial coefficients. Now suppose that a set of control points is known, i.e.

$$\{(x_i,y_i), (u_i,v_i) : i = 1,2,\ldots,M\}.$$

Using the relationship between u and (x,y), the following matrix equation may be developed

$$\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_M \end{bmatrix} = \begin{bmatrix} Y_1' & x_1Y_1' & x_1^2Y_1' & \cdots & x_1^NY_1' \\ Y_2' & x_2Y_2' & x_2^2Y_2' & \cdots & x_2^NY_2' \\ \vdots & \vdots & \vdots & & \vdots \\ Y_M' & x_MY_M' & x_M^2Y_M' & \cdots & x_M^NY_M' \end{bmatrix} \cdot \begin{bmatrix} K_o^u \\ K_1^u \\ \\ K_N^u \end{bmatrix}$$

where $Y_i' = (1, y_i, y_i^2, \ldots, y_i^N)$ and

$$K_i^u = (k_{i0}^u, k_{i1}^u, \ldots, k_{iN}^u).$$

The equation may be written simple as $U = T_u \cdot K_u$.

A similar equation may be developed for the v coordinates,

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_M \end{bmatrix} = \begin{bmatrix} X_1' & y_1X_1' & y_1^2X_1' & \cdots & y_1^NX_1' \\ X_2' & y_2X_2' & y_2^2X_2' & \cdots & y_2^NX_2' \\ \vdots & \vdots & \vdots & & \vdots \\ X_M' & y_MX_M' & y_M^2X_M' & \cdots & y_M^NX_M' \end{bmatrix} \cdot \begin{bmatrix} K_o^v \\ K_1^v \\ \\ K_N^v \end{bmatrix}$$

where $X_i' = (1, x_i, x_i^2, \ldots, x_i^N)$ and

$$K_i^v = (k_{io}^v, k_{i1}^v, \ldots, k_{iN}^v).$$

This equation also may be written simply as $V = T_v \cdot K_v$.

Since there are $(N+1)^2$ unknown coefficients k_{ij}, the number of control points M must be greater than or equal to $(N+1)^2$. For the case $M \geq (N+1)^2$, we may use least squares or minimum mean squared error as a criterion to provide a best estimate for coefficients k_{ij} ([1,2]).

Alternatively, an interactive procedure can be used: elementary polynomial transformations are sequentially applied and their parameters are modified according to a subjective criterion. In the vast majority of application, often the transformation can be modeled as a polynomial of second order or first order (affine transformation).

2.2. The local model

In effect, this model is based on piecewise linear approximations to the unknown distortion. If the number M of control points is small, this approximation may be performed by the use of interactively selected control points, by triangulating a plan set ([1]), by using the Gabriel graph, or the Voronoi diagram ([4]). However, for large values of M, a certain clustering analysis for control points is necessary. In the sequel, the method and algorithms for solving this problem are briefly reviewed.

2.2.1. The basic idea of the method.

Generally speaking, the aim of a clustering process is to be able to insert the objects of a given set X into k clusters. An important variant of a clustering process is characterized by the principle Dynamic Cluster Method ([3]).

Suppose that $\mathrm{I\!P}_k$ is the set of all partitions of a certain set of objects of X in k classes (clusters), $\mathrm{I\!L}_k$ is a set of

vectors $L = (A_1, A_2, \ldots, A_k)$, and G is a mapping from $\mathbb{P}_k$ into $\mathbb{L}_k$, i.e. for all $P \in \mathbb{P}_k$,

$$G(P) = L = (A_1, A_2, \ldots, A_k).$$

A mapping $\mathcal{E}$ from L_k into P_k is introduced which will be used for the identification of a "feature vector" $(A_1, A_2, \ldots, A_r)$ with a certain partition of X.
Let D be a dissemble measure $D : X \times \mathbb{L} \rightarrow R^+$.
Then $\mathcal{E}(L) = P = (P_1, P_2, \ldots, P_k)$ with

$P_i = \{x \in X \mid D(x, A_i) \leq D(x, A_j)$, for all $j \neq i\}$, where $1 \leq i \leq k$.

The basic idea of the method is to use successively the function G and $\mathcal{E}$ as defined until a certain stable state is obtained.

2.2.2. A dynamic cluster algorithm for clustering the control point set C

Let $C \subseteq X$ be a set of control points in the sense of the geometric transform problem.
Step 1: Let k elements of C be chosen at random to be the "representation" of the k clusters. Let us denote them by $A_1, A_2, \ldots, A_k$.
Step 2: For all i, any element c_i of C is assigned to class j iff $d(c_i, A_j)$ is minimum, for $1 \leq j \leq k$.
Step 3: For all j, a new mean A_j of class j is computed. $(A_1, A_2, \ldots, A_r)$ is the new cluster representation.
Step 4: If no A_j has changed - stop, else goto step 2.

After using the above algorithm, the local approximation T_i for each class A_i is determined and the process of the approximation for the non-control points is as follows:
The coordinates of the non-control point P are computed by transform T_i iff $D(P, A_j) \leq D(P, A_i)$, for all $j \neq i$, where $D(P, A_i)$ can be chosen as the Euclid distance from P to mean A_i, see Fig. 1.

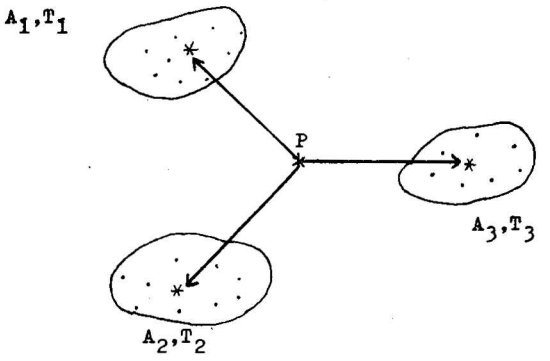


Fig. 1. The coordinates of P are computed with T_3 because $D(P,A_3) < D(P,A_1)$ and

$$D(P,A_3) < D(P,A_2)$$

2.2.3. A dynamic cluster algorithm for determining local transformations

Step 1: Let k classes of C (the control points) be chosen at random. Let us denote them by $A_1, A_2, \ldots, A_k$.

Step 2: The transforms T_i for class A_i are determined, $1 \leq i \leq k$.

Step 3: For all i, any c_i of C is assigned to A_j iff $d(c_i,A_j)$ is minimum, for $1 \leq j \leq k$, where $d(c_i,A_j)$ can be determined as error of A_j for c_i using the transform T_j.

Step 4: If no A_j has changed - stop, else go to step 2.

2.3. The interpolating function

Regarding the procedures outlined above, it should be noted that coordinates resulting from polynomial transforms T do not fall, in general on the sampling grid. Therefore, the gray levels g(u,v) of the points (u,v) have to be interpolated in some cases,

$$g(u,v) = \sum_m \sum_n g(m,n) \cdot R(u-m, v-n)$$

where m and n span the desired grid points around (u,v) and R is an interpolating function. For example, we can use linear interpolation or nearest-neighbour interpolation.

2.3.1. The linear interpolation ([1,2])

Suppose that (u,v) is surrounded by the four digital points (u_1,v_1), (u_1+1,v_1), (u_1,v_1+1), (u_1+1,v_1+1), i.e. $u_1 \leq u \leq u_1 + 1$ and $v_1 \leq v \leq v_1 + 1$.

Determine the gray level of (u,v) in g by

$$g(u,v) = (1-\alpha)(1-\beta)g(u_1,v_1) + \alpha(1-\beta)g(u_1+1,v_1) + (1-\alpha)\beta \cdot g(m,v_1+1) + \alpha \cdot \beta g(u_1+1,v_1+1)$$

where $\alpha = u-u_1$ and $\beta = v-v_1$.

2.3.2. The nearest-neighbour interpolation ([1,2,5])

Suppose that [u] denotes the integer closest to u. Determine the gray level of (u,v) in g by copying the value g([u],[v]). It is clear that ([u],[v]) is the nearest grid point for (u,v), see Fig. 2.

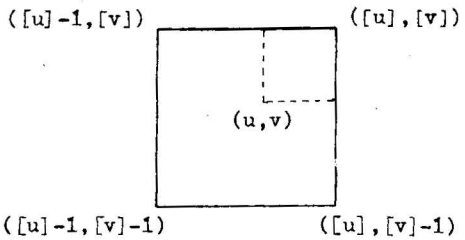


Fig. 2. Nearest-neighbour mapping

It is well known that the nearest-neighbour interpolation produces false-contour effects for high enlargement factors. When the enlargement factor is less then 1 (i.e. when the image is reduced), very good results are obtained ([5]).

2.3.3. A Criterion for selecting the interpolating function

Suppose that we have a set of control points (x_i,y_i), (u_i,v_i), $i = 1,2,\ldots,M$, and transforms T_u, T_v. The mean error of the interpolating function F is defined as follows:

$$\mathcal{E}(F) = \frac{1}{M} \cdot \sum_{i=1}^{M} \left\{ g\left[F(T_u(x_i,y_i)),F(T_v(x_i,y_i))\right] - g(u_i,v_i) \right\}^2 .$$

The interpolating function F is chosen from the list $(F_1,F_2,\ldots,F_k)$ iff

$$\mathcal{E}(F) = \min_{i=1,2,\ldots k} \mathcal{E}(F_i) .$$

For the nearest-neighbour interpolating function, the mean error can be defined as follows:

$$\mathcal{E}([\,]) = \frac{1}{2M} \cdot \sum_{i=1}^{M} \left\{ \text{Sign} \left| [T_u(x_i,y_i)] - u_i \right| + \text{Sign} \left| [T_v(x_i,y_i)] - v_i \right| \right\}.$$

The nearest-neighbour interpolating function is chosen if

$$\mathcal{E}([\,]) \leq \mathcal{E}^o,$$

where $\mathcal{E}^o$ denotes a threshold defined by users.

3. The direct approach for geometrical correction

In many cases, one can use the affine transform T for geometrical correction ([5]). Then, computation of T can be omitted by providing a suitable algorithm and a direct mapping of one grid onto the other. It requires only simple and fast integer arithmetic operations. This algorithm is based on the following theoretical result.

(3.1) Proposition: The affine transformation can be implemented by applying two different scale factors (r,r′) and rotation angles (θ, θ') to the two space directions.

Demonstration: In fact, the equations defining the affine transformation

$$u = a_o + a_1 x + a_2 y$$

$$v = b_o + b_1 x + b_2 y$$

can be written as

$$u = a_o + r(x \cdot \cos \Theta + y \cdot \sin \Theta)$$

$$v = b_o + r(-x \cdot \sin \Theta' + y \cdot \cos \Theta')$$

where

$$a_1 = r \cdot \cos \Theta \,, \quad a_2 = r \cdot \sin \Theta \,,$$

$$b_1 = -r \cdot \sin \Theta \,, \quad b_2 = r \cdot \cos \Theta' .$$

From here, it is clear that affine transformations can be realized by three simple operations: shifts, rotations and scale changes.

(3.2) Proposition: The affine transformations can be performed by using an algorithm for digital straight line computation.

Demonstration: The digital representation of straight lines is a classical topic in computer graphics. Many algorithms have been developed to produce such representations.

Let us consider a straight line l described by equation $y = (m/n)x + q$ where m and n are integers with no common factor, i.e. $\gcd(m,n) = 1$. For simplicity, suppose that the slope of line l lying between 0 and $\pi/4$, i.e. $0 \leq n \leq m$ and that $q = 0$. The point (j,k) with j,k integers belongs to the digital representation of line l iff $k = (n/m)j + \varepsilon$ with $|\varepsilon| \leq 1/2$. The scale factor problem can be restated as the selection or insertion of r equally spaced pixels from among the m pixels of an image line, and therefore it corresponds to the problem of generating a digital straight line, too. It should be noted that a digital image can be rotated by simply using the set of parallel strips with the desired rotation angle of the given image, i.e. our algorithm is based on the computation of straight lines $y = (n/m)x$ on a discrete plane, where $\frac{n}{m} = \operatorname{tg} \Theta$. Thus, the affine transform can be performed by using only the algorithm for the computation of digital straight lines.

(3.3) The algorithm generating a digital straight line
$k = (n/m)j$
This algorithm selects the points (j,k) with j,k integers belonging to the digital representation of line l, i.e.
$k = (n/m)j + \varepsilon$ where $|\varepsilon| \leq 1/2$.
By some algebra, one gets from the above

$$-m \leq 2m \cdot k - 2n \cdot j \leq m.$$

Without loss of generality our considerations are restricted to the first octant since any straight line can be mapped into a first octant line by simple mappings from (j,k) to $(\pm j, \pm k)$.

By means of an incremental procedure, the computation is reduced to very simple operations (such as add and compare), a possible implementation of the algorithm is the following:

```
      SUBROUTINE LINE(N,M)
C     The algorithm generating a digital straight line K = (N/M)J
      INTEGER R,DN,DM
      DATA J,K /0,0/
      DN = 2 * N
      DM = 2 * M
      R = 0
      WRITE ( ... ) J,K
      DO 1 J = 1,M
      R = R + DN
      J = J + 1
      IF(R.LE.M) GOTO 2
      R = R - DM
      K = K + 1
2     WRITE ( ... ) J,K
1     CONTINUE
      END
```

This procedure is the basis for implementing fast affine transforms of digital images, because based on it the rotation and scalar transformation can be implemented by subroutine ROTAS and SCALA as follows:

```
      SUBROUTINE SCALA (L,M)
C     The scala factor problem corresponds to generating
C     a digital straight line K = (L/M)J
C     The scale factor S = L/M
      J = 0
      K = 0
      R = 0
      DL = 2 * L
      DM = 2 * M
      PJ = J
      PK = K
      WRITE ( ...) PJ,PK
      DO 1 J = 1,L
      R = R + DL
      J = J + 1
      IF (R.LE.M) GOTO 2
      R = R - DM
      K = K + 1
2     PJ = J
      PK = K
      WRITE ( ... ) PJ,PK
1     CONTINUE
      END
```

```
      SUBROUTINE ROTAS (N,M)
C     The image rotation algorithm deals with the representation
C     of straight line K = (N/M)J
C     The desired rotation angle T = N/M
      DATA J,K /0,0/
      DN = 2 * N
      DM = 2 * M
      R = 0
      WRITE ( ... ) J,K
      DO 1 J = 1,M
      R = R + DN
      J = J + 1
      WRITE (...) J,K
      IF (R.LE.M) GOTO 1
      R = R - DM
      K = K + 1
      WRITE (...) J,K
1     CONTINUE
      END
```

For preserving the continuity in the straight lines generated by means of the algorithm above, we can use the image smoothing algorithm. It is well known that there exists a class of iterative local image smoothing techniques in which a neighbourhood of each pixel is examined and the pixel is replaced by an average of a selected set of its neighbours chosen so as to make it likely that they belong to the same region as the pixel ([6]). Here, we can use the approach which attempts to choose neighbours that belong to the same histogram peak as the pixel, but are more typical of that peak, i.e. uses those neighbours whose probabilities, as estimated from the histogram are higher than that of the given pixel ([7]).

Acknowledgements
The author wishes to thank the referee of Rostock. Math. Kolloq. for his helpful comments and suggesting references on this paper.

References

/1/ Rosenfeld, A., and Kak, A.: Digital Picture Processing. New York 1976

/2/ Hall, E.: Computer Image Processing and Recognition. New York 1979

/3/ Diday, E., and Simon, J. C.: Clustering analysis. In: Fu, K.S. (Ed.): Digital Pattern Recognition. Berlin 1976, 47 - 94

/4/ Toussaint, G.: Pattern recognition and geometrical complexity. IEEE Trans. Comput. C-75, 1324 - 1345 (1980)

/5/ Bracini, C., et al: Fast geometrical manipulation. Computer Graphics and Image Processing 13, 127 - 141 (1980)

/6/ Jain, A., et al: Image restoration, modelling and reduction of dimensionality. IEEE Trans. Comput. C-23, 470 - 476 (1974)

/7/ Narayana, K. A., et al: Image smoothing by local use of global information. IEEE Trans. Systems Man Cybernet. 12, 826 - 831 (1981)

received: July 17, 1984
revised version October 20, 1984

Author's address:

Dr. sc. nat. Hoang Kiem
Institute of Informatics and Cybernetics
Nghia Do - Tu liem
Ha Noi - Viet Nam

Rostock. Math. Kolloq. 28, 99 - 110 (1985)

94A11
68C05

Reiner Creutzburg
Hans-Jörg Grundmann

Determination of convenient moduli for 16-bit mixed-radix number-theoretic transforms

1. Introduction

With the rapid advances in large scale integration, a growing number of digital signal processing operations becomes attractive. The number-theoretic transform (NTT) was introduced as a generalization of the discrete Fourier transform (DFT) over residue class rings of integers in order to implement fast cyclic convolution and correlation without round-off errors and with better efficiency than the fast Fourier transform (FFT) (cf. /1/, /2/). Other interesting applications of the NTT are in fast digital filtering (/2/), image processing (/3/), fast coding and decoding of error-correcting codes (/4/) and very fast DFT computation (/5/). A large number of transform methods are developed (/1/, /2/, /6/) to relieve some of the length limitations of conventional Fermat-number and Mersenne-number transforms (/1/, /2/, /8/). It is always a hard problem to find convenient moduli that are large enough to avoid overflow and to find primitive N-th roots of unity modulo m with minimal binary weight for transform lengths N that are highly factorizable and large enough for practical applications. In the recent papers /6/, /7/ a useful way was shown to solve this problem by studying cyclotomic polynomials. In this note we determine all convenient prime moduli $p \leq 2^{16} + 1$, so that $\alpha = 2$ is a primitive N-th root of unity modulo p for all mixed radix lengths N in the range $2 \leq N \leq 65536$. A large number of interesting cases for practical applications is obtained.

2. Number-theoretic transforms

Let Z be the ring of integers and $m > 1$ an odd integer with

prime factorization

$$m = p_1^{r_1} p_2^{r_2} \dots p_s^{r_s}. \tag{1}$$

Then $\alpha \in Z$ is called <u>primitive N-th root of unity modulo m</u> if (cf. /7/)

$$\alpha^N \equiv 1 \bmod m, \tag{2}$$

$$\gcd(\alpha^n - 1, m) = 1 \text{ for every } n = 1, \dots, N-1. \tag{3}$$

A necessary and sufficient condition for the existence of such primitive N-th roots of unity modulo m is (/1/)

$$N \mid \gcd(p_1 - 1, \dots, p_s - 1). \tag{4}$$

Note that (3) is always fulfilled if the modulus m is a prime number.

The NTT of length N with α as primitive N-th root of unity modulo m and its inverse are defined between N-point integer sequences

$[x_0, x_1, \dots, x_{N-1}]'$ and $[X_0, X_1, \dots, X_{N-1}]'$ where

$$X_n \equiv \sum_{k=0}^{N-1} x_k \alpha^{nk} \bmod m, \qquad (n = 0, \dots, N-1)$$

$$x_k \equiv N' \sum_{n=0}^{N-1} X_n \alpha^{-nk} \bmod m, \qquad (k = 0, \dots, N-1),$$

and $N \cdot N' \equiv 1 \bmod m$. Note that the components of a signal

$$\underline{x} = [x_0, x_1, \dots, x_{N-1}]'$$

have to be quantized to integers before using the NTT. However, in many practical applications this is already done by sensors with analog/digital conversion.

The NTT has a similar structure and properties like the DFT, particularly the cyclic convolution property (/1/).

From the numerical point of view the following three essential conditions on NTT are required:

- N has to be large enough and highly factorizable in order to implement fast algorithms like prime-factor-, Winograd-, single-radix-, mixed-radix algorithms (/2/),

- α must have a simple binary representation so that the arithmetic modulo m is easy to perform,
- m has to be large enough to avoid overflow but on the other hand small enough, so that the machine word length is not exceeded. Furthermore m should have a simple binary arithmetic.

For instance, the Fermat-number transform with $N = 2^{d+1}$, $d \geq 0$, $\alpha = 2$ and $m = 2^{2^d} + 1$ is a compromise between these various conditions (/8/). The Fermat-number transform uses the well-known FFT- algorithm for radix-2-transform lengths. But for large transform lengths the modulus m exceeds the machine word length. So one has to look for convenient moduli for fixed machine word lengths (for example 8 bit or 16 bit) and given $\alpha = 2$. This is possible if one chooses not only radix-2-lengths but other highly factorizable transform lengths, so-called mixed-radix lengths (/2/).

3. Determination of convenient moduli

Mixed-radix lengths have the form

$$N = 2^a \cdot 3^b \cdot 5^c \cdot 7^d \qquad (5)$$

with integers a, b, c, d ≥ 0 and allow the implementation of the mixed-radix fast Fourier transform algorithm (MFFT) (cf. /2/). The MFFT can be obtained from the single-radix FFT algorithms of lengths $N = 2^a; 3^b; 5^c; 7^d$, respectively. A special class of mixed-radix lengths are the Winograd-numbers.

Winograd-numbers are a product of pairwise relative prime numbers of the set

$$\{2, 3, 4, 5, 7, 9, 16\} \qquad (6)$$

and allow the implementation of the very fast and efficient Winograd-Fourier transform algorithm (WFTA) and the prime-factor algorithm (PFA) (cf. /2/). The biggest Winograd-number is $N = 5 \cdot 7 \cdot 9 \cdot 16 = 5040$. Therefore only 59 Winograd numbers exist in the range $2 \leq N \leq 5040$. The WFTA is derived from short and very efficient transforms of lengths (6). It is clear, that every winograd number is a mixed-radix length. With the help of

a computer program 613 mixed-radix lengths N in the range $2 \leq N \leq 65536$ were found. It is shown in /9/, that the MFFT algorithm is the most flexible and uses the least memory in the computer, while the WFTA and the PFA are the most efficient ones. For the applications of these algorithms to the conventional case of the field of complex numbers the reader is referred to /2/, /10/, /11/.

If one wishes to implement these algorithms in a finite field or a finite ring, for example the ring $Z_m = Z \bmod m$ of residue classes of integers modulo m ($m > 1$ is a given integer), one has to look for convenient moduli m, so that an element $\alpha \in Z$ is a primitive N-th root of unity modulo m where N is of mixed-radix form. In order to simplify the arithmetic modulo m the element $\alpha \in Z$ must have a simple binary representation. We choose $\alpha = 2$. In this note we determine all possible prime moduli $p \leq 2^{16} + 1$ for computers with 16 bit word length, so that $\alpha = 2$ is a primitive N-th root of unity modulo p, where N is of mixed-radix form. From number theory it is known, that the condition

$$N \mid p - 1 \tag{7}$$

must be valid. In /6/ was shown, that the following conditions are equivalent

- α is a primitive N-th root of unity modulo m, (8)

- $\chi_N(\alpha) \equiv 0 \bmod m$; $\gcd(N,m) = 1$. (9)

This means that a convenient modulus m has to be a divisor of $\chi_N(\alpha)$, where χ_N denotes the N-th cyclotomic polynomial. However, if N is large ($N > 250$) the prime factorization of $\chi_N(2)$ is not known. Another concept of determination of moduli are the primitive divisors of $\alpha^N - 1$. By definition an integer m with

$$m \mid \alpha^N - 1, \tag{10}$$

$$\gcd(m, \alpha^n - 1) = 1 \text{ for all } n = 1, \ldots, N-1 \tag{11}$$

is called <u>a primitive divisor of $\alpha^N - 1$</u> (cf. /12/).

Note that (10) - (11) is equivalent to (2) - (3). Therefore m

is a primitive divisor of $\alpha^N - 1$, if and only if α is a primitive N-th root of unity modulo m. The primitive divisors of $2^N - 1$ are listed in /13/, but for large N ($N > 1200$) such divisors are not known. Therefore the only possibility to obtain convenient moduli is the direct computation of the order N of $\alpha = 2$ modulo each prime number $p \leq 2^{16} + 1$ and select those p for which N is of mixed-radix form. Using the property of the primitive N-th root of unity modulo m

$$2^{N/2} \equiv -1 \bmod m, \text{ if N is even,}$$

the order of the element $\alpha = 2$ modulo each prime number p ($p \leq 2^{16} + 1$) was determined by the help of a computer PDP 11/40. In order to avoid errors, the results were compared with the number-theoretic tables of Kraitchik /14/ and are shown in Table 1. Another good possibility to control the results is now described. We consider the structure of the primitive divisors of $\alpha^N - 1$ (cf. /12/).

Lemma 1: Let $N > 1$ be an odd integer. The primitive divisors of $\alpha^N - 1$ and $\alpha^{2N} - 1$ have the form

$$2kN + 1 \quad (k > 0, \text{ integer}). \tag{12}$$

Proof: Because of the Chinese Remainder Theorem and the equivalence of (1) - (2) and (10) - (11) we can restrict our attention to the primitive prime divisors p of $\alpha^N - 1$ and $\alpha^{2N} - 1$, respectively. From $\alpha^N \equiv 1 \bmod p$ it follows $N \mid p - 1$. By assumption the number N is odd, and $p - 1$ is always even. This gives $2N \mid p - 1$ and $p = 2kN + 1$ with an integer $k > 0$. Further by $\alpha^{2N} \equiv 1 \bmod p$ it follows $2N \mid p - 1$ and $p = 2kN + 1$. ■

Table 1: Mixed-radix lengths N ($2 \leq N \leq 65536$) and prime moduli $p \leq 2^{16} + 1$ for number-theoretic transforms with $\alpha = 2$ as primitive N-th root of unity modulo p

N	p	N	p	N	p
2	3	72	433, 38737	480	23041
3	7	81	2593	486	1459
4	5	84	1429, 14449	490	491
5	31	96	193	500	7001, 28001
7	127	100	101, 8101	504	1009, 21169
8	17	105	29191	525	4201, 7351
9	73	112	5153	540	541, 30241, 49681
10	11	135	271	560	4481
12	13	144	577	576	3457
14	43	162	163	648	1297, 3889
15	151	168	3361	700	701
16	257	175	39551	735	41161
18	19	180	181, 54001	750	2251
20	41	196	197	756	757
21	337	200	401	784	3137, 50177
24	241	210	211	810	6481, 9721
25	601, 1801	224	449, 2689	882	883, 3529, 22051
28	29, 113	243	487	960	26881
30	331	245	1471	972	2917, 4861
32	65537	270	811, 15121	1000	4001
35	71	288	1153, 6337	1008	34273
36	37, 109	300	1201, 6391	1080	2161, 21601
40	61681	336	2017	1200	4801, 55201
42	5419	350	1051	1260	2521
45	631, 23311	375	751	1296	10369
48	97, 673	378	379	1323	2647
50	251, 4051	384	769	1372	1373
60	61, 1321	392	7057	1400	2801
64	641	400	1601, 25601	1440	37441
70	281	420	421	1470	5881

N	p	N	p	N	p
1620	1621	4860	19441	12000	24001
1680	13441	4900	44101	12250	12251
1792	10753	5120	61441	12348	24697
1875	33751	5292	15877	13720	54881
1920	49921	5600	33601	14700	29401
1960	7841, 35281	6048	12097	15435	30871
2058	8233	6144	12289	15750	47251
2100	6301	6174	18523, 49393	15876	47629
2187	39367	6300	12601	16128	32257
2250	9001, 11251	6860	13721, 34301	17010	17011
2268	2269	7200	43201	21168	42337
2304	18433	7290	36451	22500	22501
2400	9601, 57601	8000	16001	22680	45361
2401	14407	8505	51031	23520	47041
2835	28351	8820	8821	26250	26251
3240	32401	8960	17921	28812	28813
3500	21001, 52501	9408	37633	29160	58321
3600	14401	10080	20161	32256	64513
3780	7561	10240	40961	37500	37501
3840	7681, 15361	10500	10501	56700	56701
4320	8641	10935	21871	62500	62501
4374	17497				

Lemma 2: Let $N > 1$ be an odd integer. The primitive divisors of

$$\left.\begin{matrix} 2^N - 1 \\ 2^{2N} - 1 \end{matrix}\right\} \quad \text{have the form } 2kN + 1, \tag{13}$$

$$2^{4N} - 1 \quad \text{have the form } 4kN + 1, \tag{14}$$

$$2^{2^n N} - 1 \quad \text{have the form } 2^{n+1}kN + 1, \ (n \geq 3). \tag{15}$$

Proof: There is only necessary to show (15), since (13) follows directly from (12), and (14) can be proved in analogy to Lemma 1.

Let p be a prime divisor of $2^{2^n N} - 1$. From (2) - (3) and

$$2^{2^n N} \equiv 1 \bmod p$$

we get $2^n N \mid p-1$ and $p = 2^n N q + 1$ with an integer $q > 0$. Note that the element 2 belongs to the exponent $2^n N$ modulo p. With the help of the quadratic residue character of 2

$$\left(\frac{2}{p}\right) = \begin{cases} +1, & \text{if } p \equiv 1;\ 7 \bmod 8, \\ -1, & \text{if } p \equiv 3;\ 5 \bmod 8 \end{cases}$$

and Euler's criterion

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \bmod p \qquad (a \in Z)$$

we obtain for $n \geq 3$

$$\left(\frac{2}{2^n N q + 1}\right) \equiv 2^{2^{n-1} N q} \equiv 1 \bmod p.$$

Therefore

$$2^n N \mid 2^{n-1} N q$$

and q is even. With $k > 0$ and $q = 2k$ one obtains (15). ■

Example 1: Let the mixed-radix length $N = 6860$ be given. From Lemma 2 we obtain, that the primitive divisors of $2^{6860} - 1$ have the form $m = 6860kN + 1$. With the help of a prime number table we can find the following possibilities for primitive prime divisors $p < 2^{16} + 1$ of the form $p = 6860k + 1$:

13721, 34301, 41161, 54881.

This set of prime numbers must contain all primitive prime divisors $p < 2^{16} + 1$ of $2^{6860} - 1$. Therefore one has to compute the order of the element $\alpha = 2$ modulo each of these prime numbers. So only two primitive prime divisors of $2^{6860} - 1$ were found: 13721 and 34301. Note that the product of these two primitive prime divisors is a primitive divisor of $2^{6860} - 1$, but it is greater than $2^{16} + 1$.

Example 2: In the same way as in example 1 one can find the two

primitive prime divisors 97 and 673 of $2^{48} - 1$. Therefore three primitive divisors $m \leq 2^{16} + 1$ occur, so that $\alpha = 2$ is a primitive 48-th root of unity modulo m, where m is equal to 97, 673 or $97 \cdot 673 = 65281$.
It is clear that this method is a good control of the direct computation described above. A computer program was written to determine the structure of the primitive divisors of $2^N - 1$ with N of mixed-radix form. The resulting sets of possible primitive prime divisors were compared with the direct computed results and the table of Kraitchik and are shown in table 1.

4. Results

Only 143 from 613 possible mixed-radix lengths N in the range $2 \leq N \leq 65536$ were found, so that $\alpha = 2$ is a primitive N-th root of unity modulo p with $p \leq 2^{16} + 1$. However, there are a lot of other important cases of practical interest. For example the mixed-radix length $N = 250$ does not occur in the table, but if one chooses $\alpha = 2^2 = 4$ as a primitive 250-th root of unity one can use the moduli 7001 or 28001 given in the table for $N = 500$ and $\alpha = 2$. Using this method one can find with the help of the given table a lot of suitable cases, because the arithmetic with α as a power of two is easy to perform.
Only 39 from 59 possible Winograd-numbers N were found, so that $\alpha = 2$ is a primitive N-th root of unity modulo p with a prime number $p \leq 2^{16} + 1$. However, there are also a lot of other interesting cases for Winograd-numbers N. For example the Winograd-numbers $N = 360$ and $N = 720$ don't occur in the table, but $N = 1440$, $N = 3 \cdot 1440 = 4320$, $N = 5 \cdot 1440 = 7200$ can be used to find elements with simple binary representation, that are primitive 360-th or 720-th roots of unity modulo m. For example $\alpha = 2^{10}$ is a primitive 720-th root of unity modulo 43201 and $\alpha = 2^4$ is a primitive 360-th root of unity modulo 37441. This method is also useful for finding suitable moduli in order to avoid overflow modulo m in the computation of cyclic convolutions (cf. /7/).

The most interesting cases of the table are those with two or three moduli for given mixed-radix length N, because then two or three processors can perform the NTT in parallel. This is also important to avoid overflow. If the maximal amplitudes of the signals $\underline{x}$ and $\underline{h}$ are scaled by

$$\max|x_i| \leq A, \quad \max|h_i| \leq A \quad (i = 0, \ldots, N-1),$$

then (cf. /8/)

$$A_1 \leq \sqrt{\frac{m-1}{2N}}$$

gives the dynamic range A_1 for one-dimensional overflow-free cyclic convolution modulo m, where m has the prime factorization

$$m = p_1^{r_1} p_2^{r_2} \ldots p_s^{r_s}.$$

In the case of two-dimensional signals of size N x N one obtains the dynamic range A_2 for two-dimensional convolution

$$A_2 \leq \frac{1}{N} \sqrt{\frac{m-1}{2}}.$$

For example N = 540 is a good mixed-radix length for signal processing operations. If two processors work in parallel modulo 30241 and modulo 49681, both numbers are smaller than $2^{16} + 1$, one can obtain a dynamic range of $A_1 \leq 1179$ for one-dimensional convolutions. If three processors work in parallel with moduli 541, 30241, 49681 one can obtain the excellent dynamic ranges of $A_1 \leq 27433$ and $A_2 \leq 1180$ for one- and two-dimensional convolutions, respectively. Unfortunately, the maximal number of processors for parallel computation is limited to three in the range $2 \leq N \leq 65536$ and $p \leq 2^{16} + 1$ ($\alpha = 2$). The largest transform lengths N for parallel computing were found to be N = 6174 and N = 6860. If one uses 8-bit processors, then one has to look for primitive prime divisors $p \leq 2^8 + 1$. Two interesting cases for N = 28 and N = 36 can be obtained for parallel computation. The dynamic range for both cases in one-dimensional convolutions is $A_1 \leq 8$.

If one wishes to apply the radix-2 or radix-4 algorithm, one can use for example $N = 960 = 15 \cdot 64$ and find that $\alpha = 2^{15}$ is a primitive 64-th root of unity modulo 26881. Another interesting case is $N = 3840 = 15 \cdot 256$. The element $\alpha = 2^{15}$ is a primitive 256-th root of unity modulo 7681 and modulo 15361 and parallel processing is possible. In this way efficient transforms with $\alpha = 2$ or α as power of two as primitive N-th root of unity of mixed-radix lengths N for practical applications can be constructed.

References

/1/ McClellan, J. H., and Rader, C. M.: Number Theory in Digital Signal Processing. Englewood Cliffs 1979

/2/ Nussbaumer, H. J.: Fast Fourier Transform and Convolution Algorithms. Berlin 1981

/3/ Reed, I. S., Truong, T. K., Kwoh, Y. S., and Hall, E. L.: Image processing by transforms over a finite field. IEEE Trans. Comput. C-26, 874 - 881 (1977)

/4/ Reed, I. S., Scholtz, R. A., Truong, T. K., and Welch, I. R.: The fast decoding of Reed-Solomon codes using Fermat theoretic transforms and continued fractions. IEEE Trans. Inform. Theory IT-24, 100 - 106 (1978)

/5/ Siu, W. C., and Constantinides, A. G.: Very fast discrete Fourier transform, using number theoretic transform. IEE Proceedings G-130, 201 - 204 (1983)

/6/ Creutzburg, R., and Tasche, M.: F-Transformation und Faltung in kommutativen Ringen. Elektron. Informationsverarb. Kybernet. 21, 129 - 149 (1985)

/7/ Creutzburg, R., and Tasche, M.: Zahlentheoretische Transformationen und primitive Einheitswurzeln in einem Restklassenring modulo m, I, II. Rostock. Math. Kolloq. 25, 4 - 22 (1984); 26, 103 - 109 (1984)

/8/ Creutzburg, R., and Grundmann, H.-J.: Die Fermattransformation und ihre Anwendung bei der schnellen Berechnung digitaler Faltungen. Rostock. Math. Kolloq. 24, 77 - 98 (1983)

/9/ Blanken, J. D., and Rustan, P. L.: Selection criteria for efficient implementation of FFT algorithms. IEEE Trans. Acoust. Speech Signal Process. ASSP-30, 107 - 109 (1982)

/10/ Silverman, H. F.: An introduction to programming the Winograd Fourier transform algorithm (WFTA). IEEE Trans. Acoust. Speech Signal Process. ASSP-25, 152 - 165 (1977)

/11/ Kolba, D. P., and Parks, T. W.: A prime factor FFT algorithm using high-speed convolution. IEEE Trans. Acoust. Speech Signal Process. ASSP-25, 281 - 294 (1977)

/12/ Kraitchik, M.: Introduction à la Théorie des Nombres. Paris 1952

/13/ Brillhart, J., Lehmer, D. H., Selfridge, J. L., Tuckerman, B., and Wagstaff Jr., S. S.: Factorizations of $b^n \pm 1$ up to High Powers. Contemporary Mathematics, Vol. 22, AMS, Providence 1983

/14/ Kraitchik, M.: Recherches sur la Théorie des Nombres. Paris 1924

received: December 3, 1984

Authors addresses:

Dr. R. Creutzburg
Akademie der Wissenschaften
der DDR
Zentralinstitut für Kybernetik
und Informationsprozesse
Kurstraße 33
DDR-1086 Berlin

Dipl.-Phys. H.-J. Grundmann
Akademie der Wissenschaften
der DDR
Institut für Kosmosforschung
Kalkhorstweg
DDR-2800 Neustrelitz

Rostock. Math. Kolloq. 28, 111 - 120 (1985) 60A99
62A99

Rainer Hellmann
Ingo Kölbl

Zu Problemen der Einführung und Behandlung von Elementen der Wahrscheinlichkeitsrechnung und Statistik im Mathematikunterricht

Elementare Kenntnisse aus der Wahrscheinlichkeitsrechnung und Statistik (Stochastik) gehören zur Allgemeinbildung eines jeden Menschen, um Aussagen aus den Naturwissenschaften und der Gesellschaft richtig interpretieren und einordnen zu können. Neben der Kenntnis einiger wichtiger Begriffe, wie z. B. Zufall, Ereignis, Häufigkeit, Wahrscheinlichkeit, ist das Verständnis für stochastische Vorgänge und das sogenannte stochastische Denken von besonderer Bedeutung. Dieses bei der heranwachsenden Generation herauszubilden, muß Anliegen der Schule sein und insbesondere im Mathematikunterricht im Zusammenhang mit einer engen Koordination mit naturwissenschaftlichen Unterrichtsfächern erfolgen.

Eine Analyse der zur Zeit gültigen Lehrpläne für den Mathematikunterricht und denen der Unterrichtsfächer Biologie, Chemie und Physik zeigt, daß es schon eine Reihe von Ansatzpunkten für die Behandlung stochastischer Probleme in diesen Fächern gibt. Beispiele sind

- die Behandlung der Mendelschen Gesetze und der Häufigkeitsverteilungen im Biologieunterricht, der Brownschen Molekularbewegung und des Begriffs der Aufenthaltswahrscheinlichkeit im Physik- und Chemieunterricht,
- die Auswertung von Daten und insbesondere Meßwerten bei Experimenten,
- die Arbeit mit Diagrammen u. a. graphischen Darstellungen.

Leider ist es aber so, daß die vorhandenen Möglichkeiten zur Herausbildung stochastischer Denk- und Arbeitsweisen bei den Schülern nicht oder nur zum Teil genutzt werden. Es erweist

sich unseres Erachtens als notwendig, Zielvorstellungen und daraus resultierende Inhalte zur Aufnahme von Elementen der Stochastik im Mathematikunterricht mit bereits vorhandenen Möglichkeiten im mathematisch-naturwissenschaftlichen Unterricht zu verbinden.
So bietet es sich z. B. an, die Mittelwertbildung im Mathematikunterricht so einzuführen, daß die Untersuchung an einer Stichprobe aus einer Grundgesamtheit erfolgt und die gewonnene Aussage auf die Grundgesamtheit übertragen wird. Ausgegangen wird von einer Größenmessung von 5 - 7 Schülern einer Klasse und der Mittelwertbildung der erhaltenen Meßwerte. Die nun folgende Überlegung mit den Schülern befaßt sich damit, ob und wann von dieser Mittelwertbildung auf eine durchschnittliche Größe der Schüler der Klasse geschlossen werden kann. Im Zusammenhang damit sind Fragen der Auswahl der Schüler (zufällige Stichprobe) aus der Klasse (Grundgesamtheit) zu diskutieren mit dem Ziel, die Erkenntnis bei den Schülern herauszubilden, daß ein Schluß auf die Gesamtheit mit kalkulierbarem Risiko möglich ist und daß nur durch eine Zufallsauswahl im Mittel Verzerrungen der Aussage vermieden werden können. Dabei können bei einzelnen Schülern durchaus "größere" Abweichungen auftreten. Um dieses noch deutlicher werden zu lassen, ist anhand der Einzelfehler (Abweichungen vom Mittelwert) über die zufällige Variabilität und über Maßzahlen zu ihrer Beschreibung (wie mittlerer und größter absoluter Fehler, Varianz usw.) zu sprechen. So kann der relative (und prozentuale) Fehler eingeführt und berechnet sowie eine Fehlerabschätzung vorgenommen werden. Zusammenfassend wird verallgemeinert, daß von Daten einer Stichprobe und deren Auswertung auf allgemeingültige Aussagen der Grundgesamtheit geschlossen werden kann, ohne diese vollständig auswerten zu müssen. Die Risiken eines solchen statistischen Schlusses müssen diskutiert werden. Eine weitere Möglichkeit, stochastische Gedankengänge bei den Schülern zu entwickeln, liegen in der Betrachtung des "Schätzwertcharakters" von Aussagen, denen statistisches Material zugrunde liegt. So können mit den Schülern im Biologieunterricht durchgeführte Keimprobenuntersuchungen, in denen aus einer bestimmten Anzahl

von Versuchsobjekten nach bestimmten Merkmalen einige Versuchsobjekte ausgewählt und ausgezählt wurden, mathematisch dahingehend ausgewertet werden, von einem in Prozenten angegebenen Ergebnis Rückschlüsse auf die Qualität der Gesamtheit zu ziehen. Es wird hierbei der statistische Charakter des Prozentbegriffs in Form einer Schätzung hervorgehoben. Im Vergleich mit anderen Aufgaben zur Prozentrechnung, in denen die Prozentsätze genaue Angaben darstellen, da zu ihrer Berechnung eine Grundgesamtheit herangezogen wurde, ist den Schülern anzuerziehen, daß sie eine Sicht dafür bekommen, wie man die auszuwertenden Daten gewinnt und wie die erzielten Ergebnisse zustandekommen bzw. gewertet werden müssen.

Die beiden angeführten Beispiele verdeutlichen das methodische Vorgehen, welches für eine Einführung von Elementen der Stochastik im Mathematikunterricht und für die Herausbildung von stochastischen Denk- und Arbeitsweisen vorgeschlagen wird. Ausgangspunkt sind dabei Aufgaben aus dem unmittelbaren Lebensbereich der Schüler und aus anderen Unterrichtsfächern, an denen die Begriffe erarbeitet und aus theoretischer Sicht verallgemeinert werden. Bei diesem mehr induktiven Vorgehen werden neben der Vermittlung von Kenntnissen und der Herausbildung von Fähigkeiten zum Aufgabenlösungsprozeß auch erzieherische Potenzen des Unterrichts umittelbar wirksam. Auch für die Einführung des Wahrscheinlichkeitsbegriffs in der Mittelstufe wird ein solches Vorgehen vorgeschlagen, wobei sich verschiedene Konzeptionen für eine Definition dieses Begriffs anbieten.

Der "klassische" Wahrscheinlichkeitsbegriff hat den Vorteil, bei praktischen Berechnungen oft schnell und bequem zum Ziel zu führen, da es in vielen Fällen zulässig ist, Gleichwahrscheinlichkeit der Elementarereignisse vorauszusetzen. Es lassen sich viele Aufgaben und Anwendungsbeispiele formulieren, die auch auf die Erfahrungen der Schüler (z. B. Glücksspiele) zurückgreifen. Schon für die Mittelstufe besteht die Möglichkeit, den klassischen Wahrscheinlichkeitsbegriff in Aufgaben mit einem gewissen Praxisbezug anzuwenden, zum Beispiel zur Erläuterung der Diffusion, der Behandlung des 2. Mendelschen Gesetzes und der Berechnung des Risikos rezessiver Erbkrankheiten. Man muß

jedoch darauf achten, daß nicht dadurch, daß die Kombinatorik in den Vordergrund rückt, die Entwicklung einer stochastischen Denkweise gehemmt wird und solche Interpretationen von Ergebnissen auftreten wie: "Jeder sechste Wurf ist eine Sechs" oder "Jedes zweite Kind wird ein Junge". Da beim Arbeiten mit dem klassischen Wahrscheinlichkeitsbegriff kein umfangreiches Datenmaterial gegeben bzw. experimentell erarbeitet werden muß, wird sich solch eine Unterrichtsstunde nicht wesentlich von einer sonstigen Arithmetikstunde unterscheiden bis auf die Besonderheit, daß im Stochastikunterricht größeres Gewicht auf eine richtige Interpretation der Ergebnisse gelegt werden muß.
Zur Einführung des Wahrscheinlichkeitsbegriffs ist die "klassische Definition" u. E. jedoch nicht geeignet, denn neben der unzulässigen Einengung auf gleichwahrscheinliche Elementarereignisse ergibt diese "Definition" ja auch den bekannten Zirkelschluß. Solche Zugeständnisse sollte man auch in der Polytechnischen Oberschule nicht machen, sondern den Wahrscheinlichkeitsbegriff so einführen, daß er erweiterungs- bzw. entwicklungsfähig ist, auch z. B. im Hinblick auf eine Axiomatik in Klasse elf oder der Hochschule (vgl. /1/). Wir werden darauf im weiteren noch näher eingehen.
Bei der Einführung eines "statistischen" Wahrscheinlichkeitsbegriffs gibt es eine Reihe von Vorzügen für die Herausbildung einer stochastischen Denkweise. Die Schüler erkennen, wie sich Gesetzmäßigkeiten als Folge von Zufallserscheinungen realisieren. Das Gesetz der großen Zahlen von Bernoulli wird dabei implizit benutzt und stellt einen besonderen Anwendungsbezug dar. Vorausgesetzt, daß nicht nur Experimente mit Münze und Würfel durchgeführt werden, kann ein anwendungsorientierter Stochastikunterricht bei den Schülern auch die Überzeugung herausbilden, daß die Mathematik, obwohl sie sich mit ideellen Objekten beschäftigt, ihren Ursprung in der Praxis hat.
Problematisch ist die Bereitstellung von (tatsächlich anwendungsbezogenen) Aufgaben zum statistischen Wahrscheinlichkeitsbegriff für die Polytechnische Oberschule. Es müssen Experimente durchgeführt werden oder größeres Datenmaterial verarbeitet werden. Letzteres wird durch die Verwendung des elektronischen

Taschenrechners in der Schule erleichtert. Es liegt also nahe, ihn in der beschreibenden Statistik bei der Auswertung empirischer Verteilungen anzuwenden. Ein Beispiel wird im weiteren noch folgen.

Zur Einführung des Wahrscheinlichkeitsbegriffs ist die "statistische Definition" u. E. jedoch auch nicht geeignet, weil zum einen, wie Sill in /2/ feststellt, "dadurch in unzulässiger Weise der Wahrscheinlichkeitsbegriff mit der Durchführung von Versuchen gekoppelt ist" und der Eindruck entsteht, die Wahrscheinlichkeit sei kein exaktes Maß. Es muß herausgestellt werden, daß die Wahrscheinlichkeit ein Maß ist, das den gleichen Charakter hat wie z. B. die Länge, das Volumen, und daß sie mit hinreichender Genauigkeit gemessen werden kann. Zum anderen muß man, will man exakt vorgehen, die stochastische Konvergenz der relativen Häufigkeit betrachten. Die dazu notwendige Benutzung von Bernoullis Theorem zum Zwecke der Definition ergibt dann aber wieder einen Zirkelschluß.

Die "statistische Definition" würde also ebenso wie die "klassiche Definition" den Wahrscheinlichkeitsbegriff nicht nur einengen, sondern auch eine spätere Weiterentwicklung des Begriffs in oberen Klassen blockieren. Trotzdem wird man natürlich nicht darauf verzichten, an geeigneter Stelle herauszuarbeiten, daß die relative Häufigkeit in unabhängigen Versuchen um die Wahrscheinlichkeit variiert.

Eine axiomatische Definition der Wahrscheinlichkeit halten wir für die Polytechnische Oberschule für indiskutabel und wollen darauf an dieser Stelle auch nicht weiter eingehen.

Unseres Erachtens sollte der Wahrscheinlichkeitsbegriff mehr in philosophischer Richtung im Mathematikunterricht eingeführt werden und zwar in Anlehnung an Hörz (vgl. /3/) als "Maß für die zufällige Verwirklichung von Möglichkeiten". Diese Begriffsbestimmung beinhaltet keinerlei Einschränkungen. Die "klassische" und "statistische" Definition sollten den Schülern lediglich als Berechnungsmöglichkeiten gegeben werden, was natürlich in engem Zusammenhang mit der Einführung des Wahrscheinlichkeitsbegriffs geschehen muß. Ein Beispiel:

Der Wahrscheinlichkeitsbegriff soll am Würfel eingeführt werden (wir sind zwar auch gegen "Würfelbudenmathematik", aber der

Würfel ist meist der einzige Zufallsgenerator, mit dem die Schüler tatsächlich Erfahrungen gemacht haben). Nachdem die Schüler die Elemente des "philosophischen Wahrscheinlichkeitsbegriffs" auf den Würfel angewendet haben (Möglichkeiten entspricht Augenzahl; zufällige Verwirklichung entspricht Würfeln; das zufällige Ereignis ist die gewürfelte Augenzahl) soll nun ein "Maß" für die Wahrscheinlichkeit bestimmt werden. Jetzt wird zunächst über die relative Häufigkeit der statistische Wahrscheinlichkeitsbegriff als Berechnungsmöglichkeit erarbeitet, und anschließend wird erörtert, unter welchen Bedingungen der klassische Wahrscheinlichkeitsbegriff anwendbar wäre. Dadurch wird bei der Einführung des Wahrscheinlichkeitsbegriffs erreicht, daß sich die Nachteile der beiden Berechnungsmöglichkeiten (bez. der Interpretation des Ergebnisses) zu einem gewissen Grade kompensieren und keine einseitige Interpretation des Ergebnisses suggeriert wird. In den folgenden Unterrichtsstunden wird der statistische Wahrscheinlichkeitsbegriff jedoch nur noch in den Fällen angewendet, in denen der klassische versagt. Damit er trotzdem nicht unterrepräsentiert bleibt, erfordert das eine entsprechende Bereitstellung von Aufgaben, die für den Unterricht geeignet sind.

Eine Möglichkeit wäre das Anwenden des statistischen Wahrscheinlichkeitsbegriffs in der beschreibenden Statistik. Die anwendungsorientierte Betrachtung der Wahrscheinlichkeitsrechnung in enger Verbindung mit der beschreibenden Statistik fördert die Entwicklung einer stochastischen Denkweise. Eine Verbindung ist durch den Zusammenhang von Zufallsgröße und zufälligem Ereignis gegeben, den man aber in der Mittelstufe natürlich nicht vollständig betrachten kann.

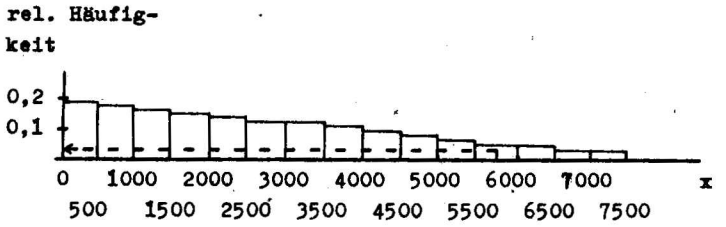
Prinzipiell setzt die Anwendung des Wahrscheinlichkeitsbegriffs eine Zufallsauswahl der Stichprobenelemente voraus. Es gibt jedoch Beispiele, wo die zufällige Auswahl nicht von uns selbst vorgenommen wird, sondern quasi an uns herangetragen wird, z. B. bei der Wahrscheinlichkeit für einen Raucher, an Lungenkrebs zu erkranken. Ähnlich verhält es sich mit Unfallstatistiken, Sterbestatistiken usw.

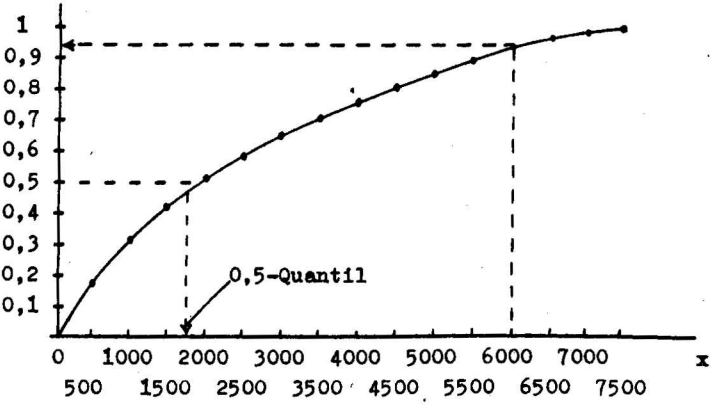
Im folgenden wird eine Aufgabe vorgestellt, bei der die zufällige Auswahl ebenfalls naturgemäß gegeben ist.

Im Biologieunterricht der Klasse 8 wird die Lebenserwartung für Menschen behandelt. Ein gegebener "Erwartungswert" (für die Schüler wäre der Begriff "Mittelwert" wohl besser) wird leicht fehlinterpretiert, und es wäre eine lohnende Aufgabe für einen zeitlich vorhergehenden Stochastikunterricht, einmal eine Lebensdauerverteilung mit Hilfe der beschreibenden Statistik auszuwerten und zu interpretieren. Dabei ist es ganz natürlich, Wahrscheinlichkeiten für verschiedene Bereiche der Lebensdauer anzugeben.
Das folgende Beispiel, das einen Dauertest mit 100 Bauelementen beschreibt, läßt sich dann ohne weiteres auf die Belange des Biologieunterrichts übertragen. Folgende Daten liegen vor:

Lebensdauer x in Betriebsstunden		abs. Häufigkeit	rel. Häufigkeit	rel. Summenhäufigkeit
von 0	bis unter 500	18	0,18	0,18
500	1000	15	0,15	0,33
1000	1500	12	0,12	0,45
1500	2000	9	0,09	0,54
2000	2500	8	0,08	0,62
2500	3000	7	0,07	0,69
3000	3500	7	0,07	0,76
3500	4000	6	0,06	0,82
4000	4500	5	0,05	0,87
4500	5000	4	0,04	0,91
5000	5500	3	0,03	0,94
5500	6000	2	0,02	0,96
6000	6500	2	0,02	0,98
6500	7000	1	0,01	0,99
7000	7500	1	0,01	1,00

Mittelwert (mittlere Lebensdauer) $\bar{x} = 2230$





Werden die beiden graphischen Darstellungen gleichzeitig ausgewertet, so bietet gerade dieses Beispiel die Möglichkeit, durch Vergleich mit dem Histogramm die spezielle Bedeutung der Summenkurve zu erläutern, die von den Schülern oft nicht erkannt wird. Denn bei einem Bauelement oder Individuum ist es ja in den meisten Fällen von größerem Interesse zu erfahren, wie groß die Wahrscheinlichkeit ist, eine bestimmte Lebensdauer zu erreichen, während die Frage nach dem Ausfall innerhalb eines relativ kleinen Zeitintervalls für die Praxis von geringer Bedeutung ist.

Bei unserem speziellen Beispiel bieten sich folgende Möglichkeiten an:
Die Wahrscheinlichkeit, im Zeitraum von 5500 bis 6000 Betriebsstunden auszufallen, beträgt etwa 0,02 (vgl. Histogramm).
Die Wahrscheinlichkeit bis zur Zeit von 6000 Betriebsstunden auszufallen beträgt etwa 0,96 (vgl. Summenkurve).
Die Wahrscheinlichkeit bis zur Zeit von 6000 Betriebsstunden nicht auszufallen, also eine höhere Lebensdauer zu erreichen, ist etwa 1 - 0,96 = 0,04.
Interessant ist noch das Ablesen des P-Quantils für P = 0,5 (Halbwertzeit) aus der Darstellung der relativen Summenhäufigkeit. Es beträgt rund 1750 Betriebsstunden, also ist die Hälfte der Bauelemente bereits vor Erreichen der mittleren Lebensdauer ausgefallen. Durch die Behandlung der Halbwertzeit im Stochastikunterricht würde auch die spätere Anwendung dieses Begriffs im Physikunterricht der zehnten Klasse erleichtert werden.

Die aus den relativen Häufigkeiten der Stichprobe ermittelten Wahrscheinlichkeiten sind Schätzwerte, was den Schülern immer wieder deutlich vor Augen geführt werden muß. Welche der hier gezeigten Möglichkeiten zur Auswertung eines Histogramms bzw. einer Summenkurve im Unterricht jeweils genutzt werden, hängt vom jeweiligen Ziel der Unterrichtsstunde ab. Wie wir gesehen haben, wird durch die Anwendung des Wahrscheinlichkeitsbegriffs in der beschreibenden Statistik diese selbst interessanter, und die graphische Darstellung von Häufigkeitsverteilungen bleibt kein formaler Selbstzweck. Kosswig gibt in /4/ ein weiterführendes Konzept der Verbindung von Wahrscheinlichkeitsrechnung und beschreibender Statistik an und beschreibt, wie bestimmte Begriffe der Wahrscheinlichkeitsrechnung mit Hilfe der beschreibenden Statistik interpretiert und motiviert werden können.
Unsere Untersuchungen zur Einführung von Elementen der Wahrscheinlichkeitsrechnung und Statistik und zu einer damit zusammenhängenden Herausbildung von stochastischen Denkweisen bei den Schülern im obligatorischen Mathematikunterricht unserer POS werden in der aufgezeigten Richtung weiter fortgesetzt.

Der Schwerpunkt liegt dabei auf Anwendungen aus naturwissenschaftlichen Unterrichtsfächern bzw. der Verwendung von im Mathematikunterricht erworbenem Wissen und Können zur Wahrscheinlichkeitsrechnung und Statistik in anderen Unterrichtsfächern.

Literatur

/1/ Steinbring, H, und Strässer, R.: Rezension von Stochastiklehrbüchern beider Sekundarstufen. Zentralbl. Didakt. Math. 13, 236 - 247 (1981)

/2/ Sill, D.: Internes unveröffentlichtes Material. Pädagogische Hochschule Güstrow 1983

/3/ Hörz, H.: Zufall - eine philosophische Untersuchung. Berlin 1980

/4/ Kosswig, F. W.: Beschreibende Statistik als Anwendung und Motivation von Begriffen der Stochastik. Math. Naturwiss. Unterr. 33, 78 - 87 (1980)

eingegangen: 03. 12. 1984

Anschrift der Verfasser:

Dipl.-Lehrer R. Hellmann
Dr. paed. I. Kölbl
Wilhelm-Pieck-Universität Rostock
Sektion Mathematik
Universitätsplatz 1
DDR-2500 Rostock

Hinweise für Autoren

Manuskripte (in deutscher, ggf. auch in russischer oder englischer Sprache) bitten wir, an die Schriftleitung zu schicken. Die gesamte Arbeit ist linksbündig zu schreiben. Eine Ausnahme hiervon bilden hervorzuhebende Formeln und das Literaturverzeichnis. Der Kopf der Arbeit soll folgende Form haben: Rostock. Math. Kolloq. /Leerzeile/ Vorname Name/ Leerzeile/ Titel der Arbeit/ 1 Zeilenumschaltung/ Unterstreichung/ Leerzeile. Der Text der Arbeit ist eineinhalbzeilig (= 3 Zeilenumschaltungen) zu schreiben mit maximal 63 Anschlägen je Zeile und maximal 37 Zeilen je Seite. Zwischenüberschriften sind wie folgt einzuordnen: 6 Zeilenumschaltungen/ Zwischenüberschrift/ Unterstreichung (ohne Zeilenumschaltung)/ 5 Zeilenumschaltungen. Hervorhebungen sind durch Unterstreichen und Sperren möglich. Ankündigungen wie Satz, Definition, Bemerkung, Beweis u. ä. sind zu unterstreichen und mit einem Doppelpunkt abzuschließen. Vor und nach Sätzen, Definitionen u. ä. ist ein Zeilenabstand von 5 Umschaltungen zu lassen. Fußnoten sind möglichst zu vermeiden. Sollte doch davon Gebrauch gemacht werden, so sind sie durch eine hochgestellte Ziffer im Text zu kennzeichnen und innerhalb des oben angegebenen Satzspiegels unten auf der gleichen Seite anzugeben. Formeln und Bezeichnungen sollen möglichst mit der Schreibmaschine zu schreiben sein. Hervorzuhebende Formeln sind drei Leerzeichen einzurücken und mit 6 Umschaltungen zum übrigen Text zu schreiben. Formelzähler sollen am rechten Rand stehen. Der Platz für Abbildungen ist beim Schreiben auszusparen; die Abbildungen selbst sind in der dem ausgesparten Platz entsprechenden Größe gesondert nach TGL-Vorschrift auf Transparentpapier beizufügen. Der zugehörige Begleittext ist im Manuskript mitzuschreiben. Sein Abstand nach unten beträgt 5 Umschaltungen. Literaturzitate im Text sind durch laufende Nummern in Schrägstrichen (vgl. /8/, /9/ und /10/) zu kennzeichnen und am Schluß der Arbeit unter der Zwischenüberschrift Literatur zusammenzustellen.

Beispiele: (Zeitschriftenabkürzungen nach Math. Reviews)

/8/ Zariski, O., and Samuel, P.: Commutative Algebra. Princeton 1958

/9/ Steinitz, E.: Algebraische Theorie der Körper. J. Reine Angew. Math. 137, 167 - 309 (1920)

/10/ Gnedenko, B. W.: Über die Arbeiten von C. F. Gauß zur Wahrscheinlichkeitsrechnung. In: Reichardt, H. (Ed.): C. F. Gauß, Gedenkband anläßlich des 100. Todestages. S. 193 - 204, Leipzig 1967

Die Angaben sollen in Originalsprache erfolgen; bei kyrillischen Buchstaben soll die bibliothekarische Transkription (Duden) verwendet werden.

Am Ende der Arbeit stehen folgende Angaben zum Autor und zur Arbeit: eingegangen: Datum/ Leerzeile/ Anschrift des Verfassers:/ Titel Initialen der Vornamen Name/ Institution/ Struktureinheit/ Straße Hausnummer/ Land Postleitzahl Ort.

Der Autor wird gebeten, eine Korrektur des Durchschlags vor Offsetmanuskript zu lesen und dabei die mathematischen Symbole einzutragen. Ferner sollte er 1 - 2 Klassifizierungsnummern (entsprechend der "1980 Mathematics Subject Classification" der Math. Reviews) zur inhaltlichen Einordnung seiner Arbeit angeben.