

ROSTOCKER

MATHEMATISCHES KOLLOQUIUM

Heft 43

UNIVERSITÄT



ROSTOCK

Herausgeber:

Der Rektor der Universität Rostock

Wissenschaftliche Leitung:

Prof. Dr. Günther Wildenhain
(Sprecher des Fachbereichs Mathematik)
Dr. Werner Plischke

Redaktionelle Bearbeitung:

Dr. Klaus-Dieter Drews

Herstellung der Druckvorlage:

Dr. Andreas Straßburg

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

Redaktionsschluß: 30. Januar 1991

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

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

Zitat-Kurztitel: Rostock. Math. Kolloq. (1991) 43

Universität Rostock
Abteilung Wissenschaftspublizistik
Vogelsang 13/14, O-2500 Rostock, Bundesrepublik Deutschland
Telefon 36 95 77

Druck: Ostsee-Druck Rostock, BT Greifswald

01000

ROSTOCKER MATHEMATISCHES KOLLOQUIUM

Heft 43

| | | |
|---|--|----|
| DIETLINDE LAU | <i>Kongruenzen auf abgeschlossenen Mengen linearer Funktionen in mehrwertigen Logiken</i> | 3 |
| ROGER LABAHN; SERGEJ L. BESRUKOV | <i>Asymptotic solution of a special equation with binomial coefficients</i> | 17 |
| LOTHAR BERG | <i>Asymptotic expansions of a special integral</i> | 23 |
| MUHAMMAD BERISHA | <i>On coefficients of double lacunary trigonometric series</i> | 27 |
| EDUARD A. DANIELJAN; FRIEDRICH LIESE | <i>The analysis of a $M_r/G_r/1/\infty$ model with time dependent priorities</i> | 31 |
| VOLKER ISERNHAGEN | <i>Die Strategie der orthogonalen Übergänge und ihre Anwendung auf zwei Klassen von Iterationsverfahren zur Lösung von linearen Gleichungen in Hilberträumen</i> | 55 |
| STEFAN WELLEK | <i>Prinzipien der Planung und Auswertung von Äquivalenzstudien</i> | 63 |

DIETLINDE LAU

Kongruenzen auf abgeschlossenen Mengen linearer Funktionen in mehrwertigen Logiken

In der Arbeit werden für die in [14] von A. Szendrei ermittelten Teilklassen linearer Funktionen, die die Funktion $x + y - z$ enthalten, sämtliche Kongruenzen bestimmt. Außerdem wird eine Eigenschaft der Kongruenzen auf abgeschlossenen Mengen aus linearen, idempotenten Funktionen hergeleitet und gezeigt, daß es auf Klassen linearer Funktionen über endlichen Mengen nur abzählbar viele Kongruenzen geben kann.

1. Grundbegriffe und Bezeichnungen

Für eine endliche Menge E sei P_E^n die Menge aller n -stelligen Funktionen über E und $P_E := \bigcup_{n \geq 1} P_E^n$. Falls $E_k := \{0, 1, \dots, k-1\}$, $k \geq 2$, so sei $P_k := P_{E_k}$. Zur Kennzeichnung der Stellenzahl (Arität) einer Funktion f aus P_E^n schreiben wir statt f oft f^n . Die durch $e_i^n(x_1, \dots, x_n) = x_i$ ($1 \leq i \leq n$) und $c_a^n(x_1, \dots, x_n) = a$ ($a \in E$) definierten Funktionen e_i^n und c_a^n aus P_E^n nennen wir Projektionen bzw. Konstanten. Wie in [9] (siehe auch [11]) führen wir $\zeta, \tau, \Delta, \nabla$ und $*$ als Operationen über P_E ein, die für beliebige $f^n, g^n \in P_E$ durch $\zeta f \in P_E^n$, $\tau f \in P_E^n$, $\Delta f \in P_E^{n \max(1, n-1)}$, $\nabla f \in P_E^{n+1}$, $f * g \in P_E^{m+n-1}$ und

$$(\zeta f)(x_1, \dots, x_n) = f(x_2, x_3, \dots, x_n, x_1),$$

$$(\tau f)(x_1, \dots, x_n) = f(x_2, x_1, x_3, \dots, x_n),$$

$$(\Delta f)(x_1, \dots, x_{n-1}) = f(x_1, x_1, x_2, \dots, x_{n-1}) \quad \text{für } n \geq 2 \quad \text{und}$$

$$\zeta f = \tau f = \Delta f = f \quad \text{für } n = 1,$$

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

$$(f * g)(x_1, \dots, x_{m+n-1}) = f(g(x_1, \dots, x_m), x_{m+1}, \dots, x_{m+n-1}) \quad (x_1, \dots, x_{m+n-1} \in P_E)$$

definiert sind.

Eine Funktion f , die sich aus Funktionen einer Menge $F (\subseteq P_E)$ durch endliche Anwendungen der Operationen $\zeta, \tau, \Delta, \nabla, *$ erzeugen läßt, heißt Superposition über F . Die Menge aller Superpositionen über F bezeichnen wir mit $[F]$. Eine Teilmenge $F \subseteq P_E$ heißt abgeschlossene (Teil)menge bzw. Teilklasse von P_E , wenn $[F] = F$ ist.

Eine Äquivalenzrelation κ auf einer Teilklasse F von P_E nennen wir wie üblich eine Kongruenz(relation) von F , wenn F mit $\zeta, \tau, \Delta, \nabla, *$ kompatibel ist, d.h. wenn für beliebige $f, g, s, t \in F$ aus $(f, g), (s, t) \in \kappa$ stets $(f * s, g * t) \in \kappa$ und $(\alpha f, \alpha g) \in \kappa$ für jedes $\alpha \in \{\zeta, \tau, \Delta, \nabla\}$ folgt.

Auf jeder abgeschlossenen Teilmenge F von P_E gibt es nach [9] mindestens drei Kongruenzen $\kappa_0, \kappa_a, \kappa_1$ (die sogenannten trivialen Kongruenzen):

$$(f^n, g^n) \in \kappa_0 \iff \{f, g\} \subseteq F \wedge f = g,$$

$$(f^n, g^n) \in \kappa_a \iff \{f, g\} \subseteq F \wedge n = m,$$

$$(f^n, g^m) \in \kappa_1 \iff \{f, g\} \subseteq F.$$

Die Mengen aller Kongruenzen auf F sei mit $Con F$ bezeichnet. Kongruenzen $\kappa \in Con F$ mit $\kappa \subseteq \kappa_a$ nennen wir Aritätskongruenzen und fassen sie zur Menge $Con_a F$ zusammen.

Für die Erläuterung weiterer im folgenden verwendeten Begriffe und Bezeichnungen sei auf [11] und [4] verwiesen.

2. Kongruenzen auf Teilklassen von L_M

Seien $R = (R, +, \cdot)$ ein unitärer Ring und $M = (M, +, \cdot)$ ein Linksmodul über R . Dann läßt sich eine abgeschlossene Teilmenge L_M von P_M wie folgt definieren:

$$L_M := \bigcup_{n \geq 1} \left\{ f^n \in P_E \mid \exists a_0 \in M \exists a_1, \dots, a_n \in R : f(\tilde{x}) = a_0 + \sum_{i=1}^n a_i \cdot x_i \right\}.$$

Für jede Teilklasse A von L_M , die die Funktion d mit

$$d(x, y, z) := x + y - z$$

enthält, existiert nach [14] ein eindeutig bestimmter Unterring T_A von R und ein eindeutig

bestimmter Untermodul N_A von $R \times M$ über T_A mit

$$A = \bigcup_{n \geq 1} \{f^n \in L_M \mid \exists a_0 \in M \exists a_1, \dots, a_n \in T_A : f(\tilde{x}) = a_0 * \sum_{i=1}^n a_i x_i \wedge (1 - a_1 - a_2 - \dots - a_n, a_0) \in N_A\}.$$

Die Mengen T_A und N_A erhält man wie folgt:

$$\begin{aligned} T_A &:= \{a \in R \mid ax + (1-a)y \in A\}, \\ N_A &:= \{(1-a, b) \in R \times M \mid ax + b \in A^1\}. \end{aligned}$$

Zwecks Bestimmung der Kongruenzen auf A bezeichne I ein beliebiges Ideal von T_A und U einen dazu passenden Untermodul von N_A mit der Eigenschaft:

$$\forall \alpha, \beta \in T_A \forall u, v \in N_A : (\alpha - \beta \in I \wedge u - v \in U \implies \alpha \cdot u - \beta \cdot v \in U). \quad (1)$$

Mit Hilfe von I und U läßt sich nun eine Relation $\kappa(I, U) (\subseteq \kappa_a)$ wie folgt definieren:

$$\left(a_0 + \sum_{i=1}^n a_i x_i, b_0 + \sum_{i=1}^n b_i x_i \right) \in \kappa(I, U) :\iff \forall i \in \{1, \dots, n\} : a_i - b_i \in I \wedge (1 - a_1 - \dots - a_n, a_0) - (1 - b_1 - \dots - b_n, b_0) \in U.$$

Lemma 1 $\kappa(I, U)$ ist eine Kongruenz auf A .

Beweis: Die Äquivalenzrelationeneigenschaften von $\kappa(I, U)$ prüft man leicht nach. Ebenfalls offensichtlich ist die Kompatibilität von $\kappa(I, U)$ mit den Operationen ζ, τ, Δ und ∇ . Sei nun $(f^n, g^n), (s^m, t^m) \in \kappa(I, U)$ mit $f(\tilde{x}) = a_0 + a_1 x_1 + \dots + a_n x_n, g(\tilde{x}) = b_0 + b_1 x_1 + \dots + b_n x_n, s(\tilde{x}) = c_0 + c_1 x_1 + \dots + c_m x_m$ und $t(\tilde{x}) = d_0 + d_1 x_1 + \dots + d_m x_m$. Dann gilt $b_i = a_i + \alpha_i, d_j = c_j + \beta_j$ für gewisse $\alpha_i, \beta_j \in I$ ($i = 1, \dots, n, j = 1, \dots, m$),

$$\begin{aligned} (f * s)(\tilde{x}) &= a_0 + a_1 c_0 + a_1 c_1 x_1 + \dots + a_1 c_m x_m + a_2 x_{m+1} + \dots + a_n x_{m+n-1}, \\ (g * t)(\tilde{x}) &= b_0 + b_1 d_0 + b_1 d_1 x_1 + \dots + b_1 d_m x_m + b_2 x_{m+1} + \dots + b_n x_{m+n-1}, \\ a_1 c_i - b_1 d_i &= -\alpha_1 \beta_i - \alpha_1 c_i - a_1 \beta_i \in I \quad (i = 1, \dots, m) \quad \text{und} \end{aligned}$$

$$\begin{aligned}
& (1 - a_1c_1 - \dots - a_1c_m - a_2 - \dots - a_m, a_0 + a_1c_0) \\
& - (1 - b_1d_1 - \dots - b_1d_m - b_2 - \dots - b_m, b_0 + b_1d_0) \\
= & a_1(1 - c_1 - \dots - c_m, c_0) - b_1(1 - d_1 - \dots - d_m, d_0) \\
& + (1 - a_1 - \dots - a_n, a_0) - (1 - b_1 - \dots - b_n, b_0) \in U
\end{aligned}$$

nach (1) und $(f, g), (s, t) \in \kappa$, womit $\kappa(I, U)$ eine Kongruenz von A ist. ■

Für den Nachweis, daß auf A neben den trivialen Kongruenzen nur noch Kongruenzen des Typs $\kappa(I, U)$ existieren, benötigen wir noch einige Beziehungen und die nachfolgenden drei Lemmata. Mit q_a (für $a \in R$) bezeichnen wir die durch

$$q_a(x, y) := a \cdot x + (1 - a) \cdot y$$

definierte Funktion.

Lemma 2 *Sei $\kappa \subseteq \kappa_a$ eine Kongruenz auf A und sei $f(\tilde{x}) = a_0 + a_1x_1 + \dots + a_nx_n \in A$ sowie $g(\tilde{x}) = b_0 + b_1x_1 + \dots + b_nx_n \in A$. Dann gilt:*

$$(f, g) \in \kappa \iff \forall i \in \{1, \dots, n\} : (q_{a_i}, q_{b_i}) \in \kappa \wedge (\Delta^{n-1}f, \Delta^{n-1}g) \in \kappa$$

mit $\Delta^{i+1}f := \Delta(\Delta^i f)$, $\Delta^1 f := \Delta f$.

Beweis Sei $h \in \{f, g\}$, $h(x_1, \dots, x_n) = \alpha_0 + \alpha_1x_1 + \dots + \alpha_nx_n$, und bezeichne d_n die durch $d_n(x_1, \dots, x_{n+2}) := x_1 + x_2 + \dots + x_{n+1} - n \cdot x_{n+2}$ definierte $n+2$ -stellige Funktion aus $\{\{d\}\}$. Dann ergibt sich die Aussage des Lemmas leicht aus den Identitäten:

$$\begin{aligned}
q_{\alpha_i}(x, y) &= d(h(y, \dots, y, \overset{i}{x}, y, \dots, y), y, h(y, \dots, y)) \text{ und} \\
h(\tilde{x}) &= d_n(q_{\alpha_1}(x_1, x_1), q_{\alpha_2}(x_2, x_1), \dots, q_{\alpha_n}(x_n, x_1), (\Delta^{n-1}h)(x_1), x_1).
\end{aligned}$$
■

Sei $Q_A := \{q_a \in A^2 \mid a \in T_A\}$. Auf Q_A lassen sich zwei Operationen \oplus, \odot auf folgende Weise definieren:

$$\begin{aligned}(q_a \oplus q_b)(x, y) &:= d(q_a(x, y), q_b(x, y), y) \text{ und} \\ (q_a \odot q_b)(x, y) &:= q_a(q_b(x, y), y).\end{aligned}$$

Wegen $q_a \oplus q_b = q_{a+b}$ und $q_a \odot q_b = q_{a \cdot b}$ – was man leicht nachprüft – gilt offenbar:

Lemma 3 Die Algebra (Q_A, \oplus, \odot) ist isomorph zu $(T_A, +, \cdot)$. ■

Für beliebige $g, h \in A^1$ und beliebiges $\alpha \in T_A$ sei nun

$$\begin{aligned}(g \oplus h)(x) &:= d(g(x), h(x), x) \text{ und} \\ (\alpha \odot g)(x) &:= q_\alpha(g(x), x)\end{aligned}$$

gesetzt. Mit Hilfe der Abbildung

$$\varphi : A^1 \longrightarrow N_A, \quad ax + b \longmapsto (1 - a, b)$$

läßt sich ohne Mühe folgendes Lemma beweisen:

Lemma 4 Die Linksmoduln (A^1, \oplus, \odot) und $(N_A, +, \cdot)$ über T_A sind isomorph. ■

Satz 1 Auf einer beliebigen Teilklasse A von L_M , die die Funktion d enthält, gibt es nur die Kongruenzen $\kappa_0, \kappa_a, \kappa_1$ sowie die des Typs $\kappa(I, U)$.

Beweis Bezeichne κ eine Kongruenz auf A , die von κ_1 verschieden ist. Da jede Teilklasse A von L_M mit $d \in A$ auch die Projektionen enthält, gilt dann bekanntlich $\kappa \subseteq \kappa_a$ (siehe [2] oder [1]). Nach Lemma 2 und 3 (bzw. 4) bewirkt κ auf T_A (bzw. N_A) eine eindeutig bestimmte Kongruenz μ_κ (bzw. ν_κ):

$$(q_a, q_b) \in \kappa \iff (a, b) \in \mu_\kappa \quad (\text{bzw.} \tag{2}$$

$$(ax + b, cx + d) \in \kappa \iff ((1 - a, b), (1 - c, d)) \in \nu_\kappa. \tag{3}$$

Bekanntlich gibt es dann ein Ideal I von T_A und einen Untermodul U von N_A mit den Eigenschaften:

$$(a, b) \in \mu_\kappa \iff a - b \in I, \quad (4)$$

$$(u, v) \in \nu_\kappa \iff u - v \in U. \quad (5)$$

Zwecks Nachweis von (1) seien $a, b \in T_A$ und $u = (1 - \alpha, \beta)$, $v = (1 - \gamma, \delta) \in N_A$ mit $a - b \in I$ sowie $u - v \in U$. Dann gilt $\{(q_a, q_b), (\alpha x + \beta, \gamma x + \delta)\} \subseteq \kappa$. Folglich haben wir

$$(q_a(\alpha x + \beta, x), q_b(\gamma x + \delta, x)) = ((a\alpha + 1 - a)x + a\beta, (b\gamma + 1 - b)x + b\delta) \in \kappa,$$

womit

$$((1 - (a\alpha + 1 - a), a\beta), (1 - (b\gamma + 1 - b), b\delta)) = ((a - a\alpha, a\beta), (b - b\gamma, b\delta)) \in \nu_\kappa$$

und daher $a \cdot (1 - \alpha, \beta) - b \cdot (1 - \gamma, \delta) \in U$, also (1) gilt. Wegen Lemma 2 und der Beziehungen (2)–(5) kann κ nur die Kongruenz $\kappa(I, U)$ sein. ■

Als nächstes wollen wir uns eine Eigenschaft der Kongruenzen auf der abgeschlossenen Teilmenge

$$L_{M;id} := \bigcup_{n \geq 1} \{f^n \in L_M \mid \exists a_0, \dots, a_n : f(\tilde{x}) = a_0 + \sum_{i=1}^n a_i x_i \wedge a_1 + a_2 + \dots + a_n = 1\}$$

(einer Menge sogenannter idempotenter Funktionen) von L_M überlegen. Wir betrachten dazu die bijektive Abbildung

$$\alpha : f^n \longrightarrow f^{n-1}$$

von $L_{M;id} := L_{M;id} \setminus L_M^1$ auf L_M , wobei $F(x_1, \dots, x_{n-1}) := f(0, x_1, \dots, x_{n-1})$ gelte.

Lemma 5 Die Abbildung α hat folgende Eigenschaften:

(a) Für die Operationen $\hat{\zeta}$, $\hat{\tau}$, $\hat{\Delta}$, $\hat{\nabla}$ und $\hat{\star}$, die definiert sind durch

$$(\hat{\zeta}f)(x_1, \dots, x_n) = f(x_1, x_3, x_4, \dots, x_n, x_2),$$

$$(\hat{\tau}f)(x_1, \dots, x_n) = f(x_1, x_3, x_2, \dots, x_n),$$

$$(\hat{\Delta})(x_1, \dots, x_{n-1}) = f(x_1, x_2, x_2, \dots, x_{n-1}) \text{ für } n \geq 3,$$

$$\zeta f^n = \tau f^n = \Delta f^n = f \text{ für } n = 2,$$

$$(\hat{\nabla})(x_1, \dots, x_{n+1}) = f(x_1, x_3, x_4, \dots, x_{n+1}) \text{ und}$$

$$(f\hat{\star}g)(x_1, \dots, x_{m+n-2}) = f(x_1, g(x_1, \dots, x_m), x_{m+1}, \dots, x_{m+n-2}) \quad (n, m \geq 2),$$

gilt $\alpha(\hat{\gamma}f) = \gamma(\alpha(f))$ für jedes $\gamma \in \{\zeta, \tau, \Delta, \nabla\}$ und $\alpha(f\hat{\star}g) = \alpha(f)\hat{\star}\alpha(g)$, d.h., die Algebra $\langle L_{M;\text{id}}; \hat{\zeta}, \hat{\tau}, \hat{\Delta}, \hat{\nabla}, \hat{\star} \rangle$ ist isomorph zur Algebra $\langle L_M; \zeta, \tau, \Delta, \nabla, \star \rangle$.

(b) Für jede nichtlineare abgeschlossene Teilmenge A von $L_{M;\text{id}}$ ist $\alpha(A)$ abgeschlossen, und es gilt

$$\alpha(A) \not\subseteq L_{M;\text{id}}, \quad A \subseteq \alpha(A) \text{ sowie } \alpha(A) \cap L_{M;\text{id}} = A.$$

Beweis Die Behauptung (a) prüft man leicht nach.

Bezeichne $A (\neq \emptyset)$ eine abgeschlossene Teilmenge von $L_{M;\text{id}}$. Aus (a) folgt dann unmittelbar, daß auch $\alpha(A)$ abgeschlossen ist. Wegen $e_1^2 \in A$ haben wir $\alpha(e_1^2) = c_0 \in \alpha(A)$, womit $\alpha(A) \not\subseteq L_{M;\text{id}}$ gilt. Sei nun $f \in A$. Dann haben wir $\nabla f \in A$ und folglich $\alpha(\nabla f) = f \in \alpha(A)$. Also ist $A \subseteq \alpha(A)$. $L_{M;\text{id}} \cap \alpha(A) \subseteq A$ und damit $L_{M;\text{id}} \cap \alpha(A) = A$ ergibt sich aus $\Delta(\alpha^{-1}f) = f \in A$ für jedes $f \in L_{M;\text{id}} \cap \alpha(A)$. ■

Satz 2 Seien $A (\neq \emptyset)$ eine Teilklasse von $L_{M;\text{id}}$, κ eine Kongruenz auf A und $\alpha(\kappa)$ die durch

$$(F, G) \in \alpha(\kappa) : \Longleftrightarrow (\alpha^{-1}F, \alpha^{-1}G) \in \kappa$$

definierte Relation auf $\alpha(A)$. Dann ist

(a) $\alpha(\kappa)$ eine Kongruenz auf A , und es gilt

(b) $\alpha(\kappa)_{|A} = \kappa$, d.h., die Kongruenzen auf A erhält man durch Beschränkung der Kongruenzen von $\alpha(A)$ auf A .

Beweis Da κ laut Voraussetzung eine Kongruenz auf A ist, ist κ auch mit den Operationen $\hat{\zeta}, \hat{\tau}, \hat{\Delta}, \hat{\nabla}$ und $\hat{*}$ verträglich. Hieraus und aus Lemma 5a) folgt unsere Behauptung (a).

Nach Lemma 5b) haben wir $A \subset \alpha(A)$, womit $\alpha(\kappa)_{/A}$ eine Kongruenz auf A ist. Seien nun f und g Funktionen aus A . Falls $(f, g) \in \kappa$, gilt $(\nabla f, \nabla g) \in \kappa$ und nach Definition von $\alpha(\kappa)$ ferner $(\alpha(\nabla f), \alpha(\nabla g)) = (f, g) \in \alpha(\kappa)_{/A}$, d.h. $\kappa \subseteq \alpha(\kappa)_{/A}$. Ist $(f, g) \in \alpha(\kappa)_{/A}$, so haben wir $(\alpha^{-1}f, \alpha^{-1}g) \in \kappa$. Da f und g aus $L_{M, id}$ sind, gilt $\alpha^{-1}f = \nabla f$ und $\alpha^{-1}g = g$. Folglich ist $(\nabla f, \nabla g) \in \kappa$ und $(\Delta(\nabla f), \Delta(\nabla g)) = (f, g) \in \kappa$, d.h. $\alpha(\kappa)_{/A} \subseteq \kappa$, woraus (b) folgt. ■

Aus den Sätzen 1 und 2 läßt sich mit Hilfe bekannter Eigenschaften von Gruppen, Körpern und Vektorräumen für die in [8], [12]–[15] untersuchten Klassen linearer Funktionen eine Reihe von Folgerungen ziehen. Hier sollen anschließend ohne Beweis nur einige *Beispiele* angegeben werden.

1. Falls $R = M$ und R ein endlicher Körper ist, hat nach [14] bis auf Isomorphie jede Teilklasse A von L_M mit $d \in A$ und $A \not\subseteq L_{M, id}$ die Struktur

$$L_W \cap Pol V \left(= \bigcup_{n \geq 1} \{f^n \mid \exists a_0 \in V \exists a_0, \dots, a_n \in W : f(\tilde{x}) = a_0 + a_1 x_1 + \dots + a_n x_n\} \right),$$

wobei W ein Unterkörper von R und V ein Untervektorraum von R über W ist. Für jede nichttriviale Kongruenz κ auf A existiert dann ein $I \in \{\{0\}, W\}$ sowie ein Untervektorraum U von V mit der Eigenschaft

$$\forall \alpha - \beta \in I \quad \forall a - b \in U : a \cdot \alpha - b \cdot \beta \in U,$$

und es gilt

$$\kappa = \bigcup_{n \geq 1} \left\{ \left(a_0 + \sum_{i=1}^n a_i x_i, b_0 + \sum_{i=1}^n b_i x_i \right) \mid a_0 - b_0 \in U \wedge \forall i \in \{1, \dots, n\} : a_i - b_i \in I \right\}.$$

Die einzige nichttriviale Kongruenz auf L_M ist insbesondere

$$\kappa_c := \{(f, g) \in \kappa_a \mid \exists a \in M : f = g + a\}.$$

2. Auch auf einer maximalen Klasse \mathbf{L} von quasilinearen Funktionen von P_k mit $k = p^m$,

p prim, $m \geq 1$, ist κ_c die einzige nichttriviale Kongruenz (siehe [12], [5]). Bekanntlich ([12]) ist \mathbf{L} isomorph zu einer gewissen Menge L_M , wobei R ein der Ring aller $m \times m$ -Matrizen über E_p und M der Vektorraum aller $m \times 1$ -Matrizen über E_p ist.

3. Wählt man $R = M = (E_k; + \bmod k, \cdot \bmod k)$, so gibt es auf L_M nur folgende Aritätskongruenzen:

$$\kappa_{s,t} := \bigcup_{n \geq 1} \left\{ (a_0 + \sum_{i=1}^n a_i x_i, b_0 + \sum_{i=1}^n b_i x_i) \mid s \mid a_0 - b_0 \wedge \forall i \in \{1, \dots, n\} : t \mid a_i - b_i \right\},$$

wobei s und t beliebige Teiler von k mit $s \nmid t$ bezeichnen.

Es sei noch bemerkt, daß $\text{Con } A$ nicht für jede Teilklasse von L_M eine endliche Menge ist. Findet man z.B. in $R \setminus \{0\}$ ein Element z mit $z^2 = 0$, so besitzt die abgeschlossene Menge

$$Z := \bigcup_{n \geq 1} \{f^n \in L_M \mid \exists a_1, \dots, a_n \in [\{0, z\}]_+ : f(\tilde{x}) = a_1 x_1 + \dots + a_n x_n\}$$

unendlich viele verschiedene Kongruenzen. Zwecks Beschreibung einiger dieser Kongruenzen bezeichne $r(f)$ die Anzahl der wesentlichen Variablen von $f \in Z$, und es sei

$$(f^n, g^n) \in \chi_i : \Longleftrightarrow f = g \vee (n = m \wedge r(f) \leq i \wedge r(g) \leq i) \\ (f, g \in Z, i \in N).$$

Man prüft leicht nach, daß für jedes $i \in N$ die Relation χ_i eine Kongruenz auf Z ist, wobei $|\text{Con } Z| \geq \aleph_0$ gilt.

Im folgenden Abschnitt soll nun gezeigt werden, daß es höchstens abzählbar viele Kongruenzen auf Teilklassen von L_M ($|M| < \aleph_0$) geben kann.

3. Ein Kriterium für den Nachweis der Abzählbarkeit von $\text{Con } A$ für gewisse $A \subseteq P_E$

Bekanntlich (siehe [4], S. 29) lassen sich Kongruenzrelationen als Grundmengen von Unter-algebren beschreiben. Für eine Unter-algebra A von $P_E = \langle P_E; \zeta, \tau, \Delta, \nabla, * \rangle$ wählt man dazu als neue Grundmenge $A \times A$ und definiert:

$$\begin{aligned}
\forall \alpha \in \{\zeta, \tau, \Delta, \nabla, *\} : \alpha(f, g) &:= (\alpha f, \alpha g), \\
(f, g) * (p, q) &:= (f * p, g * q), \\
\sigma(f, g) &:= (g, f) \text{ sowie} \\
\varrho((f, g), (p, q)) &:= \begin{cases} (f, q) & \text{für } g = p, \\ (f, g) & \text{sonst} \end{cases}
\end{aligned}$$

für beliebige $(f, g), (p, q) \in A \times A$. Die Kongruenzrelationen auf A sind dann genau die Grundmengen von Unteralgebren der Algebra $\langle A \times A; \zeta, \tau, \Delta, \nabla, *, \sigma, \varrho \rangle$, d.h., es gilt

$$\text{Con } A = \text{Sub}\langle A \times A; \zeta, \tau, \Delta, \nabla, *, \sigma, \varrho \rangle.$$

Speziell für die Aritätskongruenzen haben wir außerdem

$$\text{Con}_n A = \text{Sub}\left(\bigcup_{n \geq 1} A^n \times A^n; \zeta, \tau, \Delta, \nabla, *, \sigma, \varrho\right).$$

In [7] wurde nun folgender Satz bewiesen:

Satz 3 Sei $A = \langle A; F \rangle$ eine Algebra mit $|A| \leq \aleph_0$. Außerdem existiere auf A eine Halbordnungsrelation \leq mit folgenden drei Eigenschaften:

1. $f \leq g \implies [f]_F \subseteq [g]_F$,
2. jede Kette bez. \leq ist wohlgeordnet, d.h., jede Kette bez. \leq hat ein minimales Element,
3. jede Antikette bez. \leq besteht aus endlich vielen Elementen.

Dann hat A höchstens abzählbar viele verschiedene Unteralgebren. ■

Bezeichne \circ eine Operation über E mit den Eigenschaften

- a) $\forall a, b, c \in E : a \circ (b \circ c) = (a \circ b) \circ c,$
- b) $\forall a, b \in E : a \circ b = b \circ a,$
- c) $\exists r \geq 1 : (\forall a \in E : \underbrace{a \circ a \circ \dots \circ a}_{(1+r)-mal} = a).$

Die Eigenschaften a)–c) sind z.B. erfüllt, wenn (E, \circ) eine abelsche Gruppe oder $x \circ y := \max_{\omega}(x, y)$ bzw. $x \circ y := \min_{\omega}(x, y)$ (ω Halbordnungsrelation, für die max bzw. min existiert) ist.

Mit Hilfe von \circ läßt sich folgende Teilmenge von P_E definieren:

$$K_0 := \bigcup_{n \geq 1} \{f^n \in P_E \mid \exists f_0, f_1, \dots, f_n \in P_E^1 : f(x_1, \dots, x_n) = f_0(f_1(x_1) \circ f_2(x_2) \circ \dots \circ f_n(x_n))\}.$$

Die Menge K_0 ist i.allg. nicht abgeschlossen, enthält jedoch bei geeigneter Wahl von \circ unsere oben betrachteten Mengen L_M .

Analog zum Nachweis von $|Sub A| \leq \aleph_0$ für alle $A = [A] \subseteq K_0$ in [7] wollen wir abschließend folgenden Satz beweisen:

Satz 4 Eine Teilklasse A von P_E , für die eine Operation \circ mit den Eigenschaften a)–c) und $A \subseteq K_0$ existiert, besitzt höchstens abzählbar viele Kongruenzen.

Beweis Da nach [10] jede Teilklasse von P_E stets nur endlich viele Nichtaritätskongruenzen besitzt, genügt es, die Abzählbarkeit der Menge $Con \langle \bigcup_{n \geq 1} A^n \times A^n; \zeta, \tau, \Delta, \nabla, *, \sigma, \varrho \rangle$ zu beweisen. Dazu definieren wir auf $K_0^* := \bigcup_{n \geq 1} K_0^n \times K_0^n$ eine Halbordnungsrelation \leq , die die Bedingungen 1)–3) aus Satz 3 erfüllt, womit Satz 4 aus Satz 3 folgt.

Jedes Paar $(f, g) \in K_0^n \times K_0^n$ mit $f(\tilde{x}) = f_0(f_1(x_1) \circ \dots \circ f_n(x_n))$ und $g(\tilde{x}) = g_0(g_1(x_1) \circ \dots \circ g_n(x_n))$ ist durch die Angabe des Tupels $\varphi(f, g) := ((f_0, g_0), (f_1, g_1), \dots, (f_n, g_n))$

eindeutig bestimmt. Zwei Paare $(f, g), (p, q)$ mit $\varphi(f, g) = ((f_0, g_0), \dots, (f_n, g_n))$ und $\varphi(p, q) = ((p_0, q_0), \dots, (p_m, q_m))$ heißen genau dann \sim -äquivalent, wenn sie folgende drei Bedingungen erfüllen:

1. $(f_0, g_0) = (p_0, q_0)$,
2. $\{(f_1, g_1), \dots, (f_n, g_n)\} = \{(p_1, q_1), \dots, (p_m, q_m)\}$,
3. Wenn (f_i, g_i) in $((f_1, g_1), \dots, (f_n, g_n))$ genau u -mal vorkommt, dann findet man eine gewisse ganze Zahl v , so daß (f_i, g_i) in $((p_1, q_1), \dots, (p_m, q_m))$ genau $(u + v \cdot r)$ -mal auftritt (r bezeichne dabei die kleinste Zahl, für die o die Bedingung c) erfüllt).

Die Faktormenge $K_{0/\sim}^*$ ist offensichtlich endlich. Ihre Elemente bezeichnen wir mit K_1, \dots, K_t . Auf den Mengen K_1, \dots, K_t läßt sich eine Halbordnungsrelation \leq wie folgt definieren:

Für $(f^n, g^n), (p^m, q^m) \in K_i$ ($i = 1, \dots, t$) mit $\varphi(f, g) = ((f_0, g_0), \dots, (f_n, g_n))$ und $\varphi(p, q) = ((p_0, q_0), \dots, (p_m, q_m))$ schreiben wir $(f, g) \leq (p, q)$ genau dann, wenn für jedes (f_i, g_i) , das in $((f_1, g_1), \dots, (f_n, g_n))$ j_1 -mal und in $((p_1, q_1), \dots, (p_m, q_m))$ j_2 -mal auftritt, stets $j_1 \leq j_2$ gilt. Diese Halbordnungsrelation \leq erfüllt die Bedingung 1) von Satz 3, da man, falls $(f, g) \leq (p, q)$ ist, (f, g) aus (p, q) durch Anwendung der Operationen ζ, τ, Δ wegen der Eigenschaft c) von o erhält. Eigenschaft 2) aus Satz 3 gilt hier offenbar ebenfalls. Es bleibt also noch zu zeigen, daß jede Antikette in der Menge K_0^* bez. \leq nur aus endlich vielen Elementen besteht. Auf $P_E^1 \times P_E^1$ läßt sich eine totale Ordnung einführen. Für $\varphi(f, g) = ((f_0, g_0), \dots, (f_n, g_n))$ vereinbaren wir bez. dieser Ordnung, mit (f_{i_0}, g_{i_0}) das kleinste Element aus $\{(f_1, g_1), \dots, (f_n, g_n)\}$, mit (f_{i_1}, g_{i_1}) das kleinste Element aus $\{(f_1, g_1), \dots, (f_n, g_n)\} \setminus \{(f_{i_0}, g_{i_0})\}$, \dots und mit (f_{i_s}, g_{i_s}) das größte Element aus $\{(f_1, g_1), \dots, (f_n, g_n)\}$ zu bezeichnen. Wenn (f_{i_j}, g_{i_j}) in $\{(f_1, g_1), \dots, (f_n, g_n)\}$ genau b_j -mal auftritt, $j = 0, 1, \dots, s$, so sei $\alpha(f, g) := \langle b_0, b_1, \dots, b_s \rangle$. Ferner gelte $D_i := \{\alpha(f, g) \mid (f, g) \in K_i\}$, $i = 1, 2, \dots, t$. Wir vereinbaren, genau dann $\alpha(f, g) \leq \alpha(p, q)$ zu setzen, wenn $(f, g) \leq (p, q)$ ist. Die halbgeordneten

Mengen K_i und D_i sind zueinander isomorph. Im Beweis von Satz 2 aus [7] wurde nun gezeigt, daß eine beliebige Teilmenge C von D_i , die aus paarweise unvergleichbaren Elementen besteht, endlich ist. Also erfüllt \leq auch die Bedingung 3) von Satz 3. ■

Literatur

- [1] Denecke, K., und Lau, D.: *Kongruenzen auf Klons und vollinvariante Kongruenzen relativ freier Algebren II*. Rostock. Math. Kolloq. **29**, 4–20 (1986)
- [2] Gorlov, V.V.: *Über Kongruenzen auf abgeschlossenen Postschen Klassen* (in Russisch). Mat. Zametki **13**, 725–734 (1973)
- [3] Gorlov, V.V., and Lau, D.: *Congruences on closed sets of selfdual functions in many-valued logics and on closed sets of linear functions in prime-valued logics*. Közl.-MTA Számítástech. Automat. Kutató Int. Budapest **29**, 31–39 (1983)
- [4] Ihringer, T.: *Allgemeine Algebra*. Stuttgart 1988
- [5] Lau, D.: *Congruences on closed sets of k -valued logic*. Colloq. Math. Soc. Janos Bolyai. **28**, 417–440 (1981)
- [6] Lau, D.: *Über die Anzahl von abgeschlossenen Mengen linearer Funktionen der n -wertigen Logik*. Elektron. Informationsverarb. Kybernet. EIK **14**, 561–563 (1978)
- [7] Lau, D.: *Ein Kriterium für den Nachweis der Abzählbarkeit gewisser Teilverbände des Verbandes der abgeschlossenen Mengen von Funktionen der k -wertigen Logik*. Rostock. Math. Kolloq. **30**, 11–18 (1986)
- [8] Lau, D.: *Über abgeschlossene Mengen linearer Funktionen in mehrwertigen Logiken*. J. Inform. Process. Cybernet. EIK **24**, 367–381 (1988)

- [9] **Mal'cev, A.I.:** *Iterative Algebren und Postsche Mannigfaltigkeiten* (in Russisch). Algebra i Logika **5**, 5–24 (1966)
- [10] **Mal'cev, A.I.:** *Über Kongruenzen auf Unterhalbgebren iterativer Algebren* (in Russisch). Metody Diskret. Analiz **29**, 40–52 (1976)
- [11] **Pöschel, R., und Kaluznin, L.A.:** *Funktionen- und Relationenalgebren*. Berlin 1979
- [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] **Szendrei, A.:** *On closed sets of linear operations over a finite set of square-free cardinality*. Elektron. Informationsverarb. Kybernet. EIK **14**, 547–559 (1978)
- [14] **Szendrei, A.:** *On closed classes of quasilinear functions*. Czechoslovak. Mat. J. **30**, 489–509 (1980)
- [15] **Szabo, L., and Szendrei, A.:** *Slupecki-type criteria for quasilinear functions over a finite dimensional vector space*. Elektron. Informationsverarb. Kybernet. EIK **17**, 601–611 (1981)

eingegangen: 12. September 1989

revidierte Fassung: 15. Februar 1990

Verfasser:

Dr. D.Lau
Universität Rostock
Fachbereich Mathematik
Universitätsplatz 1
O-2500 Rostock
Germany

ROGER LABAHN; SERGEJ L. BESRUKOV

Asymptotic solution of a special equation with binomial coefficients

Let any positive integer p and any positive real number s be given parameters. For positive integers k , our aim is to solve the equation

$$s \cdot \sum_{i=0}^k \binom{n}{i} = \binom{n}{k+p} \quad (1)$$

in the real variable n with $n > k + p - 1$.

It is easy to prove that there is a unique solution $n(k)$ (see Section 1). Our main result is the following

Theorem $n(k) \sim c \cdot k$, where $c - 1$ is the positive root of $x^p - x^{p-1} - s$.

After proving this in Section 2, we shall give an application of the case $s = 1$, $p = 2$ to the subword poset $B_{2,m}$ in Section 3.

This paper continues our investigations presented in [6] where we dealt with the case $s = p = 1$ only. While the asymptotics $n(k) \sim 3k$ if $k \rightarrow \infty$ had been shown by Berg [1], we could even prove that $|n(k) - 3k| \rightarrow 0$ if $k \rightarrow \infty$. Moreover, in [6] we gave the final solution to the according "discrete" problem, i.e. determine that (uniquely determined) value $N(k)$ of n for which $s \cdot \sum_{i=0}^k \binom{n}{i} \geq \binom{n}{k+p}$ but $s \cdot \sum_{i=0}^k \binom{n+1}{i} < \binom{n+1}{k+p}$. It holds

$$N(k) = \begin{cases} 3k & \text{if } 1 \leq k \leq 4, \\ 3k - 1 & \text{if } k \geq 5. \end{cases}$$

1. Uniqueness

For arbitrarily fixed k , we consider $q(n) := \sum_{i=0}^k \binom{n}{i} / \binom{n}{k+p}$ for $n > k + p - 1$. Since each summand is of the form

$$\frac{(i+1) \cdot \dots \cdot (k+p)}{(n-i) \cdot \dots \cdot [n - (k+p-1)]},$$

$q(n)$ is monotone decreasing and $q(n) \rightarrow 0$ if $n \rightarrow \infty$.

Moreover, if $n \rightarrow k+p-1$ then $\binom{n}{k+p} = \binom{n}{k+p-1} \cdot [n-(k+p-1)]/(k+p) \rightarrow 0$ but for $n \leq k+p$, $\sum_{i=0}^k \binom{n}{i} \leq \sum_{i=0}^{k+p} \binom{k+p}{i} = 2^{k+p}$, i.e. $\sum_{i=0}^k \binom{n}{i}$ is bounded and $q(n) \rightarrow \infty$ if $n \rightarrow k+p-1$. Because $q(n)$ is continous in the considered interval, there is exactly one value for n with $q(n) = 1/s$. This is the unique solution $n(k)$ of (1).

2. Asymptotic behaviour

We have $\sum_{i=0}^k \binom{2k+1}{i} = 2^{2k}$, and by Stirling's formula:

$$\binom{2k+1}{k+p} < \binom{2k+1}{k} < 2 \cdot \binom{2k}{k} < 2^{2k+1}/\sqrt{\pi k}.$$

Hence, $q(2k+1) > \sqrt{\pi k}/2$, i.e. for sufficiently large k , $q(2k+1) > 1/s$. Because $q(n)$ is monotone decreasing in n this implies that for sufficiently large k ,

$$n(k) \geq 2k+1. \quad (2)$$

Next we prove by taking into consideration only a certain fixed number of the last summands of $\sum_{i=0}^k \binom{n}{i}$ that $n(k)$ must be essentially larger than $2k+1$. Suppose that $n(k) \sim 2k$, for any subsequence of $\{k\}$. Then for the constant value $m := \lceil 1/s \rceil$ and $j = 0, 1, \dots, m$,

$$\begin{aligned} \binom{n}{k-j} &= \frac{k+p}{n-(k+p)+1} \cdot \dots \cdot \frac{k-j+1}{n-(k-j)} \cdot \binom{n}{k+p} \\ &= \frac{1+p/k}{n/k-1-(p-1)/k} \cdot \dots \cdot \frac{1-(j-1)/k}{n/k-1+j/k} \cdot \binom{n}{k+p} \sim \binom{n}{k+p}, \end{aligned}$$

because $p=\text{const.}$, $j=\text{const.}$, and we have assumed that for the considered subsequence of $\{k\}$, $n/k \rightarrow 2$. Note that there are at most $m+p$, i.e. finitely many factors.

This implies $\sum_{i=0}^k \binom{n}{i} \geq \sum_{j=0}^m \binom{n}{k-j} \sim (m+1) \binom{n}{k+p}$ because $m=\text{const.}$ Hence for the considered values $n(k)$, $q(n(k)) \rightarrow m+1 > 1/s$. But on the other hand we know from (1) by the definition of $n(k)$ that $q(n(k)) = 1/s$ for all k . This contradiction means:

$$\text{There is no subsequence of } \{k\} \text{ such that } n(k)/k \rightarrow 2. \quad (3)$$

Now we can apply an interesting estimation given in [4]:

For $n \geq 5$, $k \geq 1$, $n > 2k+1$,

$$\sum_{i=0}^k \binom{n}{i} = \binom{n}{k} \cdot \frac{n-k}{n-2k-1} \cdot \frac{1}{1+\beta}, \quad \text{where } 0 \leq \beta \leq \frac{n \cdot \ln 2\pi n}{2(n-2k-1)^2}. \quad (4)$$

Let us investigate the behaviour of the upper bound for β in (4). Because of (3) for $n = n(k)$, $\left(1 - 2\frac{k}{n} - \frac{1}{n}\right)^2$ does not tend to 0 if $n \rightarrow \infty$, and consequently $\beta \rightarrow 0$ if $k \rightarrow \infty$. Note that because of (2), $k \rightarrow \infty$ implies $n \rightarrow \infty$. Therefore in our situation, (4) means

$$\sum_{i=0}^k \binom{n}{i} \sim \frac{n-k}{n-2k} \cdot \binom{n}{k} \quad (5)$$

and together with (1) we get for $n = n(k)$

$$s \cdot \frac{n-k}{n-2k} \cdot \binom{n}{k} \sim \binom{n}{k+p}. \quad (6)$$

On the other hand for $k \rightarrow \infty$,

$$\begin{aligned} \frac{n-2k}{n-k} \binom{n}{k+p} &= \frac{n-2k}{n-k} \cdot \frac{n-(k+p)-1}{k+p} \cdot \dots \cdot \frac{n-k}{k+1} \cdot \binom{n}{k} \\ &= \frac{n/k-2}{n/k-1} \cdot \frac{n/k-1-(p-1)/k}{1+p/k} \cdot \dots \cdot \frac{n/k-1}{1+1/k} \cdot \binom{n}{k} \\ &\sim \left(\frac{n}{k}-2\right) \cdot \left(\frac{n}{k}-1\right)^{p-1} \cdot \binom{n}{k}. \end{aligned}$$

Comparing this with (6) we finally have for $k \rightarrow \infty$,

$$\left(\frac{n(k)}{k}-2\right) \cdot \left(\frac{n(k)}{k}-1\right)^{p-1} = \left(\frac{n(k)}{k}-1\right)^p - \left(\frac{n(k)}{k}-1\right)^{p-1} \rightarrow s. \quad (7)$$

Obviously, $x^p - x^{p-1} = x^{p-1}(x-1)$ is negative for $0 < x < 1$, and monotone increasing and positive for $x \geq 1$. Therefore there is exactly one positive solution of $x^p - x^{p-1} = s$ for any $s > 0$. This is greater than 1 and we may finally conclude from (7) that $n(k) \sim ck$ where $(c-1)^p - (c-1)^{p-1} = s$.

Remark We should sketch how to get the asymptotic formula (5) for $\sum_{i=0}^k \binom{n}{i}$ without using Fjodorov's formula (4). On the one hand it is well known that

$$\sum_{i=0}^k \binom{n}{i} \leq \frac{n-k}{n-2k} \cdot \binom{n}{k} \quad (\text{see e.g. [5], p.276}),$$

on the other hand we have for any constant $m \leq k$,

$$\begin{aligned}
\sum_{i=0}^k \binom{n}{i} &\geq \sum_{j=0}^m \binom{n}{k-j} = \binom{n}{k} \cdot \sum_{j=0}^m \frac{k}{n-k+1} \cdot \frac{k-1}{n-k+2} \cdots \frac{k-(j-1)}{n-k+j} \\
&\sim \binom{n}{k} \cdot \sum_{j=0}^m \left(\frac{k}{n-k} \right)^j = \binom{n}{k} \cdot \frac{1}{1 - k/(n-k)} \cdot \left[1 - \left(\frac{k}{n-k} \right)^{m+1} \right] \\
&= \frac{n-k}{n-2k} \cdot \binom{n}{k} \cdot \left[1 - \left(\frac{1}{n/k-1} \right)^{m+1} \right].
\end{aligned}$$

Because of (2) and (3) for sufficiently large k , $1/(n/k-1) < \lambda < 1$. Consequently for any real $\varepsilon > 0$, we can find an constant m such that for sufficiently large $k > m$

$$1 - \left(\frac{1}{n/k-1} \right)^{m+1} > 1 - \varepsilon.$$

This means: By taking into consideration enough large" summands of $\sum_{i=0}^k \binom{n}{i}$ we can prove an asymptotic lower bound as closed to the upper bound as we wish. Therefore

$$\sum_{i=0}^k \binom{n}{i} \sim \frac{n-k}{n-2k} \cdot \binom{n}{k}$$

if n is asymptotically greater than $2k$.

3. An application

A binary word $\underline{a} = a_1, a_2, \dots, a_r$ of length r is said to be a subword of a binary word $\underline{b} = b_1, b_2, \dots, b_t$ of length $t \geq r$ iff there are indices $1 \leq i_1 < i_2 < \dots < i_r \leq t$ such that $a_j = b_{i_j}$ for $j = 1, 2, \dots, r$. For fixed $m \geq 2$, the set of all subwords of $u_{2,m} := \underbrace{010101\dots}_{m \text{ letters}}$

ordered by " $\underline{a} < \underline{b}$ iff \underline{a} is a subword of \underline{b} " forms the poset $B_{2,m}$. It was introduced by Burosch at al. [3]. The levels of $B_{2,m}$ are formed by subwords of a given length and [3] it has been shown that the r -th level $B_{2,m}^{(r)}$ of $B_{2,m}$ contains exactly $\sum_{i=0}^{m-r} \binom{r}{i}$ words. We shall use the previous results to determine the level $r(m)$ with the largest number of elements asymptotically. For this we consider

$$\begin{aligned}
|B_{2,m}^{r+1}| - |B_{2,m}^r| &= \sum_{i=0}^{m-r-1} \binom{r+1}{i} - \sum_{i=0}^{m-r} \binom{r}{i} \\
&= \sum_{i=1}^{m-r-1} \binom{r}{i-1} - \binom{r}{m-r} = \sum_{i=0}^{m-r-2} \binom{r}{i} - \binom{r}{m-r}.
\end{aligned}$$

Setting $s = 1$, $p = 2$ we know from our Theorem that there is exactly one value $r(m)$ such

that this difference is non-negative for all $r \leq r(m)$ and negative for all $r > r(m)$, i.e. $|B_{2,m}^{(r)}|$ is unimodal and the level $r(m)$ is one of those with the largest number of elements. Moreover, setting $n = r$, $k = m - r - 2$ it follows that $r(m) \sim c \cdot (m - r(m) - 2)$ where $c - 1$ is the positive root of $x^2 - x - 1$, i.e. $c - 1 = \tau := \frac{1}{2}(1 + \sqrt{5})$. Using $\tau^2 - \tau = 1$ we finally get

$$r(m) \sim \frac{c}{1+c} \cdot m = \frac{\tau+1}{\tau+2} \cdot m = \frac{\tau+2}{5} \cdot m = \frac{1}{10}(5 + \sqrt{5}) \cdot m = 0.72336 \dots \cdot m.$$

Clearly, one can use (4) to find bounds of $n(k)$ in the case $s = 1$, $p = 2$ in our original problem. Due to (4) both bounds differ in a term containing $\ln k$ and we conjecture that the upper bound

$$n(k) \leq \frac{3}{2}k + 1 + \frac{1}{2}\sqrt{5k^2 + 12k + 8}$$

is more realistic. Moreover, basing on some calculations we should formulate the following conjecture about the solution $N(k)$ of the according "discrete" problem:

Conjecture For $s = 1$, $p = 2$, $\lfloor \tau^2 k \rfloor \leq N(k) \leq \lceil \tau^2 k \rceil$.

Provided that this conjecture is true we could improve the result for $r(m)$ to

$$-1.7236 \dots = -\left(1 + \frac{\tau+2}{5}\right) \leq r(m) - \frac{\tau+2}{5} \cdot m \leq \frac{3\tau+1}{5} = 1.1708 \dots$$

Acknowledgement We are indebted to Prof. L. Berg for his helpful comments to the first version of this paper.

References

- [1] **Berg, L.:** *Das asymptotische Verhalten einer speziellen Funktion von Binomialkoeffizienten.* Rostock. Math. Kolloq. **31**, 93–94 (1987)
- [2] **Besrukov, S.L., and Labahn, R.:** *On an inequality with binomial coefficients*(in Russian). Proc. 8. Soviet Conf. Theor. Cyb. 41–42, Gorki 1988
- [3] **Burosch, G., Gronau, H.–D.O.F., Laborde, J.–M., and Warnke, I.:** *The poset of subwords of an n -ary word.* (submitted to Discrete Math.)
- [4] **Fjodorov, S.A.:** *About one possibility to estimate volumes of balls in binary Hamming spaces* (in Russian). Dokl. Akad. Nauk. SSSR **286**, 1360–1364 (1984)
- [5] **Gavrilov, G.P., and Saposhenko, A.A.:** *Problems in Discrete Mathematics.* (in Russian). Moscow 1977
- [6] **Labahn, R.:** *More on a relation between binomial coefficients.* Rostock. Math. Kolloq. **34**, 21–27 (1988)
- [7] **Labahn, R.** *On sums of binomial coefficients.* Ernst–Moritz–Arndt–Universität Greifswald, Preprint–Reihe Mathematik **25**, 54–55 (1990)

received May 2, 1990

revised version May 22, 1990

Authors

Dr. R. Labahn
Universität Rostock
Fachbereich Mathematik
Universitätsplatz 1
O–2500 Rostock
Germany

Dr. S.L. Besrukov
Moscow State University
Department of Mathematics
Moscow
USSR

LOTHAR BERG

Asymptotic expansions of a special integral

Usually, one obtains an asymptotic expansion of a parameter integral

$$J(s) = \int_0^{\infty} G(s, t) dt$$

by expanding the integrand into a suitable series, i.e.

$$G(s, t) = \sum_{n=0}^{\infty} G_n(s, t),$$

and applying termwise integration. This method fails, if the arising integrals

$$\int_0^{\infty} G_n(s, t) dt \tag{1}$$

do not exist. Under certain assumptions it is possible, cf. [1], to replace the terms (1) by the limits

$$\lim_{p \rightarrow +\infty} \int_0^{\infty} e^{-pt} G_n(s, t) dt.$$

Another possibility to overcome the just mentioned deficiency consists in a splitting of the integral in such a way that the separate parts of the integral can be expanded into suitable series. Here, we demonstrate this method with respect to the special integral

$$J_{\alpha, \beta}(s) = \int_0^{\infty} \frac{dt}{(1+t)^{\alpha}(s+t)^{\beta}} \tag{2}$$

with $s > 0$ and $\alpha + \beta > 1$, for which R. Wong [2] has proved by means of distributions that for $s \rightarrow \infty$

$$J_{\frac{1}{3},1}(s) \approx \frac{2\pi}{\sqrt{3}} \sum_{n=0}^{\infty} (-1)^n \binom{-\frac{1}{3}}{n} s^{-n-\frac{1}{3}} - \sum_{n=1}^{\infty} \frac{3^n(n-1)!}{2 \cdot 5 \dots (3n-1)} s^{-n}. \quad (3)$$

In [2] it is also shown that this expansion cannot be derived by partial integration.

Without loss of generality we assume that $\alpha < 1$, because by partial integrations of the kind

$$\int_0^{\infty} \frac{dt}{(1+t)^{\alpha}(s+t)^{\beta}} = \frac{1}{(\alpha-1)s^{\beta}} - \frac{\beta}{\alpha-1} \int_0^{\infty} \frac{dt}{(1+t)^{\alpha-1}(s+t)^{\beta+1}},$$

so far as $\alpha > 1$, we can reduce the general case to the foregoing one or to the case

$$\int_0^{\infty} \frac{dt}{(1+t)(s+t)^{\beta}} = \beta \int_0^{\infty} \frac{\ln(1+t)}{(s+t)^{\beta+1}} dt,$$

which can be treated in an analogous way.

Substituting $t = x - 1$, we obtain from (2) for $s > 1$

$$J_{\alpha,\beta}(s) = \int_1^{\infty} \frac{dx}{x^{\alpha}(s-1+x)^{\beta}} = \int_0^{\infty} \frac{dx}{x^{\alpha}(s-1+x)^{\beta}} - \int_0^1 \frac{dx}{x^{\alpha}(s-1+x)^{\beta}}. \quad (4)$$

The first integral on the right-hand side can be expressed by

$$\int_0^{\infty} \frac{dx}{x^{\alpha}(s-1+x)^{\beta}} = (s-1)^{1-\alpha-\beta} \int_0^{\infty} \frac{dx}{x^{\alpha}(1+x)^{\beta}} = \frac{\Gamma(1-\alpha)\Gamma(\alpha+\beta-1)}{\Gamma(\beta)} (s-1)^{1-\alpha-\beta}. \quad (5)$$

In view of

$$\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{2}{3}\right) = \frac{2\pi}{\sqrt{3}}$$

this yields the first sum in (3). The second integral in (4) has the expansion

$$\int_0^1 \frac{dx}{x^\alpha(s-1+x)^\beta} = \sum_{n=0}^{\infty} s^{-n-\beta} \binom{n+\beta-1}{n} \int_0^1 x^{-\alpha}(1-x)^n dx, \quad (6)$$

and according to

$$\int_0^1 x^{-\alpha}(1-x)^n dx = \frac{\Gamma(1-\alpha)n!}{\Gamma(n+2-\alpha)} = \frac{n!}{(n+1-\alpha)\dots(1-\alpha)}$$

this yields the second sum in (3). However, using for $s > 2$ the expansion

$$(s-1+x)^{-\beta} = \sum_{n=0}^{\infty} \binom{n+\beta-1}{n} \frac{(-x)^n}{(s-1)^{n+\beta}},$$

we obtain instead of (6) the simpler result

$$\int_0^1 \frac{dx}{x^\alpha(s-1+x)^\beta} = \sum_{n=0}^{\infty} \binom{n+\beta-1}{n} \frac{(-1)^n}{n+1-\alpha} (s-1)^{-n-\beta}, \quad (7)$$

which together with (5) and (4) gives the wanted expansion for the integral (2). All these expansions are both convergent and asymptotical for $s \rightarrow \infty$.

The integral (2) can also be written in another way, namely, by means of substitution $x = (s-1)/u$ we obtain

$$J_{\alpha,\beta}(s) = (s-1)^{1-\alpha-\beta} \int_0^{s-1} u^{\alpha+\beta-2} (1+u)^{-\beta} du.$$

This implies that $J_{\alpha,\beta}(s)$ is the solution of the differential equation

$$J'_{\alpha,\beta}(s) = \frac{1-\alpha-\beta}{s-1} J_{\alpha,\beta}(s) + \frac{1}{s^\beta(s-1)} \quad (8)$$

with the initial value

$$J_{\alpha,\beta}(1) = \int_0^\infty \frac{dt}{(1+t)^{\alpha+\beta}} = \frac{1}{\alpha+\beta-1}.$$

Since $s=1$ is a singular point of the differential equation, $J_{\alpha,\beta}(s)$ can also be characterized as the unique solution of (8) being bounded at $s=1$.

References

- [1] **Berg, L.:** *Asymptotische Entwicklungen mit Hilfe von Neutritzen.* Arch. Math. **14**, 162–171 (1963)
- [2] **Wong, R.:** *Asymptotic Approximations of Integrals.* San Diego 1989

received: May 7, 1990

Author

Prof. Dr. L. Berg
Universität Rostock
Fachbereich Mathematik
Universitätsplatz 1
O-2500 Rostock
Germany

MUHARREM BERISHA

On coefficients of double lacunary trigonometric series

1. Introduction

The dependence of Fourier coefficients from the behaviour of the univariate modulus of continuity is examined in the papers of Konjushkov [1], Aljančić and Tomić [8]. They have proved the following result: *If the Fourier coefficients a_n of a function $f(x) \in L_p$ decrease and $a_n > 0$, then the inequality*

$$n^{1-\frac{1}{p}} a_n \leq M_p \omega^{(p)}\left(f, \frac{\pi}{2n}\right)$$

holds, where

$$\omega^{(p)}\left(f, \frac{\pi}{2n}\right) = \sup_{0 \leq h \leq \frac{\pi}{2n}} \|f(\circ + h) - f(\circ - h)\|_{L_p}.$$

In [4] an estimate for the modulus of smoothness from below and above via Fourier coefficients is given.

In [2], [3], and [5] Fourier coefficients of a bivariate function from L_p are estimated.

In this paper we estimate the modulus of smoothness from below and above by Fourier coefficients of double lacunary trigonometric series, and we present necessary and sufficient conditions in terms of Fourier coefficients which ensure that the function $f(x, y)$ belongs to the Nikolski class of the type $S^0 H_p^{r_1, r_2}$. In fact, we prove the following results:

Theorem 1 *Let $f(x, y) \in L_p$ for fixed p_i ($i = 1, 2$) from the interval $1 \leq p_i < \infty$ and*

$$f(x, y) \sim \sum_{\nu=1}^{\infty} \sum_{\mu=1}^{\infty} c_{\nu\mu} \cos \nu x \cos \mu y,$$

where

$$c_{\nu\mu} = \begin{cases} a_{l_1 l_2}, & \text{if } \nu = 2^{l_1} \text{ and } \mu = 2^{l_2}, \\ 0 & \text{elsewhere,} \end{cases}$$

provided that l_1 and l_2 are non-negative integers. Then

$$A_5 B(c_{\nu\mu}, k_1, k_2, n, m) \leq \omega_{k_1 k_2}\left(f, \frac{1}{n}, \frac{1}{m}\right)_{\bar{p}} \leq A_6 B(c_{\nu\mu}, k_1, k_2, n, m),$$

where the constants A_5 and A_6 do not depend on $f(x, y)$, n and m .

Theorem 2 *The function $f(x)$, which Fourier series is*

$$f(x, y) \sim \sum_{\nu=1}^{\infty} \sum_{\mu=1}^{\infty} c_{\nu\mu} \cos \nu x \cos \mu y$$

with Fourier coefficients

$$c_{\nu\mu} = \begin{cases} a_{l_1 l_2}, & \text{if } \nu = 2^{l_1} \text{ and } \mu = 2^{l_2}, \\ 0 & \text{elsewhere} \end{cases}$$

(l_1 and l_2 are non-negative integers), belongs to the class $S^0 H_p^{r_1 r_2}$ if and only if the Fourier coefficients of this function satisfy the condition

$$c_{\nu\mu} = O\left(\frac{1}{\nu^{r_1} \mu^{r_2}}\right). \quad (1)$$

2. Preliminaries

Let $f(x, y)$ be a bivariate function, 2π -periodic in each variable. Let us define, that $f(x, y) \in L_{\vec{p}}$, $1 \leq p_i < \infty$ ($i = 1, 2$), if f is measurable and

$$\|f\|_{\vec{p}} \leq c < \infty,$$

where

$$\|f\|_{\vec{p}} = \left\{ \int_0^{2\pi} \left[\int_0^{2\pi} |f(x, y)|^{p_1} dx \right]^{p_2/p_1} dy \right\}.$$

We say that $f \in L_{\vec{p}}^0$, if the following conditions are satisfied:

1. $f \in L_{\vec{p}}$,
2. $\int_0^{2\pi} f(x, y) dx = 0$ for almost every y , and $\int_0^{2\pi} f(x, y) dy = 0$ for almost every x .

We use the notation $\omega_{k_1, k_2}(f, t_1, t_2)_{\vec{p}}$ for the mixed modulus of smoothness (in $L_{\vec{p}}$) of order k_1 in x and k_2 in y of the function $f \in L_{\vec{p}}$, i.e.

$$\omega_{k_1 k_2}(f, t_1, t_2)_{\vec{p}} = \sup_{|h_1| \leq t_1, |h_2| \leq t_2} \|\Delta_{h_1 h_2}^{k_1 k_2} f(x, y)\|_{\vec{p}},$$

where

$$\Delta_{h_1 h_2}^{k_1 k_2} f(x, y) = \sum_{\nu=0}^{k_1} \sum_{\mu=0}^{k_2} (-1)^{k_1+k_2-\nu-\mu} \binom{k_1}{\nu} \binom{k_2}{\mu} f(x + \nu h_1, y + \mu h_2).$$

We write $\omega_{k_2}(f, t_2)_{\vec{p}}$ for the modulus of smoothness (in $L_{\vec{p}}$) of order k_2 in y of the function $f \in L_{\vec{p}}$, i.e.

$$\omega_{k_2}(f, t_2)_{\vec{p}} = \sup_{|h_2| \leq t_2} \|\Delta_{h_2}^{k_2} f(x, y)\|_{\vec{p}},$$

where

$$\Delta_{h_2}^{k_2} f(x, y) = \sum_{\mu=0}^{k_2} (-1)^{k_2-\mu} \binom{k_2}{\mu} f(x, y + \mu h_2).$$

Let $Y_{nm}(f)_{\vec{p}}$ be the best angular approximation (in $L_{\vec{p}}$) of the function $f \in L_{\vec{p}}$, i.e. the approximation via the sum of trigonometric polynomials $T_{n\infty}(x, y)$ of order n in x and trigonometric polynomials $T_{\infty m}(x, y)$ of order m in y :

$$Y_{nm}(f)_{\vec{p}} = \inf_{T_{n\infty}, T_{\infty m}} \|f(x, y) - [T_{n\infty}(x, y) + T_{\infty m}(x, y)]\|_{\vec{p}},$$

where

$$T_{n\infty}(x, y) = \sum_{l_1=0}^n (a_{l_1}(y) \cos l_1 x + b_{l_1}(y) \sin l_1 x),$$

$$T_{\infty m}(x, y) = \sum_{l_2=0}^m (c_{l_2}(x) \cos l_2 y + d_{l_2}(x) \sin l_2 y),$$

where

$$T_{n\infty}(x, y) \in L_{\vec{p}} \quad \text{and} \quad T_{\infty m}(x, y) \in L_{\vec{p}}.$$

We write $E_{\infty m}(f)_{\vec{p}}$ for the best approximation (in $L_{\vec{p}}$) of the function $f \in L_{\vec{p}}$ via trigonometric polynomials of order m in x , i.e.

$$E_{\infty m}(f)_{\vec{p}} = \inf_{T_{\infty m}} \|f(x, y) - T_{\infty m}(x, y)\|_{\vec{p}},$$

where $T_{\infty m} \in L_{\vec{p}}$.

Let $f \in L_{\vec{p}}$ for fixed p_i ($i = 1, 2$) from the interval $1 < p_i < \infty$ and

$$f(x, y) \sim \sum_{\nu=1}^{\infty} \sum_{\mu=1}^{\infty} c_{\nu\mu} \cos \nu x \cos \mu y,$$

where

$$c_{\nu\mu} = \begin{cases} a_{l_1 l_2}, & \text{if } \nu = 2^{l_1} \text{ and } \mu = 2^{l_2}, \\ 0, & \text{elsewhere,} \end{cases}$$

provided that l_1 and l_2 are non-negative integers.

Denote by $s_{n\infty}(f)$ the partial sum of order n in x of the Fourier series of the function $f(x, y)$, i.e.

$$s_{n\infty}(f) \sim \sum_{\nu=1}^n \sum_{\mu=1}^{\infty} c_{\nu\mu} \cos \nu x \cos \mu y.$$

Analogously we define

$$\begin{aligned} s_{\infty m} &\sim \sum_{\nu=1}^{\infty} \sum_{\mu=1}^m c_{\nu\mu} \cos \nu x \cos \mu y, \\ s_{nm}(f) &\sim \sum_{\nu=1}^n \sum_{\mu=1}^m c_{\nu\mu} \cos \nu x \cos \mu y. \end{aligned}$$

Using the notation

$$\varphi_{nm}(x, y) = f(x, y) - s_{n\infty}(f) - s_{\infty m}(f) + s_{nm}(f),$$

one obtains

$$\varphi_{nm}(x, y) \sim \sum_{\nu=n+1}^{\infty} \sum_{\mu=m+1}^{\infty} c_{\nu\mu} \cos \nu x \cos \mu y.$$

We say that $f \in S^0 H_{\vec{p}}^{r_1 r_2}$, if

1. $f \in L_{\vec{p}}$,
2. $\omega_{k_1 k_2}(f, \delta_1, \delta_2)_{\vec{p}} \leq c \delta_1^{k_1} \delta_2^{k_2}$,

where $k_1 > r_1$ and $k_2 > r_2$.

In order to prove the basic results of this paper, we need the following lemmas.

Lemma 1 ([6]) *If the conditions $\frac{n_{l_1+1}}{n_{l_1}} \geq \lambda_1 > 1$, $\frac{n_{l_2+1}}{n_{l_2}} \geq \lambda_2 > 1$ hold and the series*

$$M = \left(\sum_{l_1=1}^{\infty} \sum_{l_2=1}^{\infty} a_{l_1 l_2}^2 \right)^{\frac{1}{2}}$$

converges, then

$$\sum_{l_1=1}^{\infty} \sum_{l_2=1}^{\infty} (a_{l_1 l_2} \cos n_{l_1} x \cos n_{l_2} y)$$

is the Fourier series of the function $f(x, y)$, which belongs to all classes $L_{\vec{p}}$, $\vec{p} = \{p_1, p_2\}$, $1 \leq p_i < \infty$ ($i = 1, 2$), and

$$A_1(p_1, p_2, \lambda_1, \lambda_2) M \leq \|f(x, y)\|_{\vec{p}} \leq A_2(p_1, p_2, \lambda_1, \lambda_2) M,$$

where the constants A_1 and A_2 depend only on p_1, p_2, λ_1 and λ_2 .

Lemma 2 ([6]) *If $f \in L_{\vec{p}}$ and*

$$f(x, y) \sim \sum_{\nu=1}^{\infty} \sum_{\mu=1}^{\infty} c_{\nu\mu} \cos \nu x \cos \mu y,$$

where

$$c_{\nu\mu} = \begin{cases} a_{l_1 l_2}, & \text{if } \nu = 2^{l_1} \text{ and } \mu = 2^{l_2}, \\ 0, & \text{elsewhere,} \end{cases}$$

then the inequalities

$$\|f - s_{\infty m}(f)\|_{\vec{p}} \leq A_3 E_{\infty[\frac{m}{2}]}(f)_{\vec{p}},$$

$$\|\varphi_{nm}\|_{\vec{p}} \leq A_4 Y_{[\frac{n}{2}][\frac{m}{2}]}(f)_{\vec{p}}$$

hold, where the constants A_3 and A_4 are independent of $f(x, y)$, n and m .

We write

$$\begin{aligned} B(c_{\nu\mu}, k_1, k_2, n, m) &= \frac{1}{n^{k_1} m^{k_2}} \left(\sum_{\nu=1}^n \sum_{\mu=1}^m c_{\nu\mu}^2 \nu^{2k_1} \mu^{2k_2} \right)^{\frac{1}{2}} + \frac{1}{n^{k_1}} \left(\sum_{\nu=1}^n \sum_{\mu=1}^{\infty} c_{\nu\mu}^2 \nu^{2k_1} \right)^{\frac{1}{2}} \\ &\quad + \frac{1}{m^{k_2}} \left(\sum_{\nu=1}^{\infty} \sum_{\mu=1}^m c_{\nu\mu}^2 \mu^{2k_2} \right)^{\frac{1}{2}} + \left(\sum_{\nu=n+1}^{\infty} \sum_{\mu=m+1}^{\infty} c_{\nu\mu}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

3. Proof of Theorem 1

Using Lemma 1 for the function $\Delta_{h_1 h_2}^{k_1 k_2} f$, one obtains

$$\|\Delta_{h_1 h_2}^{k_1 k_2} f\|_{\vec{p}} \leq A_7 \left[\sum_{\nu=1}^{\infty} \sum_{\mu=1}^{\infty} c_{\nu\mu}^2 \left(2 \sin \frac{\nu h_1}{2} \right)^{2k_1} \left(2 \sin \frac{\mu h_2}{2} \right)^{2k_2} \right]^{\frac{1}{2}}.$$

Now, since $\sin x \leq x$ for $x > 0$, we have

$$\sin \frac{\nu |h_1|}{2} \leq \frac{\nu |h_1|}{2} \quad \text{and} \quad \sin \frac{\mu |h_2|}{2} \leq \frac{\mu |h_2|}{2}.$$

The inequality $|\sin x| \leq 1$ yields

$$\begin{aligned} \|\Delta_{h_1 h_2}^{k_1 k_2} f\|_{\vec{p}} &\leq A_8 \left\{ \left[\sum_{\nu=1}^n \sum_{\mu=1}^m c_{\nu\mu}^2 \left(2 \sin \frac{\nu |h_1|}{2} \right)^{2k_1} \left(2 \sin \frac{\mu |h_2|}{2} \right)^{2k_2} \right]^{\frac{1}{2}} \right. \\ &\quad + \left[\sum_{\nu=n+1}^{\infty} \sum_{\mu=1}^m c_{\nu\mu}^2 \left(2 \sin \frac{\nu |h_1|}{2} \right)^{2k_1} \left(2 \sin \frac{\mu |h_2|}{2} \right)^{2k_2} \right]^{\frac{1}{2}} \\ &\quad + \left[\sum_{\nu=1}^n \sum_{\mu=m+1}^{\infty} c_{\nu\mu}^2 \left(2 \sin \frac{\nu |h_1|}{2} \right)^{2k_1} \left(2 \sin \frac{\mu |h_2|}{2} \right)^{2k_2} \right]^{\frac{1}{2}} \\ &\quad \left. + \left[\sum_{\nu=n+1}^{\infty} \sum_{\mu=m+1}^{\infty} c_{\nu\mu}^2 \left(2 \sin \frac{\nu |h_1|}{2} \right)^{2k_1} \left(2 \sin \frac{\mu |h_2|}{2} \right)^{2k_2} \right]^{\frac{1}{2}} \right\} \\ &\leq A_9 \left\{ |h_1|^{k_1} |h_2|^{k_2} \left(\sum_{\nu=1}^n \sum_{\mu=1}^m c_{\nu\mu}^2 \nu^{2k_1} \mu^{2k_2} \right)^{\frac{1}{2}} + |h_1|^{k_1} \left(\sum_{\nu=1}^n \sum_{\mu=m+1}^{\infty} c_{\nu\mu}^2 \nu^{2k_1} \right)^{\frac{1}{2}} \right. \\ &\quad \left. + |h_2|^{k_2} \left(\sum_{\nu=n+1}^{\infty} \sum_{\mu=1}^m c_{\nu\mu}^2 \mu^{2k_2} \right)^{\frac{1}{2}} + \left(\sum_{\nu=n+1}^{\infty} \sum_{\mu=m+1}^{\infty} c_{\nu\mu}^2 \right)^{\frac{1}{2}} \right\}. \end{aligned}$$

Since

$$\omega_{k_1 k_2}(f, \frac{1}{n}, \frac{1}{m})_{\vec{p}} = \sup_{|h_1| \leq 1/n, |h_2| \leq 1/m} \|\Delta_{h_1 h_2}^{k_1 k_2} f\|_{\vec{p}},$$

we obtain

$$\omega_{k_1 k_2}(f, \frac{1}{n}, \frac{1}{m})_{\vec{p}} \leq A_9 B(c_{\nu\mu}, k_1, k_2, n, m). \quad (2)$$

Now, we prove the opposite inequality. Considering Lemma 1, Lemma 2, the inequalities

$$\begin{aligned} Y_{nm}(f)_{\vec{p}} &\leq C \omega_{k_1 k_2}(f, \frac{1}{n}, \frac{1}{m})_{\vec{p}} \\ E_{\infty m}(f)_{\vec{p}} &\leq C \omega_{k_2}(f, \frac{1}{m})_{\vec{p}}, \end{aligned} \quad (3)$$

where the constant C does not depend on $f(x, y)$, n and m , and qualities of the modulus of smoothness, we have

$$\begin{aligned} \left(\sum_{\nu=n+1}^{\infty} \sum_{\mu=m+1}^{\infty} c_{\nu\mu}^2 \right)^{\frac{1}{2}} &\leq A_{10} \|\varphi_{nm}(x, y)\|_{\vec{p}} \\ &\leq A_{11} Y_{\left[\frac{n}{2}\right]\left[\frac{m}{2}\right]}(f)_{\vec{p}} \\ &\leq A_{12} \omega_{k_1 k_2}(f, \frac{1}{n}, \frac{1}{m})_{\vec{p}}. \end{aligned} \quad (4)$$

The inequality $(2/\pi)x \leq \sin x$, $0 \leq x \leq \pi/2$, yields

$$\frac{1}{n^{k_1} m^{k_2}} \left(\sum_{\nu=1}^n \sum_{\mu=1}^m c_{\nu\mu}^2 \nu^{2k_1} \mu^{2k_2} \right)^{\frac{1}{2}} \leq A_{13} \left[\sum_{\nu=1}^{\infty} \sum_{\mu=1}^{\infty} c_{\nu\mu}^2 \left(2 \sin \frac{\nu}{2n} \right)^{2k_1} \left(2 \sin \frac{\mu}{2m} \right)^{2k_2} \right]^{\frac{1}{2}}.$$

Then Lemma 1 and the definition of the modulus of smoothness imply

$$\frac{1}{n^{k_1} m^{k_2}} \left(\sum_{\nu=1}^n \sum_{\mu=1}^m c_{\nu\mu}^2 \nu^{2k_1} \mu^{2k_2} \right)^{\frac{1}{2}} \leq A_{14} \|\Delta_{\frac{1}{n} \frac{1}{m}}^{k_1 k_2} f\|_{\vec{p}} \leq A_{15} \omega_{k_1 k_2}(f, \frac{1}{n}, \frac{1}{m})_{\vec{p}}. \quad (5)$$

Considering the estimate $(2/\pi)x \leq \sin x$, $0 \leq x \leq \pi/2$, and Lemma 2 we have

$$\begin{aligned} I &= \frac{1}{n^{k_1}} \left(\sum_{\nu=1}^n \sum_{\mu=m+1}^{\infty} c_{\nu\mu}^2 \nu^{2k_1} \right)^{\frac{1}{2}} \\ &\leq A_{16} \left[\sum_{\nu=1}^{\infty} \sum_{\mu=m+1}^{\infty} c_{\nu\mu}^2 \left(\sin \frac{\nu}{2n} \right)^{2k_1} \right]^{\frac{1}{2}} \\ &\leq A_{17} \|\Delta_{\frac{1}{n}}^{k_1}[f - s_{\infty m}(f)]\|_{\vec{p}}, \end{aligned}$$

where

$$f - s_{\infty m}(f) \sim \sum_{\nu=1}^{\infty} \sum_{\mu=m+1}^{\infty} c_{\nu\mu} \cos \nu x \cos \mu y.$$

Using the notation

$$F(x, y) = \Delta_{\frac{1}{n}}^{k_1} f(x, y),$$

one obtains

$$\Delta_{\frac{1}{n}}^{k_1} s_{\infty m}(f) = s_{\infty m}(F).$$

Then we have

$$I \leq A_{18} \|F - s_{\infty m}(F)\|_{\vec{p}}.$$

Applying Lemma 2 and the inequalities (3), we find

$$I = A_{18} E_{\infty[\frac{m}{2}]}(F)_{\vec{p}} \leq \omega_{k_2}(F, \frac{1}{m})_{\vec{p}} \leq A_{19} \omega_{k_1 k_2}(f, \frac{1}{n}, \frac{1}{m})_{\vec{p}}. \quad (6)$$

Similarly, one obtains

$$\frac{1}{n^{k_2}} \left(\sum_{\nu=n+1}^{\infty} \sum_{\mu=1}^m c_{\nu\mu}^2 \mu^{2k_2} \right)^{\frac{1}{2}} \leq A_{20} \omega_{k_1 k_2}(f, \frac{1}{n}, \frac{1}{m})_{\vec{p}}. \quad (7)$$

The inequalities (4), (5), (6) and (7) yield

$$A_{21} B(c_{\nu\mu}, k_1, k_2, n, m) \leq \omega_{k_1 k_2}(f, \frac{1}{n}, \frac{1}{m})_{\vec{p}}. \quad (8)$$

The inequalities (2) and (8) complete the proof. ■

4. Proof of Theorem 2

We use the following notation:

$$\begin{aligned} I_1 &= \frac{1}{2^{nk_1} 2^{mk_2}} \left(\sum_{\nu=1}^{2^n} \sum_{\mu=1}^{2^m} c_{\nu\mu}^2 \nu^{2k_1} \mu^{2k_2} \right)^{\frac{1}{2}}, \\ I_2 &= \frac{1}{2^{nk_1}} \left(\sum_{\nu=1}^{2^n} \sum_{\mu=2^{m+1}}^{\infty} c_{\nu\mu}^2 \nu^{2k_1} \right)^{\frac{1}{2}}, \\ I_3 &= \frac{1}{2^{mk_2}} \left(\sum_{\nu=2^{n+1}}^{\infty} \sum_{\mu=1}^{2^m} c_{\nu\mu}^2 \mu^{2k_2} \right)^{\frac{1}{2}}, \\ I_4 &= \left(\sum_{\nu=2^{n+1}}^{\infty} \sum_{\mu=2^{m+1}}^{\infty} c_{\nu\mu}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

Obviously,

$$I_1^* = \sum_{\nu=1}^{2^{n-1}} \sum_{\mu=1}^{2^{m-1}} c_{\nu\mu}^2 \nu^{2k_1} \mu^{2k_2} = \sum_{l_1=0}^{n-1} \sum_{\nu=2^{l_1}}^{2^{l_1+1}-1} \sum_{l_2=0}^{m-1} \sum_{\mu=2^{l_2}}^{2^{l_2+1}-1} c_{\nu\mu}^2 \nu^{2k_1} \mu^{2k_2}.$$

Let $c_{\nu\mu}$ satisfy the condition (1). Using the lacunary quality of $c_{\nu\mu}$, we obtain

$$\begin{aligned} I_1^* &\leq A_{22} \sum_{l_1=0}^{n-1} \sum_{l_2=0}^{m-1} a_{l_1 l_2}^2 2^{2l_1 k_1} 2^{2l_2 k_2} \\ &\leq A_{23} \sum_{l_1=0}^{n-1} \sum_{l_2=0}^{m-1} 2^{-2l_1 r_1 + 2l_1 k_1} 2^{-2l_2 r_2 + 2l_2 k_2} \\ &= A_{23} \sum_{l_1=0}^{n-1} 2^{2l_1(k_1 - r_1)} \sum_{l_2=0}^{m-1} 2^{2l_2(k_2 - r_2)}, \end{aligned}$$

i.e.

$$I_1^* \leq A_{24} 2^{2n(k_1 - r_1)} 2^{2m(k_2 - r_2)}.$$

Therefore,

$$I_1 \leq \frac{A_{25}}{2^{nr_1} 2^{mr_2}}.$$

Note that for every δ_1 and δ_2 such that $0 < \delta_i < 1$, $i = 1, 2$, there exist non-negative integers n and m , so that

$$\frac{1}{2^{n+1}} < \delta_1 \leq \frac{1}{2^n} \quad \text{and} \quad \frac{1}{2^{m+1}} < \delta_2 \leq \frac{1}{2^m}.$$

For n and m mentioned above, we have

$$I_1 \leq A_{26} \delta_1^{r_1} \delta_2^{r_2}. \quad (9)$$

In a similar way,

$$\begin{aligned} I_2^2 &= \frac{1}{2^{2nk_1}} \sum_{\nu=1}^{2^n-1} \sum_{\mu=2^m+1}^{\infty} c_{\nu\mu}^2 \nu^{2k_1} \\ &= \frac{1}{2^{nk_1}} \sum_{l_1=0}^{n-1} \sum_{\nu=2^{l_1}}^{2^{l_1+1}-1} \sum_{\mu=2^m+1}^{\infty} c_{\nu\mu}^2 \nu^{2k_1} \\ &\leq \frac{A_{27}}{2^{nk_1}} \sum_{l_1=0}^{n-1} \sum_{l_2=m+1}^{\infty} a_{l_1 l_2}^2 2^{l_1 k_1} \\ &\leq \frac{A_{28}}{2^{nk_1}} \sum_{l_1=0}^{n-1} \sum_{l_2=m+1}^{\infty} 2^{-2l_1 r_1 + 2l_1 k_1} 2^{-2l_2 r_2} \\ &= \frac{A_{28}}{2^{nk_1}} \sum_{l_1=0}^{n-1} 2^{2l_1(k_1 - r_1)} \sum_{l_2=m+1}^{\infty} 2^{-2l_2 r_2}, \end{aligned}$$

i.e.

$$I_2 \leq \frac{A_{29}}{2^{nr_1} 2^{mr_2}}$$

or

$$I_2 \leq A_{30} \delta_1^{r_1} \delta_2^{r_2}. \quad (10)$$

One easily proves

$$I_3 \leq A_{31} \delta_1^{r_1} \delta_2^{r_2}, \quad (11)$$

$$I_4 \leq A_{32} \delta_1^{r_1} \delta_2^{r_2}. \quad (12)$$

The proof of sufficiency follows from the inequalities (9), (10), (11), (12), and also from Theorem 1.

Let us prove that the condition (1) is necessary. Since

$$\begin{aligned}
 I_1^2 &= \frac{1}{2^{2nk_1} 2^{2mk_2}} \sum_{\nu=1}^{2^n} \sum_{\mu=1}^{2^m} c_{\nu\mu}^2 \nu^{2k_1} \mu^{2k_2} \\
 &= \frac{1}{2^{2nk_1} 2^{2mk_2}} \sum_{l_1=0}^n \sum_{l_2=0}^m a_{l_1 l_2}^2 2^{2l_1 k_1 + 2l_2 k_2} \\
 &\geq \frac{1}{2^{2nk_1} 2^{2mk_2}} a_{nm}^2 2^{2nk_1 + 2mk_2} = a_{nm}^2,
 \end{aligned}$$

we have

$$a_{nm} \leq I_1. \quad (13)$$

Since

$$I_1 \leq B(c_{\nu\mu}, k_1, k_2, 2^n, 2^m),$$

and by Theorem 1,

$$B(c_{\nu\mu}, k_1, k_2, 2^n, 2^m) \leq A_{33} \omega_{k_1 k_2} \left(f, \frac{1}{2^n}, \frac{1}{2^m}\right) \vec{p},$$

we obtain

$$I_1 \leq A_{34} \omega_{k_1 k_2} \left(f, \frac{1}{2^n}, \frac{1}{2^m}\right) \vec{p}. \quad (14)$$

If $f(x, y) \in S^\circ H_{\vec{p}}^{r_1 r_2}$, the inequalities (13) and (14) imply

$$a_{nm} \leq \frac{A_{35}}{2^{nr_1} 2^{mr_2}}. \quad (15)$$

The relation (15) leads to the proof of Theorem 2. ■

References

- [1] Konjushkov, A.A.: *Best approximation by trigonometric polynomials and Fourier coefficients* (in Russian). Mat. Sb. **44**, 1, 58–84 (1958)
- [2] Zhak, I.E.: *On conjugate double trigonometric series* (in Russian). Mat. Sb. **31**, 469–484 (1952)
- [3] Kagadij, L.: *Fourier coefficients and the modulus of smoothness of functions of two variables* (in Russian). Učēn. Zap. Tartus. un-ga **253**, 229–243
- [4] Berisha, M., and Potapov, M.K.: *Moduli of smoothness and the Fourier coefficients of periodic functions of one variable* (in Russian). Publ. Inst. Math. **26**, 215–228 (1979)
- [5] Berisha, M.: *Fourier coefficients and moduli of smoothness of functions of two variables* (in Russian). Mat. Vesnik **38**, 251–262 (1986)
- [6] Berisha, M., and Kastrati, R.: *Moduli of smoothness and coefficients of lacunary double trigonometric series* (in Russian). Punime Matematike Nr. 1, 29–41, Prishtinë 1986
- [7] Potapov, M.: *Imbedding theorems in mixed metric* (in Russian). Trudy Mat. Inst. Steklov. **156**, 143–156 (1980)
- [8] Aljančić, S., und Tomič, M.: *Über den Stetigkeitsmodul von Fourier-Reihen mit monotonen Koeffizienten*. Math. Z. **88**, 274–284 (1965)

received: July 20, 1988

revised: September 15, 1988

Author

Prof. Dr. M. Berisha
University of Prishtina
Faculty of Natural Sciences
Department of Mathematics
“M. Tito” Street N.N.
38000 – Prishtina
Yugoslavia

EDUARD A. DANIELJAN; FRIEDRICH LIESE

The analysis of a $M_r/G_r/1/\infty$ model with time dependent priorities

1. The model

Priority queueing is one of the most important directions of investigations in the Theory of Queues. The complete presentation of the subject may be obtained from N. K. Jaiswal [1] or from B. V. Gnedenko et al. [2]. In both monographs the $M_r/G_r/1/\infty$ model is considered. The $M_r/G_r/1/\infty$ is a single server system with customers belonging to one of a set of r different priority classes, indexed by the subscript k ($k = \overline{1, r}$). Customers from priority class k (say k -customer) arrive in a Poisson stream at rate λ_k customers per time unit; each k -customer has a service time selected independently from the common distribution function $B_k(x)$, $B_k(0) < 1$. The condition $B_k(0) < 1$ means that "instantaneous" service for k -customers with probability one is impossible. At epoch $t = 0$ the model is assumed to be empty.

Among the queueing disciplines that impose an external priority structure on the arriving customers the head-of-the-line priority discipline is the most common and most natural. Therefore it was deeply investigated (see [1],[2]). At the same time the head-of-the-line priority discipline in $M_r/G_r/1/\infty$ model minimizes linear cost functions with regard to delay for various customers among all *conservative* disciplines without service interruptions (see [3], p.126). A queueing discipline is called *conservative* if no work (service requirement) is created or destroyed within the model; for example, destruction of work would occur if a customer is forced to leave the model before completing his service and the creation of work might correspond to a server standing idle whereas the queue is non-empty. The section of dynamik priority disciplines is relatively new for priority queues. An arriving customer is assigned a set of parameters that determine his relative position in the queue. This position may vary as a function of time owing to the appearance of customers of different priorities in the queue.

An interesting example of such a discipline first studied by L. Kleinrock [3] for $M_r/G_r/1/\infty$ model is known by the name of time-dependent priority discipline. Except Kleinrock's works about waiting times and it's generalization on some cases of the following priority function $q(t) = c + b(t - x)$ (see [4], [5]) only the publication [6] on this subject is known.

Time-dependent priority discipline provides a set of variable parameters b_k ($k = \overline{1, r}$) where

$$b_1 \geq b_2 \geq \dots \geq b_r.$$

Let us assume that some tagged k -customer ($k = \overline{1, r}$) arrives at time x and is assigned at time t a priority $q_k(t)$ defined by

$$q_k(t) = (t - x)b_k$$

where t ranges from x until the time at which this customer's service is completed. No interruptions are permitted. Whenever the service facility is ready for a new customer, the customer with the highest instantaneous priority $q(t)$ is then taken into service (that is, at time t a customer with priority $q'(t)$ is given preferential treatment over a customer with priority $q(t)$ where $q'(t) > q(t)$). Whenever a tie for the highest priority occurs, the tie is broken by a FIFO (first input – first output) rule. Obviously customers from the same class taken into service in order of a FIFO rule. The higher priority customer gains priority at a faster rate.

If the time dependent priorities $q_k(t)$ are applied then the FIFO discipline and the head-of-the-line priority disciplines occur as limit cases. One has to assume $b_1 = \dots = b_r$ and $b_k/b_{k+1} \rightarrow \infty$ for each $k = \overline{1, r-1}$, respectively.

Known results

Until now, only two results have been established concerning such a discipline with Poisson arrivals. Both of them are related to waiting times.

Let $w_k(t)$ be the virtual waiting time for k -customers at time t and

$$W_k(s, t) = E \exp(-sw_k(t))$$

for each $k = \overline{1, r}$, $s \geq 0$, $t \geq 0$, where E is the symbol for the expectation.

We use the symbol b_{ki} ($k = \overline{1, r}$, $i \geq 1$) for the i -th moment of k -customer's service time. Then $r_k = a_k b_{k1}$ is the traffic intensity for k -customers and $R = r_1 + \dots + r_r$ is the traffic intensity for the model (say traffic intensity).

For the class of conservative disciplines in $M_r/G_r/1/\infty$ equilibrium distributions exist if $R < 1$.

We reserve the symbol $W_k(s) = \lim_{t \rightarrow \infty} W_k(s, t)$ and w_{ki} for the Laplace Transform and the i -th moment of k -customers stationary waiting times, respectively. Put

$$w_0 = r_1 \cdot \frac{b_{12}}{2b_{11}} + \dots + r_r \cdot \frac{b_{r2}}{2b_{r1}}.$$

Kleinrock's theorem ([3]) *For time-dependent priorities (without interruptions) in $M_r/M_r/1/\infty$ the quantities w_{k1} ($k = \overline{1, r}$) may be calculated from the set of r linear equations*

$$w_{k1} = \frac{(w_0/(1-R)) - \sum_{i=k+1}^r r_i(1 - (b_i/b_k))w_{i1}}{1 - \sum_{i=1}^{k-1} r_i(1 - (b_k/b_i))}. \quad (1)$$

For each $s \geq 0$ let $m_r(s)$ be the root with the largest absolute value among all roots of the functional equation

$$u = s - \sum_{i=1}^{r-1} a_i(1 - (b_r/b_i))(1 - C_i(u))$$

where

$$C_i(s) = \int_0^{\infty} e^{-sx} dB_i(x), \quad i = \overline{1, r}.$$

Let $W(s, t)$ be the Laplace Transform of the virtual waiting time distribution function at time t in primary given $M_r/G_r/1/\infty$ model according to FIFO rule.

Danieljan's theorem ([6]) *For time-dependent priorities (without interruption) in the $M_r/G_r/1/\infty$ model*

$$W_r(s, t) = W(m_r(s), t) \quad (2)$$

holds for each $s \geq 0, t \geq 0$.

In the present paper we develop the method of *Collective Marks* which have been used for non-dynamic priority disciplines in [2] to investigate models with time-dependent priority disciplines. Some other characteristics connected with virtual waiting time such as busy periods are considered.

2. The discussion of the problem

For k -customers at each time t the basic interest in the $M_r/G_r/1/\infty$ model is the behaviour of the virtual waiting time $w_k(t)$. Let us make an agreement:

All notations for time-dependent priorities we also use for head-of-the-line priorities with additional upper index o .

At first let us discuss the situation for head-of-the-line priorities.

Let $\bar{w}_k^0(t)$, ($k = \overline{1, r}$) be the conditional virtual waiting time for k -customers at time t . The classical condition for the *external* (static) priority disciplines is that "no customer arrive after time t ". Let $T_k^0(u)$ be the busy period for $\overline{1, k}$ -customer (1-customer, ..., k -customer) with delay u . That is the length of a time interval that begins from delay $u \geq 0$ and finishes at the first moment after delay when no $\overline{1, k}$ -customers are present in the model. Note that $\frac{T_k^0(u)}{1, k\text{-customers}} \geq u$ and in duration of $T_k^0(u)$ after delay the server is continuously busy serving only $\overline{1, k}$ -customers.

The consideration of $w_k^0(t)$ and $T_{k-1}^0(u)$ is basic for obtaining $w_k(t)$. The key to the mentioned approach is the following important relation

$$w_k^0(t) = T_{k-1}^0(\bar{w}_k^0(t)) \quad (3)$$

for any $k = \overline{1, r}, t \geq 0$, where the processes $T_{k-1}^0(\cdot)$ and $\bar{w}_k^0(\cdot)$ are independent and $T_0^0(u) = u$.

In order to obtain a relation for the $M_r/G_r/1/\infty$ model with time-dependent priority discipline which is similar to (3) we intend to find suitable definitions of busy period with delay and conditional virtual waiting time.

Definition of busy period

Now we introduce $T_k(u)$, ($k = \overline{1, r-1}$, $u \geq 0$) as a busy period for $\overline{1, k}$ -customers with delay u according to priority rate b_{k+1} . What does it mean is shown in Figure 1.

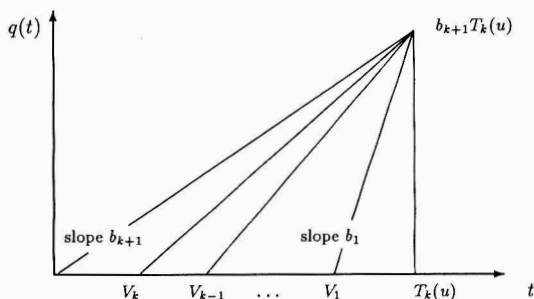


Figure 1. Diagram of priority $q(t)$ for obtaining $T_k(u)$.

Remember that service during the delay is prohibited. Now let us explain the meaning of $T_k(u)$ which begins from initial delay and therefore $T_k(u) \geq u$. $T_k(u)$ finishes at the first (after delay) moment when no $\overline{1, k}$ -customers with priorities higher than $b_{k+1}T_k(u)$ are present in the model. During the time-interval $(u, T_k(u))$ the server is continuously busy serving only $\overline{1, k}$ -customers.

It is obvious from the Figure 1 that for each $i = \overline{1, k}$, all of these i -customers arrive in time-interval $(0, V_i)$ and no one that arrives after moment V_i will be served in the time-interval $(u, T_k(u))$. Clearly, $T_k(u) = u$ if for each $i = \overline{1, k}$ no i -customers arrive in $(0, x_i)$ where

$$b_{k+1}u = b_i \cdot (u - x_i) \quad \text{and so} \quad x_i = u \cdot (1 - (b_{k+1}/b_i)).$$

The quantities V_i for each $i = \overline{1, k}$ are calculated from the similar equation

$$b_{k+1} \cdot T_k(u) = b_i \cdot (T_k(u) - V_i) \quad \text{and so} \quad V_i = T_k(u) \cdot (1 - (b_{k+1}/b_i)).$$

Our definition of $T_k(u)$ for the head-of-the-line priority's case coincides with the definition of $T_k^0(u)$.

Definition of waiting time

We reserve the designation $\bar{w}_k(t)$ for the conditional virtual waiting time of k -customers at each time t under the condition that no customers arrive after moment t , and introduce two new definitions for conditional virtual waiting time:

1. $u_k(t)$ is the sum of service times for all those $\overline{1, k}$ -customers that the server finds in the queue at time t , plus the residual service time for the customer being on server at time t ;
2. $v_k(t)$ is $w_k(t)$ without service times for all those $\overline{1, k}$ -customers that arrive after moment t and complete their service before the moment $t + w_k(t)$.

The functions $u_k(t)$ and $v_k(t)$ coincide with $\bar{w}_k^0(t)$ for the head-of-the-line priority discipline. For the time-dependent priority discipline the given definitions of conditional virtual waiting time are different. The conclusions we may make at once are:

$$\begin{aligned} u_k(t) &\leq v_k(t) \leq \bar{w}_k(t) \quad \text{for each } k = \overline{1, r-1} \text{ and } t \geq 0; \\ u_r(t) &= v_r(t) = \bar{w}_r(t) \quad \text{for each } t \geq 0. \end{aligned} \tag{4}$$

Indeed, because $v_k(t)$ and $\bar{w}_k(t)$ include the service times of all $\overline{1, k}$ -customers present in the queue at time t and the residual service time for the customer being on the server at time t , then $u_k(t)$ in (4) is the minimal among $u_k(t)$, $v_k(t)$ and $\bar{w}_k(t)$.

On the other hand, $v_k(t)$ may include the service times of only some $(k+1)$ -customers, ..., r -customers (say, $\overline{k+1, r}$ -customers) present in the queue at time t who receive service before the moment $t + w_k(t)$.

Generally speaking, in (4) there has to be *strong* inequalities. It is shown in Figure 2 where we simplify the situation as much as possible and an example of the $M_3/G_3/1/\infty$ model demonstrates the typical differences among $u_2(t)$, $v_2(t)$ and $\bar{w}_2(t)$. In Figure 2, the tagged 2-customer's arrival time coincides with point 0. Suppose there is a nonempty queue of 1, 2-customer and there is one 3-customer in the queue at time 0 that arrived to the model at time u_1 and at time u_3 still presents in the queue; $t_2 = u_2(0)$ (see definition 1.) and $t_2 - t_1$ be the last service time in $(0, t_2)$; u_2 be the first after time 0 1-customer's arrival time and $t_3 - t_2$ be it's service time.

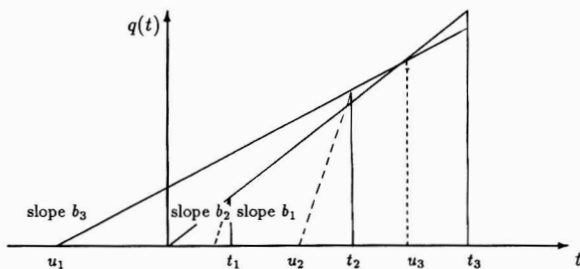


Figure 2. Diagram of $u_2(0)$, $v_2(0)$ and $\bar{w}_2(0)$ comparison in the $M_3/G_3/1/\infty$ model

For the described situation, if no arrivals are permitted after moment 0, then our 3-customer's service begins at time t_2 , and it's service time is a function of $\bar{w}_2(0)$. According to the definition 2. at time t_2 , we must take on the server the 1-customer that arrived at time u_2 . Therefore our 3-customer may be taken on server only after time t_3 . Note that at time t_3 the priority of virtual 2-customer is higher than the priority of our 3-customer. Thus, our 3-customer's service time doesn't belong to $v_2(0)$ and $v_2(0) < \bar{w}_2(0)$. Similarly we may form conditions for the inequality $u_2(0) < v_2(0)$.

$T_k(u)$ busy period's structure

Now we prove that the structure of the busy period for $\overline{1, k}$ -customers with delay u according to priority rate b_{k+1} is similar to $T_k^0(0)$'s structure.

For any $k = \overline{1, r-1}$, $i = \overline{1, k+1}$ and $s \geq 0$ set

$$\begin{aligned}\bar{a}_{ik+1} &= a_i \cdot (1 - (b_{k+1}/b_i)), \\ \bar{A}_k &= \bar{a}_{1k+1} + \bar{a}_{2k+1} + \dots + \bar{a}_{kk+1}, \\ T_k(u, s) &= E \exp(-s \cdot T_k(u)).\end{aligned}$$

If the argument u of $T_k(u)$ is a random variable with distribution function

$$B_k(x) = \sum_{i=1}^k (\bar{a}_{ik+1}/\bar{A}_k) \cdot B_i(x), \quad x \geq 0,$$

then instead of $T_k(u)$ we use a symbol T_k .

Put for each $k = \overline{1, r-1}$ and $s \geq 0$

$$\bar{T}_k(s) = E \exp(-s T_k), \quad m_{k+1}(s) = s + \bar{A}_k - \bar{A}_k \cdot \bar{T}_k(s).$$

Theorem 1 a) For each $k = \overline{1, r-1}$ and $s \geq 0$ the function $\overline{T}_k(s)$ is the root with the smallest absolute value among all roots $x = x(s)$ of the functional equation

$$\overline{A}_k \cdot x = \sum_{i=1}^k \overline{a}_{ik+1} \cdot C_i(s + \overline{A}_k - \overline{A}_k \cdot x). \quad (5)$$

b) The relation

$$T_k(u, s) = \exp(-m_{k+1}(s) \cdot u) \quad (6)$$

holds for each $u \geq 0$, $k = \overline{1, r-1}$ and $s \geq 0$.

Proof From Figure 1 we know that for each $i = \overline{1, k}$ only those i -customers that arrive in time-interval $(0, V_i)$ will be served in $(0, T_k(u))$. Assume that some tagged i -customer arrives at time $t \in (0, V_i)$. The tagged i -customer may be taken into service only after the moment

$$x = t \cdot b_i / (b_i - b_{k+1}).$$

Indeed, its priority at time $v \in (t, x)$ is less than $b_{k+1} \cdot v$ and at any time $v \in (x, T_k(u))$ more than $b_{k+1} \cdot v$. Therefore $T_k(u) > x > u$ and we can move the arrival time of the tagged i -customer from point t in point x . Let us make such kind of transformation with all i -customers ($i = \overline{1, k}$) that arrive in $(0, V_i)$, and design such a transform as T_{ik+1} . From the original stream S_i of i -customers located in $(0, V_i)$ (say, $S_i(0, V_i)$) we construct a new stream S'_i of i -customers located in $(0, T_k(u))$ (say, $S'_i(0, T_k(u))$). Generally speaking, for any w and v ($0 < v \leq w$)

$$T_{ik+1} S_i(v \cdot (1 - (b_{k+1}/b_i)), w \cdot (1 - (b_{k+1}/b_i))) = S'_i(v, w).$$

Thus, the linear transform T_{ik+1} replaces the Poisson stream $S_i(0, V_i)$ with rate a_i by a Poisson stream $S'_i(0, T_k(u))$ with rate \overline{a}_{ik+1} . Now it is easy to see that $T_k(u)$ and T_k are equal to $T_k^0(u)$ and T_k^0 respectively in the $M_k/G_k/1/\infty$ model where \overline{a}_{ik+1} ($i = \overline{1, k}$) is the arrival rate for i -customers and $B_i(x)$ their service time's distribution function. Therefore the statements a) and b) are consequences of known results for a head-of-the-line priority discipline in a $M_k/G_k/1/\infty$ model.

The main relation

The important property (3) is valid for the head-of-the-line priority discipline and may be extended to the case of time-dependent priority discipline if instead of $T_{k-1}^0(u)$ and $\overline{w}_k^0(t)$ we use their generalizations $T_{k-1}^0(u)$ and $v_k(t)$, respectively.

Let us formulate the result. Define for each $k = \overline{1, r}$, $s \geq 0$, $t \geq 0$

$$v_k(s, t) = E \exp(-s \cdot v_k(t)).$$

Theorem 2 *The relation*

$$w_k(t) = T_{k-1}(v_k(t)) \quad (7)$$

holds for each $k = \overline{1, r}$ and $t \geq 0$ where the processes $T_{k-1}(\cdot)$ and $v_k(t)$ are independent and $T_0(u) = u$.

In terms of the Laplace Transform relation (7) may be written as

$$W_k(s, t) = v_k(m_k(s), t) \quad (8)$$

for each $k = \overline{1, r}$, $s \geq 0$ and $t \geq 0$ where $m_1(s) = s$.

Proof For fixed $k = \overline{1, r}$ and each $i = \overline{1, r}$ we introduce a busy-interval h_{ik} , named ik -interval, as follows: For each i the random variable h_{ik} equals to $T_{k-1}(u)$ where u is a random variable with distribution function $B_i(x)$. It means that each i -customer generates a busy-interval h_{ik} (for fixed k) according to priority rate b_k .

Now, for each $k = \overline{1, r}$ the random process may be represented as a sum

$$v_k(t) = u_k(t) + \sum_{i=k+1}^r l_i(t) \quad (9)$$

of independent random processes $u_k(t)$, $l_{k+1}(t), \dots, l_r(t)$ where

a) $u_k(t)$ is the sum of independent service times for all those $\overline{1, k}$ -customers present in the queue at time t plus the residual service time for the customer that is in service at time t (residual service time and other service times are mutually independent);

b) for each $i = \overline{k+1, r}$ the process $l_i(t)$ is the sum of independent service times for all those i -customers present in the queue at time t and belong to $w_k(t)$.

Similarly, as the residual service time h_{ik} generates a busy-interval according to priority rate b_k each service time mentioned above individually and independently generates a busy-interval for $\overline{1, k-1}$ -customers according to priority rate b_k . It is clear that the process $w_k(t)$ is the sum of all those busy-intervals.

Now taking into account that $T_{k-1}(u)$ is a compound Poisson Process and using also the equality (9), we obtain

$$\begin{aligned} w_k(t) &= T_{k-1}(u_k(t)) + \sum_{i=k+1}^r T_{k-1}(l_i(t)) \\ &= T_{k-1} \left(u_k(t) + \sum_{i=k+1}^r l_i(t) \right) = T_{k-1}(v_k(t)). \end{aligned}$$

Applying (6) to (7) for each $k = \overline{1, r}$, $s \geq 0$, $t \geq 0$ and taking the Laplace Transform in (7) we obtain

$$\begin{aligned}
W_k(s, t) &= Ee^{-s \cdot w_k(t)} = Ee^{-s \cdot T_{k-1}(v_k(t))} = \int_0^\infty Ee^{-s \cdot T_{k-1}(u)} d_u P(v_k(t) < u) \\
&= \int_0^\infty e^{-m_k(s) \cdot u} d_u P(v_k(t) < u) = Ee^{-m_k(s) \cdot v_k(t)} = v_k(m_k(s), t).
\end{aligned}$$

Thus we established also the relation (8). ■

Note that the formulae (8) was obtained earlier by a different and long proof in [6]. Our proof is shorter and repairs an inaccuracy in the definition of the process $v_k(t)$ in [6].

Let us return to formulae (4). Now we have more information about the relation between processes $u_k(t)$ and $v_k(t)$. Indeed

Corollary 1 *The process $u_k(t)$ is an independent fraction of the process $v_k(t)$. Furthermore*

$$v_k(t) = u_k(t) + c_k(t)$$

where $u_k(t)$ and $c_k(t)$ are independent processes and $c_k(t)$ is the sum of service times of some $k+1, r$ -customers.

From Corollary 1 it is expedient to investigate separately the structure of $u_k(t)$.

3. General equations

Our purpose in this paragraph is to develop the approach based on probabilistic interpretations of the Laplace Transform for calculating virtual waiting time distribution function in terms of the Laplace Transform. We obtain a set of integral equations having complete information about virtual waiting time. We begin to formulate the main result.

Put for each $1 \leq n \leq k < j \leq r$, $s \geq 0$

$$\begin{aligned}
p_{nk}(s) &= s - \sum_{i=n}^k a_i \cdot (1 - C_i(s)), \\
p_{nk}^j(s) &= s - \sum_{i=n}^k \bar{a}_{ij} \cdot (1 - C_i(s)).
\end{aligned}$$

Let $P(t)dt$ be the probability that in time-interval $(t, t + dt)$ the model is empty. This quantity is invariant for the class of conservative disciplines in the $M_r/G_r/1/\infty$ model and may be calculated from the equation

$$\int_0^{\infty} e^{-st} P(t) dt = (m_{r+1}^0(s))^{-1}, \quad s \geq 0 \quad (10)$$

independently of the traffic intensity R .

Theorem 3 *For time-dependent priorities (without interruptions) in the $M_r/G_r/1/\infty$ model the set of following equations*

$$W_k(s, t) = u_k(m_k(s), t) - x_k(m_k(s), t) \quad (11)$$

holds for each $k = \overline{1, r}$, $s \geq 0$, $t \geq 0$ where

$$\begin{aligned} E e^{-s \cdot u_k(t)} &= u_k(s, t) = e^{p_{1k}(s) \cdot t} \left(1 - s \cdot \int_0^t e^{-p_{1k}(s) \cdot u} P(u) du \right) \\ &- \sum_{j=k+1}^r a_j \cdot (1 - C_j(s)) \int_0^t e^{-p_{1k}(s) \cdot v} \left[\int_0^{t-v} e^{-p_{1k}^j(s) \cdot u} d_u P(w_j(v) < u) \right] dv, \end{aligned} \quad (12)$$

$$x_k(s, t) = \sum_{j=k+1}^r \sum_{n=1}^k a_j \cdot (1 - C_j(s)) \cdot k_{nk}^j(s), \quad (13)$$

and for each $j = \overline{k+1, r}$ and $n = \overline{1, k}$:

$$x_{nk}^j(s) = e^{p_{nk}(s) \cdot t} \cdot \int_0^t e^{-p_{nk}(s) \cdot v} \cdot \left[\int_{\frac{b_{n-1}-b_j}{b_n-b_j}(t-v)}^{\frac{b_n-b_j}{b_n-b_j}(t-v)} e^{-p_{nk}^j(s) \cdot u} d_u P(w_j(v) < u) \right] dv. \quad (14)$$

Proof Below we use a probabilistic interpretation for the Laplace Transform. Let there be an independent “catastrophe” process taking place simultaneously which generates catastrophes at a rate $s > 0$ according to a Poisson process.

Then, for example, for each $k = \overline{1, r}$ the function $v_k(s, t)$ may be interpreted as probability of the event “no catastrophes occur during $v_k(t)$ ”.

Similarly, for each $i = \overline{1, r}$ the functions $C_i(s)$ and $1 - C_i(s)$ are the probabilities of the events “no catastrophes occur during i -customer’s service time”. Such i -customer we call “good” and “bad”, respectively.

Clearly the stream of “bad” i -customers is Poisson at rate $a_i \cdot (1 - C_i(s))$. Therefore, if for any $u \leq t$ $A_k((u, t)) =$ (all arriving in time-interval (u, t) $\overline{1, k}$ -customers are “good” .)

(that means no one "bad" $\overline{1, k}$ -customer appears in (u, t)); $L_k(t)$ = (no catastrophes occur in $(0, t)$ and during $v_k(t)$) and $N_k(t)$ = (no catastrophes occur in $(0, t)$ and during $u_k(t)$) then the probabilities $P(A_k((u, t)))$, $P(L_k(t))$ and $P(N_k(t))$ may be calculated from

$$P(A_k((u, t))) = \exp\left(-\sum_{i=1}^k a_i \cdot (1 - C_i(s)) \cdot (t - v)\right), \quad (15)$$

$$P(L_k(t)) = \exp(-st) \cdot v_k(s, t), \quad P(N_k(t)) = \exp(-st) \cdot u_k(s, t). \quad (16)$$

- a) All those $\overline{1, k}$ -customers that arrive in time-interval $(0, t)$ and are still present in the queue at time t form a fraction of $v_k(t)$ (also the fraction of $u_k(t)$).
- b) $v_k(t)$ begins at time t and finishes before the time $t + w_k(t)$ (the same for $u_k(t)$).
- c) $v_k(t)$ is represented as a sum of a random number of nonintersected time-intervals s_1, \dots, s_d where d is an integervalued random index and $s_q \in (t, t + w_k(t))$ for each $q = \overline{1, d}$.
- d) If some i -customer's service time for any $i = \overline{k+1, r}$ lies in $(t, t + w_k(t))$ then it essentially belongs to one time-interval s_q where $q = \overline{1, d}$.
- e) If for any $i = \overline{k+1, r}$ some i -customer's service time covers the time t then it's fraction from the time t forms the residual service time $w_0(t)$. Note that $w_0(t)$ belongs to $v_k(t)$ (also belongs to $u_k(t)$).

From property a) we have

$$A_k((0, t)) \supseteq N_k(t) \supseteq L_k(t).$$

Moreover,

$$A_k((0, t)) = L_k(t) + A_k((0, t)) \cap \overline{L_k(t)} = N_k(t) + A_k((0, t)) \cap \overline{N_k(t)} \quad (17)$$

where for any event C the complementary event denotes by symbol \overline{C} .

Taking into account the properties c)-d), we see from the event $\overline{L_k(t)} \cap A_k((0, t))$ that the first catastrophe necessarily appears:

1. Either in time-intervals $(0, t)$ when the model is empty;
2. or in time-interval $(0, t + w_0(t))$ when some j -customers ($j = \overline{k+1, r}$) is on the server;
3. or during some j -customer's ($j = \overline{k+1, r}$) service time that belongs to any time-interval s_q ($q = \overline{1, d}$).

Define the events $R_k(t)$, R_{jk}^1 , R_{jk}^2 as the intersection of $A_k((0, t))$ with the nonintersected events described in 1., 2., 3., respectively. From (17) we obtain

$$P(A_k((0, t))) = P(L_k(t)) + P(R_k(t)) + \sum_{j=k+1}^r (P(R_{jk}^1(t)) + P(R_{jk}^2(t))) \quad (18)$$

and similarly

$$P(A_k((0, t))) = P(N_k(t)) + P(R_k(t)) + \sum_{j=k+1}^r P(R_{jk}^1(t)). \quad (18')$$

We proceed to evaluate above introduced probabilities from the equality (18).

Calculation of $P(R_k(t))$

Let u ($u < t$) be the first catastrophe's occurrence time and suppose in the time-interval $(u, u + du)$ the model is empty (probability $se^{-su}P(u)du$).

It means that all arriving in time-interval $(0, u)$ $\overline{1, k}$ -customers (in particular, good $\overline{1, k}$ -customers) until the time u already completed their services. We need only all $\overline{1, k}$ -customers arriving in time-interval (u, t) to be good (probability $P(A_k((u, t)))$).

Summarizing over u in $(0, t)$ and applying (15) we get

$$P(R_k(t)) = s \cdot \int_0^t e^{-su} \cdot e^{-\sum_{i=1}^k a_i \cdot (1 - C_i(s)) \cdot (t-u)} \cdot P(u) du. \quad (19)$$

Calculation of $R_{jk}^1(t)$ for some $j = \overline{k+1, r}$

Let the first catastrophe occurs in time-interval $(0, t + w_0(t))$ when some j -customer (say, the tagged j -customer) is on the server. Then obviously the tagged j -customer's service begins at some time $u + v \in (0, t)$ where we denote by v the arrival time of this j -customer (the probability of such an event equals to $a_j \cdot du P(w_j(v) < u) dv$). It is clear that no catastrophe occurs in the time-interval $(0, v + u)$ (probability $\exp(-s \cdot (u + v))$).

The situation is given in Figure 3. The quantity x_m is calculated from $(u + v - x_m) \cdot b_m = u \cdot b_j$ and so $x_m = u + v - ub_j/b_m$, $t - x_m = t - v - u \cdot (1 - (b_j/b_m))$.

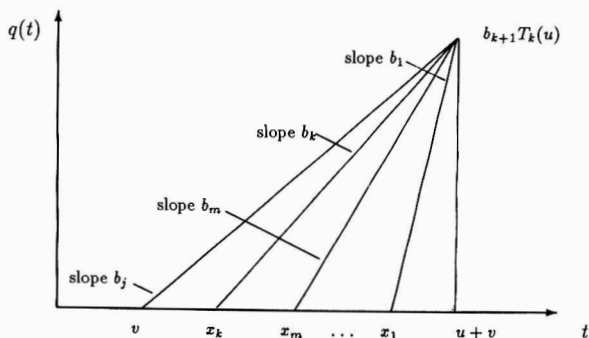


Figure 3. Calculation of $R_{jk}^1(t)$

All m-customers (in particular, good m-customers) arriving in $(0, x_m)$ already completed their services until the time $u + v$. But no one m-customer arriving in time-interval (x_m, t) will be taken on service before the tagged j-customer does.

Therefore we need all of them to be good. The probability is

$$\exp\left(-\sum_{m=1}^k a_m \cdot (1 - C_m(s)) \cdot (t - x_m)\right) = \exp\left(-\sum_{m=1}^k a_m \cdot (1 - C_m(s)) \cdot (t - v - u \cdot (1 - (b_j/b_m)))\right).$$

Note that during the tagged j-customer's time, catastrophe occurs (probability $1 - C_j(s)$) and it's service time either finishes before t or forms the residual service time $w_0(t)$.

Now, summarizing over u, v ($0 \leq u + v \leq t$), m ($m = \overline{1, k}$), and applying (15) we get

$$P(R_{jk}^1(t)) = a_j \cdot (1 - C_j(s)) \cdot \int_0^t \int_0^{t-v} e^{-s \cdot (u+v)} \exp\left(-\sum_{i=1}^k a_i \cdot (1 - C_i(s)) \cdot (t - v - u \cdot (1 - (b_j/b_i)))\right) d_u P(w_j(v) < u) dv. \quad (20)$$

Calculation of R_{jk}^2 for some $j = \overline{k+1, r}$

Let the first catastrophe occur during some j-customer's (say, the tagged j-customer) service time that belongs to any time-interval s_q ($q = \overline{1, d}$). Then obviously the arrival time of the tagged j-customer equals to some $v \in (0, t)$ and it's service beginning equals to some $u \in (t, t + w_j(t))$ (the probability of such an event also as above for the case of $R_{jk}^1(t)$ equals to $a_j \cdot d_u P(w_j(v) < u) dv$).

It is clear that no catastrophe occurs in time-interval $(0, v + u)$.

The situation is given in Figure 4.

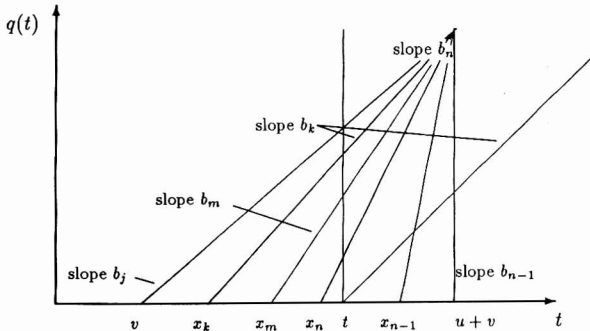


Figure 4. Calculation of $R_{jk}^2(t)$

The only restriction on u may be found from the equality for priorities for tagged j -customer, and for a virtual k -customer arriving at time t

$$x \cdot b_k = (t - v - x) \cdot b_j$$

where $x = \frac{b_j}{b_k - b_j}(t - v)$ is the moment when the mentioned priorities coincide.

Therefore the restriction on u may be written as

$$t - v \leq u \leq \frac{b_k}{b_k - b_j}(t - v).$$

The last inequality may be derived from several inequalities which are

$$\frac{b_{n-1}}{b_{n-1} - b_j}(t - v) \leq u \leq \frac{b_n}{b_n - b_j}(t - v) \quad (21)$$

where $n = \overline{1, k}$, $b_o = +\infty$, $b_o/(b_o - b_j) = 1$. In Figure 4 the exact case of the inequalities (21) is described. Indeed, for each $m = \overline{n-1, k}$ we have $ub_j = (u + v - x_m) \cdot b_m$ that yields

$$x_m = v + u \cdot (1 - (b_j/b_m))$$

so the condition $x_n \leq t \leq x_{n-1}$ is equivalent to (21).

For fixed n and every $m = \overline{n, k}$ according to (21) all arriving in time-intervals $(0, x_m)$ m -customers (in particular, good m -customers) and all arriving in time-interval $(0, t)$ $\overline{1, n-1}$ -customers (in particular, good $\overline{1, n-1}$ -customers) complete their services until the time $u + v$. On the other hand, no one m -customer ($m = \overline{n, k}$) arriving in time-interval (x_m, t) will be taken on service before the tagged j -customer does. Therefore, in the last case we need all of them to be good. The probability is

$$\exp\left(-\sum_{m=n}^k a_m \cdot (1 - C_m(s))(t - x_m)\right) = \exp\left(-\sum_{m=n}^k a_m \cdot (1 - C_m(s))(t - v - u \cdot (1 - (b_j/b_m)))\right).$$

Now it is easy to evaluate the probability of the event $R_{jk}^2(t)$:

$$\begin{aligned} P(R_{jk}^2(t)) &= a_j \cdot (1 - C_j(s)) \cdot \int_0^t \sum_{n=1}^k \int_{\frac{b_{n-1}}{b_{n-1}-b_j}(t-v)}^{\frac{b_n}{b_n-b_j}(t-v)} e^{-s \cdot (u+v)} \\ &\cdot \exp\left(-\sum_{i=n}^k a_i \cdot (1 - C_i(s))(t - v - u \cdot (1 - (b_j/b_i)))\right) \cdot d_u P(w_j(v) < u) dv. \end{aligned} \quad (22)$$

Calculation of $u_k(s, t)$

Substituting derived expressions for $P(N_k(t))$, $P(R_k(t))$, $P(R_{jk}^1(t))$ from (16), (19), (20) into equation (18') we obtain

$$\begin{aligned} \exp\left(-\sum_{i=1}^k a_i \cdot (1 - C_i(s))t\right) = \\ e^{-st} \cdot u_k(s, t) + s \cdot \int_0^t \exp\left(-\sum_{i=1}^k a_i \cdot (1 - C_i(s))(t - u)\right) P(u) du \\ + \sum_{j=k+1}^r a_j \cdot (1 - C_j(s)) \cdot \int_0^t e^{-sv} \cdot \exp\left(-\sum_{i=1}^k a_i \cdot (1 - C_i(s))(t - v)\right) \\ \cdot \int_0^{t-v} e^{-su} \cdot \exp\left(\sum_{i=1}^k a_i \cdot (1 - C_i(s))(1 - (b_j/b_i)) \cdot u\right) d_u P(w_j(v) < u) dv \end{aligned}$$

that is equivalent to (12).

Calculation of $v_k(s, t)$

Subtracting the equations (18), (18') we get

$$P(L_k(t)) = P(N_k(t)) - \sum_{j=k+1}^r P(R_{jk}^2(t)). \quad (23)$$

Substituting derived expressions for $P(N_k(t))$, $P(L_k(t))$, $P(R_{jk}^2(t))$ from (16), (22) into the equation (23) we obtain

$$\begin{aligned} e^{-st} \cdot v_k(s, t) = e^{-st} \cdot u_k(s, t) - \sum_{j=k+1}^r a_j \cdot (1 - C_j(s)) \cdot \sum_{n=1}^k \exp\left(-\sum_{i=n}^k a_i \cdot (1 - C_i(s))t\right) \\ \cdot \int_0^t e^{-P_{nk}(s) \cdot v} \int_{\frac{b_{n-1}-b_j}{b_n-b_j}(t-v)}^{\frac{b_n-b_j}{b_n-b_j}(t-v)} e^{-P_{nk}^j(s) \cdot u} d_u P(w_j(v) < u) dv. \end{aligned}$$

The last equation is equivalent to

$$v_k(s, t) = u_k(s, t) - x_k(s, t) \quad (24)$$

where $x_k(s, t)$ is calculated from (13), (14).

Finally, the relation (11) comes from the main relation in terms of the Laplace Transform (8) and equation (24). This completes the proof.

References

- [1] Jaiswal, N.K.: *Priority Queues*. New York 1968
- [2] Gnedenko, B.V. et al: *Prioritenije Sistemi Obslujivanja* (in Russian). Moscow 1973
- [3] Kleinrock, L.: *Queueing Systems: Computer Applications*. New York 1976
- [4] Netterman, A., and Adiri, A.: *A Dynamic Priority Queues with General Concave Priority Functions*. Oper. Res. **27**, 1088-1100 (1979)
- [5] Bagchi, U., and Sullivan, R.S.: *Dynamic Non-Preemptive Queues with General, Linearly Increasing Priority Function*. Oper. Res. **35**, 1278-1298 (1985)
- [6] Danieljan, E.A.: *Ob odnoi sisteme s dinamicheskimi prioritetami* (in Russian). Scientific Notes of Yerevan University **3**, 19-25 (1980)

received: January 20, 1990

Authors

Prof. Dr. E.A. Danieljan
Department of Mathematics
University of Yerevan
Mraviana 1
375049 Yerevan
USSR

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

VOLKER ISERNHAGEN

Die Strategie der orthogonalen Übergänge und ihre Anwendung auf zwei Klassen von Iterationsverfahren zur Lösung von linearen Gleichungen in Hilberträumen

Es seien X und Y zwei lineare Räume, und $L(X, Y)$ bezeichne die Menge aller linearen Operatoren von X in Y . Weiterhin bezeichne $R(A)$ den Wertebereich eines Operators $A \in L(X, Y)$. Zur Lösung bzw. verallgemeinerten Lösung der Operatorgleichung

$$Ax = b, \quad A \in L(X, Y), \quad b \in Y \quad (1)$$

eignen sich unter Verwendung einer Folge von Operatoren $D_n \in L(X, Y)$ Iterationsverfahren der Gestalt

$$x_{n+1} = x_n + D_n r_n, \quad r_n := b - Ax_n. \quad (2)$$

Die Fixpunkte dieses Iterationsverfahrens, d. h. die Elemente $x_o \in X$ mit $x_n = x_o$ für beliebiges n , erfüllen für alle n die Gleichungen $D_n Ax = D_n b$. Sie werden auch bezüglich $\{D_n : n \in N\}$ verallgemeinerte Lösungen der Gleichung (1) genannt (siehe [5], [9]). Ist die Beziehung

$$D_k A D_k = D_k \quad (3)$$

erfüllt, so ist x_{k+1} Lösung der Gleichung $D_k Ax = D_k b$ und kann somit als bezüglich $\{D_k\}$ verallgemeinerte Lösung von (1) aufgefaßt werden. Läßt sich in der Folge (D_n) ein Abschnitt von aufeinanderfolgenden Operatoren bestimmen, die paarweise die Beziehung $D_i A D_j = 0$ erfüllen, so läßt sich die Iteration (2) für diesen Abschnitt in eine einfache Form umschreiben, aus der einige interessante Eigenschaften zu ersehen sind.

Satz *Gelten für $n, m \in N$ mit $n \leq m$ und für alle i, j mit $n \leq i, j \leq m, i \neq j$, die Beziehungen*

$$D_i A D_j = 0, \quad (4)$$

so ist

$$x_{m+1} = x_n + \sum_{i=n}^m D_i (b - Ax_n). \quad (5)$$

Beweis Wir zeigen die Behauptung durch vollständige Induktion. Für $m = n$ ist (5) offensichtlich. Sei nun mit $j \leq m$ die Beziehung

$$x_j = x_n + \sum_{i=n}^{j-1} D_i(b - Ax_n)$$

erfüllt. Für x_{j+1} ergibt sich dann nach (2) und dem obigen Ausdruck für x_j

$$x_{j+1} = x_j + D_j(b - Ax_j) = x_n + \sum_{i=n}^j D_i(b - Ax_n) + D_j A \sum_{i=n}^{j-1} D_i(b - Ax_n).$$

Da nach Voraussetzung $D_j A D_i = 0$ für $i = n, n+1, \dots, j-1$ ist, gilt die Aussage auch für $j+1$ und damit der Satz.

Wir schließen dem Satz einige *Bemerkungen* an.

1. Betrachtet man den Ausdruck (5), so erkennt man, daß ein Vertauschen von Operatoren D_i mit $n \leq i \leq m$ keinerlei Auswirkungen auf die Iterierte x_{m+1} und die darauf folgenden Iterierten hat.
2. Die $(m-n)$ -malige Anwendung von (2) läßt sich unter den gegebenen Voraussetzungen auf einen Iterationsschritt

$$x_{m+1} = x_n + D_n^{(m)}(b - Ax_n)$$

des gleichen Typs mit $D_n^{(m)} = \sum_{i=n}^m D_i$ zurückführen, was von Vorteil sein kann.

3. Es sei M die Menge der Indizes von Operatoren D_i mit $n \leq i \leq m$, für die die Beziehung (3) besteht. Dann erfüllt jedes Element x_j mit $n < j \leq m+1$ sämtliche Gleichungen $D_k Ax = D_k b$ mit $k \in M$ und $k < j$, denn für $k \in M$, $k < j$ erhält man

$$D_k Ax_j = D_k A(x_n + \sum_{i=n}^{j-1} D_i(b - Ax_n)) = D_k Ax_n + D_k A D_k(b - Ax_n) = D_k b.$$

Wählt man die Folge (D_n) derart, daß die Beziehung (4) für möglichst viele benachbarte D_n gilt, so könnte man diese Wahl als Strategie der orthogonalen Übergänge bezeichnen.

Wir wollen nun die Anwendungen dieser Ergebnisse auf zwei spezielle Verfahrensklassen untersuchen.

A. PSH-ähnliche Iterationsverfahren

Es seien X und Y Hilberträume. Außerdem seien eine Folge (Z_n) von weiteren Hilberträumen und eine Folge (G_n) von Operatoren $G_n \in L(Y, Z_n)$ gegeben. Dabei setzen wir voraus, daß die Operatoren $A_n := G_n A \in L(X, Z_n)$ normal auflösbar sind ($\overline{R(A_n)} = R(A_n)$). Dann besitzen sie orthogonale verallgemeinerte Inverse $A_n^+ \in L(Z_n, X)$. Setzt man nun

$$D_n := \lambda_n A_n^+ G_n, \quad |\lambda_n - 1| < 1, \quad (6)$$

so erhält man aus (2) die von D. SCHOTT untersuchten *PSH-ähnlichen Iterationsverfahren* (siehe [9] – [11]).

Es seien $X = R^N$ und $Y = R^M$ zwei endlichdimensionale reelle Vektorräume. Man betrachte L Zeilenauswahlmatrizen E_0, E_1, \dots, E_{L-1} der Matrix $A \in L(R^N, R^M)$ derart, daß jede (nichtverschwindende) Zeile von A durch mindestens ein E_i ausgewählt wird. Entsteht die Folge $(G_n) = (E_n)$ durch zyklische Wiederholung der Matrizen E_0, E_1, \dots, E_{L-1} und setzt man alle λ_n gleich 1, so ergibt das Iterationsverfahren (2), (6) das vor allem von W. PETERS und G. MAESS untersuchte *zyklische PSH-Verfahren* (Projektion auf Schnitträume von Hyperebenen) mit

$$D_n = (E_n A)^+ E_n \quad (7)$$

(siehe u.a. [4, 5, 7, 8]). TANABE untersuchte in [12] die auf KACZMARZ [3] zurückgehende Variante des PSH-Verfahrens, bei der die E_n die transponierten Einheitsvektoren aus R^M sind (Projektion auf Hyperebenen, PH-Verfahren).

Für Operatoren D_n aus (6) besteht die Beziehung

$$D_n A D_n = \lambda_n D_n,$$

so daß die Bedingung (3) bei $\lambda_n = 1$ für alle Operatoren D_n erfüllt ist. Die Operatoren $D_n = \lambda_n A_n^+ G_n$ erfüllen die Beziehung (4), sobald $A_i A_j^+ = 0$ gesichert ist. Wegen $A_j^+ = A_j^+ A A_j^+ = (A_j^+ A_j)^* A_j^+ = A_j^* (A_j^+)^* A_j^+$ gilt $A_i A_j^+ = 0$, falls $A_i A_j^* = 0$ nachgewiesen ist. Wählt man die D_n nach (7), so ist (4) gleichbedeutend mit der Forderung, daß die von E_i ausgewählten Zeilen orthogonal zu den durch E_j ausgewählten Zeilen von A sind.

Da das Konvergenzverhalten der PSH-Verfahren bisher eingehender praktisch untersucht wurde, werden sich unsere Ausführungen nun auf diese Verfahrensklasse beziehen. Unsere Überlegungen lassen sich aber auf die allgemeinere Klasse der PSH-ähnlichen Verfahren übertragen.

Durch die Folge (E_n) , welche hier nicht notwendig zyklisch gewählt werden muß, wird eine Strategie vorgegeben, durch die die Konvergenzgeschwindigkeit wesentlich beeinflusst wird. Angestrebt wird natürlich eine Strategie, bei der der numerische Aufwand und die Konvergenzgeschwindigkeit in einem günstigen Verhältnis stehen. Bisher ist es noch nicht gelungen, für größere Matrizenklassen optimale Strategien anzugeben.

Werden durch die Matrizen E_n jeweils nur linear unabhängige Zeilen von A ausgewählt, so ist (7) gleichbedeutend mit

$$D_n = A^T E_n^T (E_n A A^T E_n^T)^{-1} E_n. \quad (8)$$

Um die Strategien einfacher beschreiben zu können, definieren wir Indexmengen $K^{(i)}$ ($i = 0, 1, \dots$), welche jeweils die Indizes der durch E_i ausgewählten Zeilen von A enthalten sollen. Die Mächtigkeit einer Indexmenge wird als Unterraumdimension bezeichnet. Zwei Strategien gelten als äquivalent (bezüglich eines Gleichungssystems), wenn sie bei Anwendung auf das System für gleichen Startvektor und gleiche Relaxationsparameter zwei Folgen von Iterierten erzeugen, die eine gemeinsame Teilfolge besitzen.

Mit Darstellung der Gleichung (1) in der Form

$$(a_i, x) = b_i, \quad i = 0, 1, \dots, M,$$

lassen sich die folgenden Aussagen über äquivalente Umformungen einer Strategie aus dem Satz und den sich anschließenden Bemerkungen für das Verfahren (2),(8) angeben. Zuerst setzen wir voraus, daß für ein n die Beziehungen $(a_i, a_j) = 0$ für alle i, j mit $i \in K^{(n)}$ und $j \in K^{(n+1)}$ erfüllt sind.

Für diesen Fall ergibt sich:

- (a) Durch Vertauschen der Indexmengen $K^{(n)}$ und $K^{(n+1)}$ entsteht eine äquivalente Strategie.
- (b) Das Ersetzen von $K^{(n)}$ und $K^{(n+1)}$ durch $K^{(n)} \cup K^{(n+1)}$ liefert ebenfalls eine äquivalente Strategie.

Insbesondere kann man somit eine Strategiemenge, die nur aus Indizes paarweise orthogonaler Zeilen der Matrix A besteht, durch eine beliebige Folge von deren Teilmengen ersetzen, sobald jeder Index der Strategiemenge in wenigstens einer Teilmenge vorkommt. Die Aussagen (a) und (b) gelten auch, wenn in D_n und D_{n+1} der gleiche Relaxationsparameter genutzt wird.

Es gelte nun für festes $i \in K^{(n)}$ die Beziehung $(a_i, a_j) = 0$ für alle $j \in K^{(n+k)}$ mit $k = 1, 2, \dots, m$ ($m \geq 1$).

- (c) Man kann jede Strategiemenge $K^{(n+k)}$ durch $K^{(n+k)} \cup \{i\}$ oder $K^{(n+k)} \setminus \{i\}$ ersetzen und erhält stets äquivalente Strategien.

Die Bemerkung (c) weist auf folgende Tatsache hin: Eine Iterierte x_{n+1} genügt einerseits denjenigen Gleichungen von $Ax = b$, die durch E_n ausgewählt werden, und andererseits erfüllt sie aber auch alle Gleichungen, die x_n als Lösung besitzen und deren Zeilenvektoren zu den von E_n ausgewählten orthogonal sind. Somit wäre das Auftreten des Index i in einer Strategiemenge $K^{(n+k)}$ ohne Bedeutung. Da der Rechenaufwand mit der Anzahl der Indizes einer Strategiemenge stark wächst, wird man bemüht sein, in den Indexmengen $K^{(n+k)}$ auf den Index i zu verzichten.

Für das KACZMARZ-Verfahren (siehe oben) gab TEWARSON [13] die Folgerungen (a) und (b) erstmals an.

Numerische Untersuchungen zeigen, daß das Kaczmarz-Verfahren schnell konvergiert, wenn die Hyperebenen "fast" orthogonal sind. Dagegen ist die Konvergenz schlecht, wenn die Hyperebenen sich unter spitzen Winkeln schneiden (siehe [4, 12, 2]).

Beim PSH-Verfahren ist die für eine vorgegebene Genauigkeit notwendige Schrittzahl i.allg. um so kleiner, je größer die Unterraumdimensionen der Strategie sind ([7, 8, 2]). Andererseits steigt mit der Unterraumdimension der Rechenaufwand pro Schritt erheblich ([2]). Mit den Aussagen (a)–(c) ist klar, daß durch das Auftreten von "orthogonalen Übergängen" in einer Strategie sich indirekt bei gleichbleibendem Aufwand die Unterraumdimension erhöht und somit eine bessere Konvergenzgeschwindigkeit zu erwarten ist. Außerdem "spart" man Rechenoperationen, wenn man eine Indexmenge äquivalent in mehrere Indexmengen kleinerer Dimension zerlegt. Dies wird deutlich, wenn man sich die für die Berechnung der Inversen $(E_n A A^T E_n^T)^{-1}$ benötigten Rechenoperationen veranschaulicht, welche i.allg. von der Ordnung der dritten Potenz der Mächtigkeit von $K^{(n)}$ sind.

Man wird also bemüht sein, daß die Indizes zueinander orthogonaler Zeilen in benachbarten Indexmengen enthalten sind.

WAINWRIGHT untersuchte in [14] die Variante des PSH-Verfahrens, bei der jeweils in den Schnittraum von zwei Hyperebenen projiziert wird (stets $|K^{(n)}| = 2$). Bei gegebenen x_n werden zwei Hyperebenen gesucht, bei denen der Rest r_{n+1} minimiert wird. Er gewann eine Abschätzung des Restes mit Hilfe des Winkels zwischen den beiden Hyperebenen und leitete daraus eine "quasi-optimale" Strategie ab. Dabei werden die $K^{(n)} = \{i, j\}$ so festgelegt, daß der Winkel zwischen den Hyperebenen, die durch die i -te und j -te Gleichung erzeugt werden, möglichst klein ist. Nach [14] ergaben numerische Tests, daß eine so gewählte Strategie die Konvergenz i.allg. verbessert. Durch diese Festlegung der Strategie wird auch der Winkel zwischen zwei durch benachbarte Indexmengen erzeugten Schnitträumen möglichst groß. Wir weisen darauf hin, daß die PSH-Verfahren besonders bei groß-dimensionierten Gleichungssystemen vorteilhaft angewandt werden können. Dabei ist die Matrix A oftmals schwach besetzt, so daß man in vielen Fällen schon aus der Besetzungsstruktur von A orthogonale Zeilen erkennen kann.

Bilden die $E_i A$ für $i = 0, 1, \dots, L-1$ ein System orthogonaler Blöcke und ist (E_n) zyklisch mit der Zyklenlänge L , so konvergiert das Verfahren nach einem Zyklus.

Wir geben nun ein aus [4] entnommenes Beispiel an, welches auch in [2] untersucht wurde.

Beispiel

$$\begin{pmatrix} 1 & 0 & 2 & -1 \\ 1 & -2 & 0 & 1 \\ 4 & 3 & -1 & 2 \\ -2 & -1 & 1 & 0 \end{pmatrix} x = \begin{pmatrix} 4 \\ 2 \\ -2 \\ 0 \end{pmatrix}$$

Die Matrix A ist "fast orthogonal", da mit Ausnahme von (a_3, a_4) alle Skalarprodukte verschwinden. KIESEWETTER, MAESS [4] gaben für die Strategie $K^{(n)} = \{i+1\}$, $i \equiv n \bmod 4$

eine langsame Konvergenz an. Mit der Folgerung (c) läßt sich nun zeigen, daß x_3 (bei beliebigem Startvektor x_0) den ersten drei Gleichungen genügt und x_4 Lösung der 1., 2. und 4. Gleichung ist. Daher gilt bei dieser Strategie $x_{4i} = x_{4i+1} = x_{4i+2}$ für $i = 1, 2, \dots$, d. h., es werden ineffektive Schritte ausgeführt. Eine verbesserte Strategie entsteht durch Weglassen dieser ineffektiven Iterationsschritte. Analog läßt sich zeigen, daß einige der in [2] untersuchten Strategien äquivalent zu der oben angegebenen sind. Stellvertretend geben wir eine davon an: $K^{(0)} = \{1\}$, $K^{(1)} = \{3\}$, $K^{(2)} = \{2\}$, $K^{(3)} = \{4\}$, $K^{(4)} = \{2\}$, $K^{(5)} = \{3\}$, $K^{(6)} = \{1\}$, $K^{(7)} = \{4\}$ und $K^{(n)} = K^{(i)}$ mit $n \equiv i \pmod{8}$.

Durch die Indext Mengen $\{1\}$, $\{2\}$ und $\{3, 4\}$ erhält man bei beliebiger Reihenfolge eine optimale Strategie, da sie die Lösung schon nach drei Iterationen liefert und den geringsten Rechenaufwand dafür erfordert.

Zum Schluß sei noch bemerkt, daß sich die bei PETERS [8] und ISERNHAGEN [2] beschriebenen Varianten der Vorbehandlung des Gleichungssystems unter Nutzung von Operatoren D_n nach (6) hier einordnen lassen. Bei PETERS werden die in einem Iterationsschritt benutzten Zahlen orthonormiert, um den Aufwand bei der Berechnung der Inversen zu senken. Das von ISERNHAGEN vorgeschlagene Orthonormierungsverfahren für das PH-Verfahren ist so aufgebaut, daß nacheinander ausgewählte Zeilen von A zueinander orthogonal sind. Damit wird die Unterraumdimension der Strategie indirekt erhöht.

Numerische Tests ergaben eine Konvergenzverbesserung.

B. SPA-ähnliche Iterationsverfahren

Es sei wiederum eine Folge (Z_n) von Hilberträumen und eine Folge (H_n) von Operatoren $H_n \in L(Z_n, X)$ gegeben. X und Y seien ebenfalls Hilberträume. Wir setzen voraus, daß die Operatoren $K_n := AH_n \in L(Z_n, Y)$ normal auflösbar sind ($\overline{R(K_n)} = R(K_n)$). Dann besitzen sie orthogonale verallgemeinerte Inverse $K_n^+ \in L(Y, Z_n)$. Setzt man

$$D_n := \lambda_n H_n K_n^+, \quad |\lambda_n - 1| < 1, \quad (9)$$

so erhält man aus (2) die von D. SCHOTT untersuchten *SPA-ähnlichen Iterationsverfahren* (siehe [9] – [11]).

Es seien $X = R^N$ und $Y = R^M$ zwei endlichdimensionale reelle Vektorräume. Man betrachte L Spaltenauswahlmatrizen F_0, F_1, \dots, F_{L-1} der Matrix $A \in L(R^N, R^M)$ derart, daß jede (nichtverschwindende) Spalte von A durch mindestens ein F_i ausgewählt wird. Entsteht die Folge $(H_n) = (F_n)$ durch zyklische Wiederholung der Matrizen F_0, F_1, \dots, F_{L-1} und setzt man alle λ_n gleich 1, so stellt das Iterationsverfahren (2), (9) das u. a. von G. MAESS und W. PETERS untersuchte *zyklische SPA-Verfahren* (Spaltenapproximation) mit

$$D_n = F_n (AF_n)^+ \quad (10)$$

dar (siehe [5], [8]). Werden durch die Matrizen F_i jeweils nur linear unabhängige Spalten von A ausgewählt, so ist (10) gleichbedeutend mit

$$D_n = F_n(F_n^T A^T A F_n)^{-1} F_n^T A^T. \quad (11)$$

Für die Operatoren D_n nach (9) ist die Bedingung

$$D_n A D_n = \lambda_n D_n$$

ebenfalls erfüllt. Die Gleichung (4) gilt, sobald $K_i^+ K_j = 0$ gesichert ist. Wegen $K_i^+ = K_i^+ K_i K_i^+ = K_i^+ (K_i K_i^+)^* = K_i^+ (K_i^+)^* K_i^*$ ist mit $K_i^* K_j = 0$ eine hinreichende Bedingung dafür gegeben.

Wählt man D_n nach (10), so ist die Gleichung (4) erfüllt, wenn die durch F_i ausgewählten Spalten orthogonal zu den durch F_j ausgewählten Spalten von A sind.

Die Anwendung des Satzes auf diese Verfahrensklasse liefert ganz analoge Aussagen wie für die PSH-Verfahren, wenn man statt der Zeilen nun die Spalten von A und damit verbundenen Räume betrachtet. Daher verzichten wir hier auf eine ausführlichere Darstellung.

Somit kann man auch für diese Verfahrensklasse in einer Reihe von Fällen eine gewisse Richtschnur für die Festlegung einer günstigen Strategie angeben.

Literatur

- [1] **Grimm, R.:** *Spaltenapproximation mit Relaxation zur iterativen Bestimmung von Kleinsten-Quadrat-Lösungen linearer algebraischer Gleichungssysteme.* Diplomarbeit, Pädagogische Hochschule Güstrow 1985
- [2] **Isernhagen, V.:** *Lösung linearer Gleichungssysteme durch Projektion auf Schnitträume von Hyperebenen.* Diplomarbeit, Pädagogische Hochschule Güstrow 1986
- [3] **Kaczmarz, S.:** *Angenäherte Auflösung von Systemen linearer Gleichungen.* Bull. Int. Acad. Polon. Sci. Lett. A., 355-357 (1937)
- [4] **Kiesewetter, H., und Maeß, G.:** *Elementare Methoden der numerischen Mathematik.* Berlin 1974
- [5] **Maeß, G.:** *Iterative Lösung linearer Gleichungssysteme.* Nova Acta Leopoldina, Neue Folge Nr. 238, Band 52, Halle 1979

- [6] **Maeß, G., und Peters, W.:** *Lösung inkonsistenter linearer Gleichungssysteme und Bestimmung einer Pseudoinversen für rechteckige Matrizen durch Spaltenapproximation.* Z. Angew. Math. Mech. **58**, 233–237 (1978)
- [7] **Peters, W.:** *Lösung linearer Gleichungssysteme durch Projektion auf Schnitträume von Hyperebenen und Berechnung einer verallgemeinerten Inversen.* Beiträge Numer. Math. **5**, 129–146 (1976)
- [8] **Peters, W.:** *Projektionsverfahren und verallgemeinerte Inverse.* Dissertation A, Universität Rostock 1977
- [9] **Schott, D.:** *Endlich erzeugte Projektionsverfahren zur Lösung linearer Gleichungen im Hilbertraum.* Rostock. Math. Kolloq. **16**, 103–128 (1981)
- [10] **Schott, D.:** *Die Methode der Projektionskerne und ihre Anwendung bei Struktur- und Konvergenzuntersuchungen von Iterationsverfahren zur Lösung linearer Operatorgleichungen in Banachräumen.* Dissertaton B, Universität Rostock 1982
- [11] **Schott, D.:** *Konvergenzsätze für Verallgemeinerungen von PSH- und SPA-Verfahren.* Math. Nachr. **118**, 89–103 (1984)
- [12] **Tanabe, K.:** *Projection method for solving a singular system of linear equations and its applications.* Numer. Math. **17**, 203–214 (1971)
- [13] **Tewarson, R.P.:** *Projection methods for solving sparse linear systems.* Comput. J. **12**, 78–81 (1969)
- [14] **Wainwright, R.L.:** *Acceleration techniques for a class of x-projection methods for solving systems of linear equations.* Comput. Math. Appl. **5**, 59–73 (1979)

eingegangen: 16. Januar 1989

Verfasser:

Dr. V. Isernhagen
Pädagogische Hochschule Güstrow
Sektion Mathematik/Physik
Goldberger Straße 12
O-2600 Güstrow
Germany

STEFAN WELLEK

Prinzipien der Planung und Auswertung von Äquivalenzstudien

Zusammenfassung

Äquivalenzstudien führen definitionsgemäß auf statistische Testprobleme, bei denen die Alternativhypothese gegeben ist durch ein beschränktes Intervall der Form $(\theta_0 - \delta_1, \theta_0 + \delta_2)$. Hierin bezeichnet θ_0 einen theoretisch spezifizierten Sollwert des jeweiligen Hauptparameters θ und δ_1 bzw. δ_2 die größte als "praktisch irrelevant" angesehene Abweichung zwischen θ und θ_0 . Für eine methodisch saubere Behandlung von Äquivalenz-Testproblemen werden eigene Entscheidungsverfahren benötigt, die sich nach Form und praktischer Durchführung von den für herkömmliche Probleme gewohnten Prozeduren erheblich unterscheiden. Im Hauptteil der vorliegenden Arbeit werden drei allgemeine Ansätze zur Konstruktion statistischer Tests auf Äquivalenz vorgestellt. Von besonderem Interesse ist dabei aus naheliegenden Gründen der klassische Ansatz, der für Äquivalenzprobleme Lösungen liefert, welche denselben Optimalitätskriterien genügen, die sich bei ein- und zweiseitigen Testproblemen erfolgreich anwenden lassen. Darüber hinaus wird aufgezeigt, daß es für die sorgfältige Planung einer Äquivalenzstudie wesentlich ist, eine von der Interpretation her sinnvolle Festlegung derjenigen parametrischen Funktion zu treffen, bezüglich welcher Äquivalenz nachgewiesen werden soll. Z.B. spricht in der Situation des Zweistichproben- t -Tests vieles dafür, von äquivalenten (varianzenhomogenen) Normalverteilungen zu fordern, daß die *standardisierte* Differenz ihrer Erwartungswerte genügend nahe bei 0 liegt.

1 Einführung

Äquivalenz- (ÄQ-) Studien unterscheiden sich von Studien herkömmlichen Typs durch die Form der statistischen Hypothesen, deren Verifizierung anhand der Daten angestrebt wird.

Bei der Formulierung der Hypothesen geht man generell aus von einem skalaren Parameter θ , welcher den interessierenden Effekt in den zugrunde liegenden Populationen sinnvoll quantifiziert. Ein in der Praxis häufiger Spezialfall ist derjenige des einfachen Parallelgruppen-Designs (ÄQ-Prüfung von zwei Behandlungen A und B in unverbundenen Stichproben), für das die Festlegung $\theta = \mu_1 - \mu_2$ naheliegend scheint, mit μ_1 bzw. μ_2 als einem Maß für die Lokation der A bzw. B entsprechenden Meßwertverteilung.

Die ÄQ-Hypothese, deren Plausibilität anhand der Studiendaten beurteilt werden soll, behauptet die Gleichheit von θ mit einem geeigneten Sollwert θ_0 "bis auf praktisch irrelevante Abweichungen". Die Menge solcher noch zu tolerierenden Abweichungen zwischen θ und θ_0 wird üblicherweise spezifiziert als ein genügend enges, θ_0 enthaltendes Intervall $(\theta_0 - \delta_1, \theta_0 + \delta_2)$ (\rightarrow Äquivalenzintervall), mit $\delta_1, \delta_2 > 0$. Im Falle des Parallelgruppen-Designs mit $\theta = \mu_1 - \mu_2$ wählt man fast immer $\theta_0 = 0$ und das ÄQ-Intervall symmetrisch um θ_0 , i.e. in der Form $(-\delta, \delta)$, mit $\delta > 0$.

Von einer ÄQ-Studie soll hier nur dann gesprochen werden, wenn die Anwendung eines statistischen Entscheidungsverfahrens vorgesehen ist, welches einen Test zum jeweiligen Niveau α ergibt für die *Nullhypothese*

$$H : \theta \leq \theta_0 - \delta_1 \vee \theta \geq \theta_0 + \delta_2$$

der Nichtäquivalenz, gegen die *ÄQ-Annahme*

$$K : \theta_0 - \delta_1 < \theta < \theta_0 + \delta_2$$

als *Alternativhypothese*. Ein solches Entscheidungsverfahren braucht nicht primär die Form eines Tests im gewohnten Sinne zu haben, sondern kann z.B. auch eine Bayes-Regel sein, von der sich zusätzlich zeigen läßt, daß die "objektive Wahrscheinlichkeit" einer falschen Entscheidung für Äquivalenz die Schranke α nicht überschreitet.

Eine schematische Übersicht über die *wichtigsten Schritte bei der sinnvollen Planung einer ÄQ-Studie* umfaßt die folgenden Punkte:

1. Festlegung der Grundstruktur der Studie (Anzahl der Behandlungen; unverbundene oder abhängige Beobachtungen; qualitative, diskrete oder stetige Variablen) sowie eines statistischen Modells für die Daten.
2. Wahl des Hauptparameters θ , bezüglich dessen Äquivalenz gezeigt werden soll.
3. Numerische Festlegung der ÄQ-Grenzen $\theta_0 - \delta_1, \theta_0 + \delta_2$.
4. Vorgabe von α ; Spezifikation eines Tests für das durch $(\theta_0 - \delta_1, \theta_0 + \delta_2)$ als ÄQ-Intervall bestimmte Testproblem (H, K) .
5. Berechnung der benötigten Fallzahlen bzw. der mit gegebenen Fallzahlen erreichbaren Trennschärfe (Power).

Die nachfolgenden Abschnitte dieser Arbeit geben eine Übersicht über Ansätze zur Lösung der in den Schritten 2-5 auftretenden Probleme. Die nach dem obigen Schema im 1. Schritt vorzunehmenden Spezifikationen haben bei der Planung von Studien beliebigen Typs am Anfang zu stehen, so daß sich hier im Hinblick auf Äquivalenz-Fragestellungen keine kommentierungsbedürftigen Besonderheiten ergeben.

2 Möglichkeiten und Probleme einer sinnvollen Festlegung des Hauptparameters θ

Der Übersichtlichkeit halber beschränken wir uns auf den Fall, daß ein Parallelgruppenversuch (unverbundene Stichproben) zur Überprüfung der Äquivalenz zweier Behandlungen A und B durchgeführt wird. Naheliegenderweise hat man hier θ so zu wählen, daß $|\theta|$ ein plausibles Maß für den Abstand zwischen den A und B zugeordneten theoretischen Verteilungen ergibt.

2.1 t -Test-Situation

Betrachtet man speziell die Situation des unverbundenen t -Tests, erkennt man leicht, daß dies zutrifft für die Wahl $\theta = (\mu_1 - \mu_2)/\sigma$, nicht aber für die unstandardisierte Mittelwertsdifferenz $\mu_1 - \mu_2$: Für jeden noch so großen festen Wert von $|\mu_1 - \mu_2|$ werden

die zugehörigen Gauß-Verteilungen $\mathcal{N}(\mu_1, \sigma^2)$, $\mathcal{N}(\mu_2, \sigma^2)$ praktisch ununterscheidbar, wenn die gemeinsame Varianz σ^2 extrem groß wird (vgl. [6], p.196) !

Gegen den Vorschlag, in der t -Test-Situation Äquivalenz durch die Forderung einer genügend kleinen standardisierten Mittelwertsdifferenz zu definieren, wird nicht selten der Einwand mangelnder Anschaulichkeit des Parameters $(\mu_1 - \mu_2)/\sigma$ erhoben (auch von Biometrikern !). Die folgende Überlegung zeigt, daß sich die Bedeutung der Maßzahl $(\mu_1 - \mu_2)/\sigma$ sogar besonders leicht veranschaulichen läßt.

Es sei X bzw. Y das Ergebnis der Anwendung von Behandlung A bzw. B bei einer zufällig ausgewählten Beobachtungseinheit, so daß $X \sim \mathcal{N}(\mu_1, \sigma^2)$, $Y \sim \mathcal{N}(\mu_2, \sigma^2)$ bei Unabhängigkeit des Paares (X, Y) . Für Behandlungen mit identischer Wirkung, d.h. mit $\mu_1 = \mu_2$, gilt offenbar $P[X < Y] = 1/2$. Andererseits ist die Unterschiedlichkeit der Behandlungen offenbar umso ausgeprägter, je weiter sich die Größe $P[X < Y]$ von $1/2$ entfernt. Eine für die Anschauung natürliche Definition von Äquivalenz ergibt sich hier somit aus der Forderung, daß

$$\frac{1}{2} - \delta_1^* < P[X < Y] < \frac{1}{2} + \delta_2^* .$$

Diese ist aber gleichbedeutend mit

$$\underbrace{-\Phi^{-1}\left(\frac{1}{2} + \delta_2^*\right)\sqrt{2}}_{=: \delta_1} < \frac{\mu_1 - \mu_2}{\sigma} < \underbrace{\Phi^{-1}\left(\frac{1}{2} + \delta_1^*\right)\sqrt{2}}_{=: \delta_2} ,$$

worin Φ^{-1} die Quantilfunktion der Standardnormalverteilung bezeichnet.

Obwohl es demnach gewichtige Argumente dafür gibt, in der t -Test-Situation Äquivalenz in terminis der standardisierten Mittelwertsdifferenz zu definieren, sieht zumindest in der medizinischen Biometrie die gegenwärtige Praxis so aus, daß geradezu routinemäßig auf ÄQ bezüglich $\mu_1 - \mu_2$ getestet wird. Dies trifft insbesondere zu auf dem durch gesundheitspolitische Maßnahmen eminent wichtig gewordenen Gebiet der Bio-Äquivalenzprüfungen. Vergleichende Bioverfügbarkeitsstudien werden standardmäßig (vgl. [1], [3], [20]) nach dem Zweiperioden-Crossoverplan vorgenommen und führen somit bei Abwesenheit heterogener Carry-over-Effekte genau auf das Problem eines Äquivalenztests für zwei Normalverteilungen mit homogenen Varianzen, aus denen unverbundene Stichproben vorliegen.

2.2 Binomialverteilungen

Wesentlich schwerer als in der t -Test-Situation tut man sich mit einer überzeugenden Begründung für die Auszeichnung einer speziellen skalaren Funktion der freien Parameter als Hauptparameter beim praktisch kaum weniger wichtigen binomialen Zweistichproben-Problem. Hier kommt die absolute Differenz $|p_1 - p_2|$ der Grundparameter genauso in Betracht wie die Abweichung $|\rho - 1|$ der Odds-Ratio von ihrem identische Binomialverteilungen charakterisierenden Wert.

Die einfache Differenz der Binomialparameter hat unbestreitbar den Vorzug, daß man mit ihr eine relativ direkte Größenordnungsvorstellung verbindet. Andererseits zeigen theoretische Überlegungen eindeutig, daß die Odds-Ratio ρ der mathematisch natürliche Parameter ist: Für das Problem des Testens auf ÄQ bezüglich ρ läßt sich ein exakter optimaler Test herleiten [13, 16]. Beim Testen auf $(p_1 - p_2)$ -Äquivalenz muß man sich hingegen mit asymptotischen Verfahren begnügen.

2.3 Allgemeine Bemerkungen zum Problem der Hauptparameter-Wahl

- (i) Die sinnvolle Festlegung eines Hauptparameters macht bei ÄQ-Problemen grundsätzlich Schwierigkeiten, wann immer eine mehrparametrische Verteilungsfamilie vorliegt. Demgegenüber spielt für die zugehörigen "konventionellen" (ein- oder zweiseitigen) Testprobleme die Nichteindeutigkeit des Hauptparameters keine Rolle. Dies hat den Grund, daß herkömmliche Hypothesen unter den in Frage stehenden Reparametrisierungen entweder invariant gelassen oder festen, nicht von unbekannten Parametern abhängenden Transformationen unterzogen werden. Beispielsweise hat man im binomialen Zweistichproben-Fall bei $\rho = p_1(1 - p_2)/p_2(1 - p_1)$ trivialerweise die Identität

$$\{(p_1, p_2) \in (0, 1)^2 \mid p_1 - p_2 = 0\} = \{(\rho, p_2) \in \mathbb{R}_+ \times (0, 1) \mid \rho = 1\},$$

so daß die herkömmliche zweiseitige Nullhypothese sich ebensogut mittels des Differenzparameters $p_1 - p_2$ wie der Odds-Ratio ρ formulieren läßt. Andererseits besitzt die auf $p_1 - p_2$ bezügliche Äquivalenzhypothese $-\delta_1 < p_1 - p_2 < \delta_2$ offenbar keine logisch gleichwertige Darstellung mittels ρ , es sei denn, einer der beiden Grundparameter p_1, p_2 ist bekannt.

- (ii) Unter Umständen wird durch eine gründliche Diskussion der Frage nach dem

von der Sache her primär interessierenden Verteilungsparameter sogar nahegelegt, ein zunächst unter Betracht genommenes ÄQ-Problem durch ein herkömmliches einseitiges Testproblem zu ersetzen. Ein interessantes Beispiel hierfür ergibt sich beim Bioäquivalenznachweis in Standard-Crossoverstudien ohne nachweisbaren Periodeneffekt. (Signifikante Periodeneffekte werden in vergleichenden Bioverfügbarkeitsstudien nicht nur faktisch sehr selten beobachtet, sondern lassen sich kaum biologisch plausibel erklären, vorausgesetzt, es liegen über beide Versuchsperioden hinweg konstante Laborbedingungen vor.)

Üblicherweise wird Bioäquivalenz der Testformulierung (T) eines Wirkstoffs mit einer Referenzformulierung (R) definiert durch das $\pm 20\%$ -Kriterium für die relative erwartete Bioverfügbarkeit von (T). D.h., für bioäquivalente Formulierungen wird gefordert, daß

$$.80 < \mu_T / \mu_R < 1.20 ,$$

mit $\mu_T = E(X)$, $\mu_R = E(Y)$ und X bzw. Y als der bei einem zufällig ausgewählten Probanden unter (T) bzw. (R) gemessenen "Bioavailability".

Vom Standpunkt eines klinischen Pharmakologen scheint es aber für wirklich gleichwertige Formulierungen kaum ausreichend, daß das 80-120%-Kriterium für die *mittleren* Bioverfügbarkeiten erfüllt ist. Vielmehr spricht vieles dafür (für eine eingehendere Diskussion siehe [14]), stattdessen zu fordern, daß der *individuelle* Bioverfügbarkeitsquotient X/Y bei einem genügend großen Anteil der Grundgesamtheit von Probanden um nicht mehr als 20% vom "Idealwert" 1 abweicht. Akzeptiert man diese Überlegung, hat man anstelle von μ_T / μ_R als Hauptparameter die Größe $P[.8 < X/Y < 1.2]$ zu betrachten. Danach hat beim Bioäquivalenznachweis in der Tat das Ziel die Bestätigung der einseitigen Alternativhypothese $P[.8 < X/Y < 1.2] > \pi_0$ zu sein, mit π_0 als genügend nahe bei 1 zu wählender Schranke.

3 Überlegungen zur numerischen Spezifikation der Äquivalenzgrenzen

An erster Stelle ist hier natürlich darauf zu achten, daß die Festlegung der ÄQ-Grenzen für den Hauptparameter θ unbedingt a priori vorgenommen werden muß und nicht etwa auf Grund eines "ersten Eindrucks" von den zu analysierenden Daten. So selbst-

verständlich dieser Grundsatz für den Statistiker scheint, so schwer fällt es oft immer noch, ihm in der Kooperation mit Klinikern und anderen Anwendern bei der Durchführung konkreter Studien effektive Geltung zu sichern.

Auch wenn man sich definitiv darauf verständigt hat, die ÄQ-Grenzen im vorhinein und damit bereits im Studienprotokoll zu fixieren, bleiben die folgenden Fragen offen, die nach Möglichkeit für jede Studie neu zu diskutieren bzw. zu entscheiden sind:

- (i) Soll das ÄQ-Intervall $(\theta_0 - \delta_1, \theta_0 + \delta_2)$ symmetrisch um den Sollwert θ_0 gewählt werden ?
- (ii) Welche Größenordnung von δ_1 bzw. δ_2 ist mit der anschaulichen Vorstellung "irrelevanter Abweichungen" zwischen θ_0 und wahren Wert des Hauptparameters noch vereinbar ?

Zu (i):

Eine symmetrische Festlegung der beiden ÄQ-Grenzen $\theta_0 - \delta_1$ und $\theta_0 + \delta_2$ scheint in den meisten Situationen sachlich vernünftig. Außerdem ergeben sich, wenn man die Herleitung optimaler Lösungen für die zugehörigen Testprobleme anstrebt, bei symmetrischer Gestalt des ÄQ-Bereichs typischerweise beträchtliche mathematische Vereinfachungen.

Andererseits trifft man in den Anwendungen mitunter auf die Situation, daß die Verwendung eines asymmetrischen ÄQ-Bereichs einer fest etablierten Konvention entspricht. Als von der Anwendungshäufigkeit her besonders wichtiges Beispiel bietet sich erneut die standardmäßige Vorgehensweise beim Bioäquivalenznachweis an. Die üblichen Verteilungsannahmen für die Analyse des Zweiperioden-Crossoverplanes werden hier zumeist nicht für die gemessenen Bioverfügbarkeiten, sondern für deren Logarithmen gemacht (\rightarrow lognormales "Model 1" im Sinne von [7]). Unter dieser Annahme ist die Differenz zwischen den Erwartungswerten der effektiv (nach Übergang zu den Differenzen zwischen den Versuchsperioden) zu vergleichenden Normalverteilungen gegeben als $\log(\mu_T/\mu_R)$. Folglich hat man, wenn man der $\pm 20\%$ -Konvention folgt und für bioäquivalente Wirkstoff-Formulierungen $|\mu_T/\mu_R - 1| < .20$ fordert, bezüglich $\theta = \log(\mu_T/\mu_R)$ gegen die unsymmetrische Alternativhypothese $\log(.80) < \theta < \log(1.20)$ zu testen.

Zu (ii):

Zur Frage nach der Größenordnung noch tolerierbarer Abweichungen des Hauptparameters vom Sollwert erhält man von substanzwissenschaftlicher Seite erfahrungsgemäß nur selten konkrete Vorgaben. Insoweit erscheint eine Orientierung an mehr oder minder schematischen Konventionen ohne weitergehende Begründung durch inhaltliche Überlegungen in der Regel unvermeidlich.

Konsensfähige Festlegungen von ÄQ-Grenzen scheinen u.a. dann gut erreichbar zu sein, wenn die ÄQ-Hypothese formulierbar ist in terminis der Differenz zweier Wahrscheinlichkeiten. So dürfte kaum bestritten werden, daß man z.B. im binomialen Zweistichproben-Fall die Forderung $|p_1 - p_2| < .20$ als sehr schwache Bedingung für die Äquivalenz der zugehörigen Verteilungen einzustufen hat. Ein wesentlich überzeugenderes Kriterium für die Äquivalenz zweier Binomialparameter oder eines einzelnen solchen Parameters mit einem theoretisch vorgegebenen Referenzwert (wie z.B. $1/2$ - vgl. Abschn. 2.1), von dem u.a. in [4] ausgegangen wird, läßt maximal 10% als Abweichung zu.

Den genannten Werten von ÄQ-Grenzen für Binomialparameter lassen sich vermöge sehr einfacher Umformungen solche für die standardisierte Differenz der Erwartungswerte zweier varianzenhomogener Normalverteilungen zuordnen. Bezeichnet nämlich wieder (X, Y) ein unabhängiges Paar von Zufallsvariablen mit $X \sim \mathcal{N}(\mu_1, \sigma^2)$, $Y \sim \mathcal{N}(\mu_2, \sigma^2)$, hat man wegen $P[X < Y] = \Phi((\mu_2 - \mu_1)/\sigma\sqrt{2})$:

$$|P[X < Y] - 1/2| < .20 \Leftrightarrow |\mu_1 - \mu_2|/\sigma < .7416 ,$$

$$|P[X < Y] - 1/2| < .10 \Leftrightarrow |\mu_1 - \mu_2|/\sigma < .3583 .$$

Folgt man den oben für Binomialparameter getroffenen Feststellungen, hätte man demnach in der Situation des unverbundenen t -Tests mit $(\mu_1 - \mu_2)/\sigma$ als Hauptparameter von äquivalenten varianzenhomogenen Normalverteilungen zumindest zu fordern, daß die Differenz ihrer Erwartungswerte nicht mehr als $\sim 75\%$ der gemeinsamen Standardabweichung beträgt.

4 Ansätze zur Konstruktion von Äquivalenztests

4.1 Das Prinzip der Konfidenzintervall-Inklusion

Der Grundidee nach geht dieser Ansatz auf [17] zurück. In der Nachfolge ist das Intervalleinschluß-Prinzip von einer Reihe von Autoren [8; 18, 19, 21; 12] in mindestens drei verschiedenen Versionen als allgemeine Lösung für ÄQ-Entscheidungsprobleme empfohlen und vielfach praktisch angewendet worden, hauptsächlich in Zusammenhang mit Bioäquivalenzprüfungen.

In der effizientesten Version arbeiten solche Intervalleinschluß-Tests wie folgt. Es sei $\underline{\theta}(\mathbf{X}; \alpha)$ bzw. $\bar{\theta}(\mathbf{X}; \alpha)$ eine untere bzw. obere Konfidenzschranke für θ zum *selben einseitigen Konfidenzniveau* $1 - \alpha$. Die Nullhypothese $H : \theta \leq \theta_0 - \delta_1 \vee \theta \geq \theta_0 + \delta_2$ der Nichtäquivalenz wird genau dann abgelehnt, wenn das Konfidenzintervall mit den Grenzen $\underline{\theta}(\mathbf{X}; \alpha)$, $\bar{\theta}(\mathbf{X}; \alpha)$ vollständig im "theoretischen" ÄQ-Intervall enthalten ist, wenn also gilt, daß $(\underline{\theta}(\mathbf{X}; \alpha), \bar{\theta}(\mathbf{X}; \alpha)) \subseteq (\theta_0 - \delta_1, \theta_0 + \delta_2)$.

Bemerkungen:

- (1) Das zweiseitige Konfidenzniveau von $(\underline{\theta}(\mathbf{X}; \alpha), \bar{\theta}(\mathbf{X}; \alpha))$ beträgt unter den obigen Voraussetzungen lediglich $1 - 2\alpha$. Trotzdem hält der zugehörige Test stets das Signifikanzniveau α ein, nicht nur 2α . Der Beweis dieser wichtigen Tatsache ist so einfach, daß er sich in wenigen Zeilen formulieren läßt: Für ein beliebiges θ im "linken Teil" der Nullhypothese, d.h. mit $\theta \leq \theta_0 - \delta_1$ hat man $P_\theta[\text{Fehler 1. Art}] = P_\theta[\theta_0 - \delta_1 < \underline{\theta}(\mathbf{X}; \alpha), \bar{\theta}(\mathbf{X}; \alpha) < \theta_0 + \delta_2] \leq P_\theta[\theta_0 - \delta_1 < \underline{\theta}(\mathbf{X}; \alpha)] \leq P_\theta[\theta < \underline{\theta}(\mathbf{X}; \alpha)]$. Da $\underline{\theta}(\mathbf{X}; \alpha)$ untere $(1 - \alpha)$ -Konfidenzschranke für θ sein soll, hat man aber definitionsgemäß $P_\theta[\theta < \underline{\theta}(\mathbf{X}; \alpha)] \leq \alpha \forall \theta$, so daß wir insgesamt schreiben können: $\theta \leq \theta_0 - \delta_1 \Rightarrow P_\theta[\text{Fehler 1. Art}] \leq \alpha$. Durch ein direkt analoges Argument folgert man, daß $\forall \theta \geq \theta_0 + \delta_2$ gilt: $P_\theta[\text{Fehler 1. Art}] \leq P_\theta[\bar{\theta}(\mathbf{X}; \alpha) \leq \theta] \leq \alpha$. Für alle anderen Werte von θ , d.h. im Falle von $\theta_0 - \delta_1 < \theta < \theta_0 + \delta_2$ gibt es nichts zu zeigen, da dann die Alternativhypothese gilt und ein Fehler 1. Art damit definitionsgemäß ausgeschlossen ist. Trotz der auch für ÄQ-Tests vorhandenen Durchsichtigkeit des Zusammenhangs zwischen Konfidenz- und Signifikanzniveau blieb die Möglichkeit, von einem $(1 - 2\alpha)$ -Konfidenzintervall auszugehen, ohne im zugehörigen Test auf Äquivalenz das Niveau α zu überschreiten,

in den ersten Arbeiten zur Anwendung des Intervallinklusions-Prinzips auf das Bioäquivalenznachweis-Problem unbeachtet (vgl. hierzu [21]).

- (2) In der t -Test-Situation mit $\theta = \mu_1 - \mu_2$, $\theta_0 = 0$ ist das Konfidenzintervall $(\underline{\theta}(\mathbf{X}; \alpha), \bar{\theta}(\mathbf{X}; \alpha))$ bekanntlich symmetrisch um $\bar{X} - \bar{Y}$ und damit fast sicher unsymmetrisch bezüglich 0. In [18] wird argumentiert, daß diese Unsymmetrie inkompatibel sei mit einer symmetrischen Festlegung des theoretischen ÄQ-Intervalls. Der Vorschlag, aus diesem Grunde mit einem um 0 symmetrischen Konfidenzintervall zu arbeiten, beruht jedoch bei näherem Hinsehen auf einem Mißverständnis: Die Anwendung des Intervallinklusions-Prinzips mit *symmetrischen Konfidenzintervallen* liefert *unsymmetrische Tests*, und vice versa.
- (3) Es läßt sich leicht zeigen (für den Spezialfall der Zweistichproben- t -Test-Situation siehe [9]), daß ein Intervalleinschlußtest im oben definierten Sinne äquivalent ist mit einer "Kombination" von zwei einseitigen Tests. Klarerweise lehnt man nämlich die Nullhypothese genau dann ab, wenn sowohl der Test von $\theta \leq \theta_0 - \delta_1$ gegen $\theta > \theta_0 - \delta_1$ auf der Basis der Statistik $\underline{\theta}(\mathbf{X}; \alpha)$ als auch derjenige von $\theta \geq \theta_0 + \delta_2$ gegen $\theta < \theta_0 + \delta_2$ mit $\bar{\theta}(\mathbf{X}; \alpha)$ als Prüfgröße signifikant wird.

4.2 Bayes-Prozeduren

Vom Bayesschen Standpunkt bieten ÄQ-Testprobleme weder in begrifflicher noch in technischer Hinsicht nennenswerte Besonderheiten. Alles, was bei festliegendem Hauptparameter θ und gegebenen ÄQ-Grenzen $\theta_0 - \delta_1$, $\theta_0 + \delta_2$ benötigt wird, ist eine gemeinsame A-priori-Verteilung $\pi(\cdot)$ für alle freien Parameter des zugrunde liegenden statistischen Modells. Der zugehörige Bayessche Test auf Äquivalenz von θ mit θ_0 lehnt ab, wenn für die A-posteriori-Wahrscheinlichkeit des ÄQ-Bereichs ein genügend großer Wert beobachtet wird, wobei es üblich ist, diese Schranke auf $1 - \alpha$ festzulegen. Mit anderen Worten ist dann der einer gegebenen A-priori-Verteilung $\pi(\cdot)$ entsprechende Bayes-Test auf ÄQ bestimmt durch die Entscheidungsregel:

$$(\star) \quad \text{Ablehnung von } H \Leftrightarrow \pi^{\Theta|\mathbf{X}=\mathbf{x}}(\theta_0 - \delta_1, \theta_0 + \delta_2) \geq 1 - \alpha \quad .$$

Wie immer ist der Bayessche Ansatz auch beim Testen auf ÄQ attraktiv durch seine konzeptuelle Einfachheit und Flexibilität bezüglich Änderungen in der Parametrisierung des Problems. Die A-posteriori-Verteilung wird ebenfalls für alle freien Modellparameter

gemeinsam bestimmt, so daß sich diejenige einer beliebigen parametrischen Funktion durch routinemäßige Anwendung der bekannten Transformationssätze für W-Dichten berechnen läßt. Beispielsweise macht es in der t -Test-Situation keinerlei Schwierigkeiten, für den Quotienten der Erwartungswerte die A-posteriori-Wahrscheinlichkeit eines als $\ddot{A}Q$ -Bereich vorgegebenen Intervalls zu berechnen. Auch zur Frage nach einer adäquaten Festlegung der A-priori-Verteilung existieren vernünftige Vorschläge, wobei man sich sehr oft für eine sog. nichtinformativ Festlegung entscheidet, d.h. eine (evtl. uneigentliche) Gleichverteilung einer geeigneten injektiven Funktion des Vektors aller freien Modellparameter (cf. [2], Sec.1.3).

Als Hauptschwierigkeit bei der Verwendung einer Entscheidungsprozedur von der Form (*) bleibt die Tatsache, daß ein solches Verfahren nicht notwendig die Einhaltung des Signifikanzniveaus α garantiert, was ja auch ein "frequentistisches" und damit für Bayesianer irrelevantes Konzept ist. Unter diesem Gesichtspunkt sind Bayessche Testprozeduren erst dann akzeptabel, wenn sich durch Simulation, exakte numerische Berechnung von Ablehnungswahrscheinlichkeiten oder sogar auf analytischem Wege zeigen läßt, daß das nominelle α zumindest in guter Näherung eingehalten wird.

Ein interessantes Beispiel eines analytischen Nachweises dafür, daß eine Bayes-Regel einen hinsichtlich des Niveaus strikt gültigen Test auf $\ddot{A}Q$ ergibt, wird in [7] betrachtet. Für den Bioäquivalenznachweis im Zweiperioden-Crossover konnten die Autoren der genannten Arbeit unter Annahme einer geeigneten nichtinformativen A-priori-Verteilung zeigen, daß der Bayes-Test auf $\ddot{A}Q$ von μ_T/μ_R mit 100% identisch ist mit einem Test nach dem Intervalleinschluß-Prinzip, was automatisch die gewünschte Niveaueigenschaft liefert. Weitere Beiträge zur Bayesschen Analyse vergleichender Bioverfügbarkeitsstudien findet man in [10] und [11].

4.3 Klassische Signifikanztests für Äquivalenzprobleme

Beim klassischen Ansatz verfährt man in der Weise, daß das Testproblem, soweit es sich auf eine mehrparametrische Verteilungsfamilie bezieht, zuerst durch Invarianzbetrachtungen oder Bedingen reduziert wird auf ein einparametrisches Problem. In der dann noch zu behandelnden einparametrischen Verteilungsfamilie läßt sich unter geeigneten Regularitätsbedingungen (Familie muß strikt totalpositiv 3. Ordnung sein, im Sinne von [5], pp.11-12) mittels des verallgemeinerten Fundamentallemmas ein gleichmäßig bester Test konstruieren ([6], p.120, Probl.30).

Wie anschaulich unmittelbar einleuchtet, haben solche direkt konstruierten Tests auf $\ddot{A}Q$ generell eine Gestalt, die zu derjenigen eines herkömmlichen zweiseitigen Tests komplementär ist. D.h., die kritische Region ist gegeben als

$$\{\mathbf{x} \in \mathcal{X} \mid C_1(\alpha) < T(\mathbf{x}) < C_2(\alpha)\} \quad .$$

Die Ablehnung der Nullhypothese der Nichtäquivalenz hat also genau dann zu erfolgen, wenn der beobachtete Wert einer geeigneten reellwertigen Teststatistik T innerhalb eines genügend kurzen Intervalls bleibt. Die Grenzen des kritischen Intervalls hängen, falls es sich um einen bedingten Test handelt, außer vom Signifikanzniveau vom Wert einer weiteren Statistik S ab, welche suffizient für den (evtl. mehrdimensionalen) Störparameter gewählt wird. Die Berechnung von $C_1(\alpha)$, $C_2(\alpha)$ hat gerade so zu erfolgen, daß die (bedingte) Ablehnungswahrscheinlichkeit sowohl unter $\theta = \theta_0 - \delta_1$ als auch unter $\theta = \theta_0 + \delta_2$ exakt gleich α ist.

In der Regel gestaltet sich die explizite numerische Berechnung optimaler Tests für $\ddot{A}Q$ -Probleme erheblich komplizierter als für die zugehörigen ein- oder zweiseitigen Probleme. Dies hat in erster Linie den Grund, daß man bei $\ddot{A}Q$ -Problemen unter der *Nullhypothese* die jeweilige Prüfverteilung in einer *nichtzentralen* Version benötigt. Außerdem entfällt bei bedingten Tests die Möglichkeit einer Reduktion auf nichtbedingte Prozeduren, welche bei ein- und zweiseitigen Problemen bekanntlich (vgl. [6], Sec.5.1) oft zu erheblichen Vereinfachungen führt. Trotzdem läßt sich ein optimaler Test auf $\ddot{A}Q$ in zahlreichen Situationen auf eine Form bringen, die bei guter Ausstattung des Anwenders mit statistischer Basissoftware den routinemäßigen Gebrauch des Verfahrens ohne weiteres zuläßt.

Daß die praktische Durchführung eines optimalen Tests auf $\ddot{A}Q$ sogar ausgesprochen einfach sein kann, findet man beispielsweise in der Situation des unverbundenen t -Tests bestätigt. Hier läßt sich zeigen [15], daß bei Verwendung von $(\mu_1 - \mu_2)/\sigma$ als Hauptparameter ein gleichmäßig bester invarianter Test auf $\ddot{A}Q$ existiert. Die Invarianzeigenschaft versteht sich hierbei bezüglich der Gruppe \mathcal{G} aller Transformationen der Form $(x_1, \dots, x_m, y_1, \dots, y_n) \rightarrow (a + bx_1, \dots, a + bx_m, a + by_1, \dots, a + by_n)$ mit $(a, b) \in \mathbb{R} \times \mathbb{R}_+$, unter denen jede Koordinate eines Punktes im \mathbb{R}^{m+n} derselben Streckung und derselben Translation unterzogen wird. Die Prüfgröße T dieses Tests ist die gewöhnliche Studentische t -Statistik. Bei symmetrischer Festlegung der $\ddot{A}Q$ -Grenzen, d.h. für das Problem

$H : |\mu_1 - \mu_2|/\sigma \geq \delta$ gegen $K : |\mu_1 - \mu_2|/\sigma < \delta$ hat der gleichmäßig beste invariante Test zum Niveau α die Ablehnungsregion

$$\{T^2 < F_{1,m+n-2,\alpha}(mn\delta^2/(m+n))\} \quad ,$$

mit

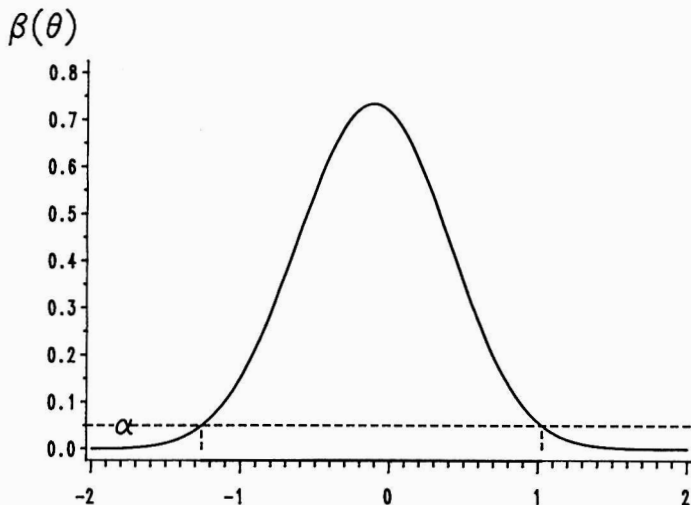
$F_{\nu_1,\nu_2,\alpha}(\psi)$:= unterer 100α -Prozentpunkt der nichtzentralen F -Verteilung
mit ν_1, ν_2 Freiheitsgraden und Nichtzentralitäts-Parameter ψ .

Numerisches Beispiel zum "Zweistichproben-t-Test auf ÄQ": In einer randomisierten zweiarmligen Therapiestudie sollten die Verteilungen der Körpergröße [cm] auf Homogenität geprüft werden. Für die $m = 178$ Patienten in Gruppe A erhielt man $\bar{X} = 169.96$, mit $S_X = 7.56$ als empirischer Standardabweichung In Gruppe B, welche $n = 175$ Patienten umfasste, ergab sich $\bar{Y} = 171.35$, $S_Y = 8.10$. Hieraus errechnet sich ein t -Wert von $T = -1.6671$, so daß man $T^2 = 2.7793$ erhält. Wählt man den ÄQ-Bereich symmetrisch, indem man unter der Alternativhypothese fordert, daß $|\mu_1 - \mu_2|/\sigma < .36$ (vgl. vorn, Abschn.3), erhält man auf dem 5%-Niveau als kritische Grenze für T^2 die Zahl 3.0089. Entsprechend kommt hier der gleichmäßig beste invariante Test zu einem positiven Resultat, d.h. er entscheidet für Äquivalenz.

Wird das Testproblem dahingehend abgeändert, daß der unter der Alternativhypothese spezifizierte ÄQ-Bereich asymmetrische Gestalt hat, benötigt man zur Berechnung der kritischen Konstanten des optimalen Tests eine Iterationsprozedur. Ein SAS-Macro-Programm, welches den entsprechenden Algorithmus implementiert, ist im Anhang zu [15] gelistet.

5 Trennschärfe und minimal erforderliche Fallzahlen

Die nachfolgende Graphik gibt die exakte Gütefunktion des gleichmäßig besten invarianten Tests auf ÄQ bezüglich $(\mu_1 - \mu_2)/\sigma$ für die Zweistichproben- t -Test-Situation wieder, mit $m = n = 12$, $\theta_0 = 0$, $\delta_1 = 1.2568$, $\delta_2 = 1.0269$, $\alpha = .05$. Man erkennt die geringgradige Unsymmetrie der Kurve, die im übrigen grob gesagt an eine Gauß-Dichte erinnert. Außerdem kann man ablesen, daß $\max_{\theta} \beta(\theta) \approx \beta(-.11) \approx .73$.



$$\theta = (\mu_1 - \mu_2) / \sigma$$

Abb.1: Gütefunktion des gleichm. besten invarianten Tests zum 5%–Niveau auf Äquivalenz von $\mathcal{N}(\mu_1, \sigma^2)$ und $\mathcal{N}(\mu_2, \sigma^2)$ bezüglich $\theta = (\mu_1 - \mu_2) / \sigma$. Theoretischer ÄQ-Bereich [Balken unten mitte]: $(-1.2568, 1.0269)$.

Allgemein erfüllt in einparametrischen Verteilungsfamilien die Gütefunktion $\theta \mapsto \beta(\theta)$ eines vernünftigen, im Sinne der Niveaueinhaltung gültigen Tests auf ÄQ stets die folgenden Bedingungen:

- (i) $\beta(\cdot)$ nimmt in einem im Innern des theoretischen ÄQ-Intervalls gelegenen Punkt θ_o^* ein strenges globales Maximum an.
- (ii) Auf $\{\theta \mid \theta \leq \theta_o^*\}$ bzw. $\{\theta \mid \theta \geq \theta_o^*\}$ nimmt $\beta(\cdot)$ streng monoton zu bzw. ab.
- (iii) Es gilt $\beta(\theta_o - \delta_1) \leq \alpha$, $\beta(\theta_o + \delta_2) \leq \alpha$.

Die für die Berechnung von Mindestfallzahlen relativ übersichtlichste Konstellation ist dann gegeben, wenn man die spezielle Alternative $\theta = \theta_0$ betrachtet (\rightarrow Nullhypothese beim zugehörigen zweiseitigen Testproblem). Mit etwas erhöhtem numerischen Aufwand ist es aber genauso gut möglich, unter irgendeiner anderen speziellen Alternative die Fallzahlen zu berechnen, die benötigt werden, um eine vorgegebene Trennschärfe zu erreichen.

Die eingehende Untersuchung von $\ddot{A}Q$ -Tests für die verschiedensten speziellen Situationen zeigt allerdings übereinstimmend, daß für die Wahl von Alternativen, unter denen die Fallzahlplanung erfolgen soll, nur eine sehr enge Umgebung von θ_0 in Betracht kommt. Denn bereits unter den bezüglich der Trennschärfe "günstigsten" Alternativen haben die Mindestfallzahlen für $\ddot{A}Q$ -Studien oft eine praktisch nicht zu realisierende Größenordnung.

Tab. 1: Ausgewählte Werte der Gütefunktion des gleichmäßig besten invarianten Tests auf $\ddot{A}Q$ bezüglich $(\mu_1 - \mu_2)/\sigma$ in der Situation des unverbundenen t-Tests mit gleichgroßen Stichproben des Umfangs n ; $\alpha = .05$, $\delta_1 = \delta_2 = \delta = .50$.

| n | $\beta(3\delta/4)$ | $\beta(\delta/2)$ | $\beta(\delta/4)$ | $\beta(0)$ |
|-----|--------------------|-------------------|-------------------|------------|
| 10 | 0.06564 | 0.07972 | 0.08957 | 0.09312 |
| 20 | 0.08564 | 0.12578 | 0.15840 | 0.17106 |
| 30 | 0.10920 | 0.19071 | 0.26645 | 0.29786 |
| 40 | 0.13223 | 0.26399 | 0.39915 | 0.45801 |
| 50 | 0.15134 | 0.32928 | 0.52149 | 0.60710 |
| 60 | 0.16735 | 0.38330 | 0.61878 | 0.72319 |
| 70 | 0.18184 | 0.42965 | 0.69400 | 0.80781 |
| 80 | 0.19563 | 0.47130 | 0.75282 | 0.86795 |
| 90 | 0.20903 | 0.50975 | 0.79956 | 0.91005 |
| 100 | 0.22216 | 0.54565 | 0.83719 | 0.93920 |

Wie der Tabelle zu entnehmen ist, trifft dies auch für optimale Tests zu, falls nicht das $\ddot{A}Q$ -Kriterium von vorneherein so schwach formuliert wird, daß die statistische Absicherung der zugehörigen Hypothese kaum noch von sachlichem Interesse ist. Die Tabelle bezieht sich auf das einfache Parallelgruppen-Design mit normalverteilten Fehlern (\rightarrow unverbundener t -Test mit gleichgroßen Stichproben). Die hierbei unter der Alternativ-

hypothese geforderte ÄQ-Bedingung kann in Anbetracht von

$$|\mu_1 - \mu_2|/\sigma < .50 \Leftrightarrow |P[X < Y] - 1/2| < .1382$$

kaum als extrem restriktiv bezeichnet werden. Trotzdem benötigt man selbst unter einer so nahe an θ_0 gelegenen Alternative wie $\theta = \delta/4 = .125 \Leftrightarrow P[X < Y] = .4648$ bereits mehr als 90 Fälle in jeder Gruppe, um eine Trennschärfe von 80% zu erreichen (\rightarrow Tab. 1, vorletzte Spalte).

6 Schlußfolgerungen für die Implementierung von Äquivalenztests in einem statistischen Expertensystem

- (i) Bei vorgegebenem Typ der zu planenden Studie sollte dem Anwender grundsätzlich eine Auswahl von verschiedenen Möglichkeiten der "Übersetzung" des Sachproblems in ein statistisches Testproblem angeboten werden. Z.B. sollte für die t -Test-Situation keinesfalls nur ein Test auf ÄQ bezüglich $\mu_1 - \mu_2$, sondern wenigstens auch bezüglich μ_1/μ_2 sowie $(\mu_1 - \mu_2)/\sigma$ zur Verfügung stehen.
- (ii) Es scheint weder sinnvoll noch praktikabel, bei der Auswahl der zu implementierenden ÄQ-Tests Einheitlichkeit hinsichtlich des Konstruktionsansatzes anzustreben. Zwar wird man sich naheliegenderweise für ein gegebenes Testproblem zuerst um eine optimale Lösung bemühen. Andererseits würde aber die Nichtzulassung von Intervalleinschluß- oder Bayes-Prozeduren zu einer für den Anwender unbefriedigenden Einengung des "Repertoires" an ÄQ-Problemen führen, deren Bearbeitung das System unterstützt.

Zu warnen ist übrigens vor der Fehleinschätzung, der Ansatz nach dem Intervalleinschluß-Prinzip sei grundsätzlich der einfachste. Beispielsweise wäre in der t -Test-Situation die Bestimmung von exakten Konfidenzgrenzen für $(\mu_1 - \mu_2)/\sigma$ keineswegs weniger aufwendig als die Berechnung der kritischen Konstanten für den gleichmäßig besten invarianten Test.

- (iii) Für die Berechnung der minimal erforderlichen Fallzahlen bietet sich ein zweistufiges Vorgehen an:
 - (a) Ermittlung des Fallzahlbedarfs unter $\theta = \theta_0$, d.h. unter der Nullhypothese

des zugehörigen zweiseitigen Problems. Liegt man bereits hier jenseits von realen Rekrutierbarkeitsgrenzen, erübrigt sich der nächste Schritt.

(b) Anwendern, die von den praktischen Randbedingungen der Studie her eine Erhöhung der unter $\theta = \theta_0$ benötigten Fallzahlen in Betracht ziehen können, sollte eine Option zur Rekalkulation der Fallzahlen unter einer beliebigen von θ_0 abweichenden Alternative angeboten werden.

Literatur

- [1] **APV – Arbeitsgemeinschaft für Pharmazeutische Verfahrenstechnik:** *Untersuchungen zur Bioverfügbarkeit, Bioäquivalenz.* Pharm. Ztg. **132**, 1952-1955 (1987)
- [2] **Box, G.P.E., and Tiao, G.C.:** *Bayesian Inference in Statistical Analysis.* Reading, Mass. 1973
- [3] **FDA:** *Code of Federal Regulations 21 (Food and Drugs), Part 320.22*
- [4] **Dunnett, C.W., and Gent, M.:** *Significance testing to establish equivalence between treatments, with special reference to data in the form of 2×2 tables.* Biometrics **33**, 593-602 (1977)
- [5] **Karlin, S.:** *Total Positivity. Volume I.* Stanford/Calif. 1968
- [6] **Lehmann, E.L.:** *Testing Statistical Hypotheses, 2nd Edition.* New York 1986
- [7] **Mandallaz, D., and Mau, J.:** *Comparison of different methods for decision-making in bioequivalence assessment.* Biometrics **37**, 213-222 (1981)
- [8] **Metzler, C.M.:** *Bioavailability – a problem in equivalence.* Biometrics **30**, 309-317 (1974)
- [9] **Schuirmann, D.J.:** *A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability.* J. Pharmacokin. Biopharm. **15**, 657-680 (1987)
- [10] **Selwyn, M.R., Dempster, A.P., and Hall, N.R.:** *A Bayesian approach to bioequivalence for the 2×2 changeover design.* Biometrics **37**, 11-21 (1981)
- [11] **Selwyn, M.R., and Hall, N.R.:** *On Bayesian methods for bioequivalence.* Biometrics **40**, 1103-1108 (1984)
- [12] **Steinijans, V.W., and Diletti, E.:** *Statistical analysis of bioavailability studies: Parametric and nonparametric confidence intervals.* Eur. J. Clin. Pharmacol. **24**, 127-136 (1983)

- [13] Wellek, S.: *Optimale Tests für Probleme der Form $H_0 : \theta \notin (\theta'_0, \theta''_0)$ gegen $H_1 : \theta'_0 < \theta < \theta''_0$ als methodische Alternative zu Äquivalenznachweis-Prozeduren vom Westlake-Typ.* **34.** Biometr. Kolloquium, Bad Nauheim, 21.-25.03.1988
- [14] Wellek, S.: *Vorschläge zur Reformulierung der statistischen Definition von Bioäquivalenz.* In G. Giani, R. Repges (Hg.), *Biometrie und Informatik – neue Wege zur Erkenntnisgewinnung in der Medizin*, 95-99. Berlin 1990
- [15] Wellek, S.: *Zur Formulierung und optimalen Lösung des Bioäquivalenznachweis-Problems in der klassischen Theorie des Hypothesentestens.* In J. Mau, H. Klinger (Hg.), *Bioäquivalenz sofort freisetzender Arzneiformen*, ***-***. Stuttgart 1990.
- [16] Wellek, S.: *Testing for equivalence with binomial data.* In Vorbereitung. (1991)
- [17] Westlake, W.J.: *Use of confidence intervals in analysis of comparative bioavailability trials.* *J. Pharm. Sc.* **61**, 1340-1341 (1972)
- [18] Westlake, W.J.: *Symmetrical confidence intervals for bioequivalence trials.* *Biometrics* **32**, 741-744 (1976)
- [19] Westlake, W.J.: *Statistical aspects of comparative bioavailability trials.* *Biometrics* **35**, 273-280 (1979)
- [20] Westlake, W.J.: *Design and statistical evaluation of bioequivalence studies in man.* In J. Blanchard (Ed.), *Principles and Perspectives in Drug Bioavailability*, 192-210. Basel 1979
- [21] Westlake, W.J.: *Response to T.B.L. Kirkwood: bioequivalence testing – a need to rethink.* *Biometrics* **37**, 591-593.

received: 15. November 1990

Verfasser:

Dr. S. Wellek
Institut für Med. Statistik u. Dokumentation
der Universität Mainz
Langenbeckstr. 1
W-6500 Mainz 1
Germany

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

Insbesondere beachte man:

1. Manuskripte sollten grundsätzlich **maschinengeschrieben** (Schreibmaschine, Drucker) in **deutscher oder englischer Sprache** abgefaßt sein.
2. Zur inhaltlichen Einordnung der Arbeit sind **1–2 Klassifizierungsnummern** (entsprechend der "1980 Mathematics Subject Classification" der Mathematical Reviews) anzugeben.
3. **Textbreite/Texthöhe** des Manuskripts sollten sich an den Maßen **160mm/230mm** orientieren.
4. Der Manuskripttext ist **eineinhalbzeilig, linksbündig**, wenn möglich, links- und rechtsbündig zu schreiben. Beim Auftreten von Formeln im laufenden Text ist der Zeilenabstand entsprechend zu vergrößern.
5. Der Platz für **Abbildungen** sollte beim Schreiben ausgespart werden; die Abbildungen selbst sind in der dem ausgesparten Platz entsprechenden Größe gesondert beizufügen.
6. **Literaturzitate** sind im Text durch laufende Nummern (vgl. [3], [4] ; [7, 8, 10]) zu kennzeichnen und am Schluß der Arbeit unter der Zwischenüberschrift **Literatur** bzw. **References** zusammenzustellen. Hierbei ist die durch die nachfolgende Beispiele veranschaulichte Form einzuhalten (die Zeitschriftenabkürzungen erfolgen nach Mathematical Reviews).
[3] **Zariski, O., and Samuel, P.:** *Commutative Algebra*. Princeton 1958
[4] **Steinitz, E.:** *Algebraische Theorie der Körper*. J. Reine Angew. Math. **137**, 167–309 (1920)
[8] **Gnedenko, B.W.:** *Über die Arbeiten von C.F. Gauß zur Wahrscheinlichkeitsrechnung*. In: Reichard, H. (Ed.): C.F. Gauß, Gedenkband anlässlich des 100. Todestages. S. 193–204, Leipzig 1967

Die Angaben erfolgen in Originalsprache ; bei kyrillischen Buchstaben sollte die bibliothekarische Transkription bzw. eine Übersetzung lt. Mathematical Reviews verwendet werden.

7. Die aktuelle, vollständige Adresse des Verfassers sollte enthalten: Titel / Vornamen
Name / Institution / Struktureinheit / Straße Hausnummer / Postleitzahl Ort / Land.

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

Schriftenreihen der Universität Rostock

| | |
|--|----------------|
| -- Archiv der Freunde der Naturgeschichte in Mecklenburg | ISSN 0518-3189 |
| -- Rostocker Agrarwissenschaftliche Beiträge | ISSN 0138-3299 |
| -- Rostocker Betriebswirtschaftliche Manuskripte | ISSN 0232-3066 |
| -- Rostocker Mathematisches Kolloquium | ISSN 0138-3248 |
| -- Rostocker Philosophische Manuskripte | ISSN 0557-3599 |
| -- Rostocker Physikalische Manuskripte | ISSN 0138-3140 |
| -- Rostocker Wissenschaftshistorische Manuskripte | ISSN 0138-3191 |
| -- Lateinamerika / Semesterberichte der Sektion Lateinamerikawissenschaften | ISSN 0458-7944 |
| -- Erziehungswissenschaftliche Beiträge | ISSN 0138-2373 |
| -- Migrationsforschung | ISSN 0863-1735 |
| -- Beiträge zur Geschichte der Universität Rostock | ISSN 0232-539X |
| -- Rostocker Beiträge zur Hoch- und Fachschulpädagogik | ISSN 0232-0539 |
| -- Rostocker Informatik-Berichte | ISSN 0233-0784 |
| -- Studien zur Geschichte der deutsch-polnischen Beziehungen | ISSN 0233-0687 |
| -- Rostocker Forschungen zur Sprach- und Literaturwissenschaft | ISSN 0233-0644 |
| -- Manuskripte zur Rostocker Universitätsgeschichte | ISSN 0863-1727 |
| -- Rostocker Universitätsreden | |
| -- Agrargeschichte | ISSN 0863-2170 |
| -- Jugendgeschichte | ISSN 0863-2391 |
| -- Leitungswissenschaftliche Manuskripte | ISSN 0863-2405 |

Bezugsmöglichkeiten

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

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